LEARNING DISABILITIES

From Identification to Intervention

SECOND EDITION

Jack M. Fletcher, G. Reid Lyon, Lynn S. Fuchs, and Marcia A. Barnes

LEARNING DISABILITIES

From Identification to Intervention

SECOND EDITION

Jack M. Fletcher
G. Reid Lyon
Lynn S. Fuchs
Marcia A. Barnes

THE GUILFORD PRESS
New York London

Epub Edition ISBN: 9781462536412; Kindle Edition ISBN: 9781462536399

A Division of Guilford Publications, Inc.
370 Seventh Avenue, Suite 1200, New York, NY 10001
www.guilford.com

Last digit is print number: 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data is available from the publisher.

ISBN 978-1-4625-3637-5

To our spouses—
Patricia McEnery, Diane Lyon,
Douglas Fuchs, and Mark Drummond—
for many years of love and support

About the Authors

Jack M. Fletcher, PhD, is the Hugh Roy and Lillie Cranz Cullen Distinguished University Professor of Psychology at the University of Houston. Since the 1970s, Dr. Fletcher, a child neuropsychologist, has completed research on many issues related to learning disabilities, including definition and classification, neurobiological correlates, and intervention, and has written over 400 articles in peer-reviewed journals. He is Principal Investigator of the Texas Center for Learning Disabilities, has served on and chaired the Mental Retardation/Developmental Disabilities study section of the National Institute of Child Health and Human Development (NICHD), and is a former member of the NICHD National Advisory Council and the President's Commission on Excellence in Special Education. He is a recipient of the Samuel Torrey Orton Award from the International Dyslexia Association and the Albert J. Harris Award from the International Literacy Association. Dr. Fletcher is a past president of the International Neuropsychological Society.

G. Reid Lyon, PhD, is Distinguished Scientist Emeritus in Neuroscience and Cognition at the Center for Brain Health, The University of Texas at Dallas. He is also Distinguished Professor Emeritus at Southern Methodist University where he served as Associate Dean of the Annette Caldwell Simmons School of Education and Human Development and Chair of the Department of Education Policy and Leadership. He served on the Maternal and Child Health study section of the NICHD and chaired the NICHD Learning Disability Research Centers Reading Development and Disorders study sections. Dr. Lyon has published over 135 articles, books, and book chapters addressing dyslexia, classification and definition of learning disabilities, developmental neuroimaging, neurophysiological correlates of reading interventions, and the development of gist reasoning in adolescents. He is a recipient of the Director's Award for Scientific Leadership in Neuroscience and Learning Disabilities from the National Institutes of Health and of the Norman Geschwind Memorial Lecture and the Samuel Torrey Orton Award from the International Dyslexia Association.

Lynn S. Fuchs, PhD, is the Dunn Family Endowed Chair of Psychoeducational Assessment, Special Education, and Human Development at Vanderbilt University. She is among the most highly cited researchers in the social sciences (according to Thomson Reuters); has published more than 350 articles in peer-reviewed journals; and serves on the editorial boards of many journals, including the *Journal of Educational Psychology, Scientific Studies of Reading, Elementary School Journal, Journal of Learning Disabilities,* and *Exceptional Children*. Her research focuses on classroom-based assessment as well as instructional methods for students with reading disabilities and math disabilities. In addition, Dr. Fuchs has conducted programmatic research on assessment methods for enhancing instructional planning and on instructional methods for improving reading and math outcomes for students with learning disabilities.

Marcia A. Barnes, PhD, is Professor in the Department of Special Education at Vanderbilt University. Previously, she was Professor & H. E. Hartfelder/Southland Corp Regents Chair in the Department of Special Education and Associate Dean for Research and Graduate Studies in the College of Education at The University of Texas at Austin. Her research focuses on math and reading comprehension disabilities in children with and without brain injuries. She also studies the typical development of reading comprehension skills, and has created several paradigms for assessing inference-making in children. Dr. Barnes has published 120 peer-reviewed papers, chapters, and books. She is a member of the editorial board of the *Journal of Learning Disabilities* and the *Journal of Educational Psychology* and has served on several national grant review panels in the United States and Canada.

Preface

In an era of increased focus on the evidence base that supports different education practices, the second edition of this book integrates multiple domains of scientific inquiry, practice, and policy involving learning disabilities (LDs). Representing several disciplines in psychology, neuroscience, genetics, and education, the book is an exposition and analysis of the scientific research base that has accumulated over the past 50 years on LDs, ranging from identification and assessment, to cognitive and neurobiological factors, to intervention, and to translation of research into practice and policy. The heart of the book is its focus on research in different domains of LDs involving reading (word recognition and comprehension), mathematics (calculations and problem solving), written language (handwriting, spelling, and composition), and the more general problem of automaticity. A clear link is made with what is known about the typical development of these skills, the manifestation of these skills as LDs, and how to support development of these skills via classroom instruction and specific interventions.

An understanding of LDs must stem from a classification framework that leads to definitions of and methods for identifying LDs that epitomize the historically central construct of *unexpected underachievement*. Based on the classification, specific LDs can be identified according to their core academic deficits, instructional response, and consideration of other disorders and contextual factors that impact treatment. This classification framework provides the capacity for systematically studying the neurobiological and environmental factors that produce LDs. Although the book's main focus is research, it also addresses practice and policy, with considerable attention paid to assessment and intervention methods that have demonstrable efficacy in each domain of LDs.

Our interest in writing the first edition of the book was stimulated in part by recognition of the major changes in U.S. public policy involving education, beginning with the focus on scientifically based instruction in the reauthorization of the Elementary and Secondary Education Act (ESEA), through the No Child Left Behind Act of 2001, and continuing with the 2004 Reauthorization of the Individuals with Disabilities Education Act (IDEA 2004). For the first time since the initial legislation supporting IDEA in 1975, IDEA 2004 allows the U.S. public education system to examine new approaches to identifying and treating LDs under the general rubric of response-to-intervention (RTI) methods and specific expectations for appropriate instruction in general education as a prerequisite to identifying LDs as *unexpected underachievement*. In the second edition, we note the continued expansion of layered approaches to service delivery as a multi-tiered system of supports (MTSS) framework in the 2015 reauthorization of the ESEA.

Although RTI methods can be used to help identify LDs, they are complementary to the MTSS framework with the goal of enhancing education outcomes for children through closer coordination between general and special education and other entitlement programs. A frequently asked question about these frameworks is whether the assessment and intervention methods are sufficiently developed. We review much of this research, identify gaps in the knowledge base, and conclude that, although some issues require additional scientific inquiry, a substantial research base exists and many of the issues regarding RTI methods and MTSS frameworks represent not an absence of assessment and intervention tools, but rather the need to scale them. In the second edition, scaling and implementing this body of research looms even larger given the policy events of the past 10 years (see Chapter 11). In the second edition, we have worked to make the intervention components more accessible while preserving the focus on evidence-based empirical research.

We hope this second edition facilitates the capacity of educators and schools to identify sound tools for assessment and instruction and to implement them in the service of better outcomes for students at risk for or with identified LDs. We believe that the research incorporated in this book shows that LDs are real, that the field does have a strong scientific basis, and that the development of the field continues in a positive direction and will continue to flourish. Most important, the robust instructional methods for each of the specific LDs described in this book reflect the accumulation of substantial scientific information on LDs that can be used to inform practice and policy.

This second edition includes new chapters on the reality of LDs, principles of intervention, the general problem of automaticity, and the problem of scaling and translating research. It thoroughly updates the research findings across multiple sources of scientific evidence. Particularly impressive is the research base on intervention in the five major domains of LDs proposed in our academic classification framework. The result is a single volume that integrates research on classification and definition, cognitive processing, neurobiological factors, and instruction.

Acknowledgments

We thank Rochelle Serwator, our editor at The Guilford Press, who proposed both editions of this book. We also thank Connie Monroe for many hours of support in completing the book and Whitney Roper for work on the figures. Drs. Catherine Jockell and Ashley Ware also helped with early chapter drafts.

Grants from the National Institute of Child Health and Human Development (NICHD) to Jack M. Fletcher helped support some of the research in the book, including P50 HD052117, Texas Center for Learning Disabilities (which also supported research by Marcia A. Barnes), and P01 HD46261, Cognitive, Instructional, and Neuroimaging Factors in Math (which also supported research by Lynn S. Fuchs and Marcia A. Barnes). Additional support for Dr. Fletcher was obtained from Grants NSF9979968, Early Development of Reading Skills: A Cognitive Neuroscience Approach, funded under the Interagency Educational Research Initiative by the National Science Foundation, the NICHD, and the Institute of Education Sciences (Jack M. Fletcher, Principal Investigator); P50 HD25802, Center for Learning and Attention Disorders (Sally E. Shaywitz and Bennett A. Shaywitz, Principal Investigators); and R01 HD38346, Brain Activation Profiles in Dyslexia (Andrew C. Papanicolaou, Principal Investigator).

Dr. Lyon's research was supported in part by Grant H327A090016, The Efficacy of Computer Assisted Tier 2 Assessment and Intervention, obtained from the Department of Education, Office of Special Education and Rehabilitation Services (Coprincipal Investigator with Patricia Mathes). Dr. Lyon's research was also supported by The Gunderson Medical Foundation, Definition and Classification of Dyslexia (G. Reid Lyon/Jane Flynn

Coprincipal Investigators).

Dr. Fuchs's work was supported by Grants R01 HD46154, Understanding/Preventing Math Problem-Solving Disability (Lynn S. Fuchs, Principal Investigator); R01 HD056109 RTI, Determining Risk, Providing Prevention, and Identifying Reading Disability (Douglas Fuchs, Principal Investigator); R01 HD059179, Calculations, Word Problems, and Algebraic Cognition (Lynn S. Fuchs, Principal Investigator); P20 HD075443, Word Problems, Language, and Comorbid Learning Disability (Lynn S. Fuchs, Principal Investigator); and R01 HD053714, Preventing and Understanding Mathematics Disability (Lynn S. Fuchs, Principal Investigator) from the NICHD. She was also supported by Grants R324C100004, Improving Understanding of Fractions among Students with Mathematical Learning Difficulties (Nancy Jordan, Principal Investigator); R305A110067, Arithmetical and Cognitive Antecedents and Concomitants of Algebraic Skill (Paul Cirino, Principal Investigator); and R324D130003, Improving Reading and Mathematics Outcomes for Students with Learning Disabilities: Next Generation Intensive Interventions (Douglas Fuchs, Principal Investigator) from the National Center of Special Education Research in the Institute of Education Sciences of the U.S. Department of Education. Further support was furnished by Grants H324V980001, Center on Accelerating Student Learning (Lynn S. Fuchs, Principal Investigator) and H324U010004, National Research Center on Learning Disabilities (Douglas Fuchs, Principal Investigator) from the Office of Special Education Programs in the U.S. Department of Education.

Dr. Barnes's work was supported by grants from the National Center of Special Education Research and the National Center for Education Research at the Institute of Education Sciences of the U.S. Department of Education, including R324A110270, A Randomized Trial of a Tutor-Based Mathematics and Attention Intervention for Low-Performing Preschoolers at Risk for Mathematical Difficulties in School (Marcia A. Barnes, Principal Investigator); R324A160052, Project Connect-IT (Connecting Text by Inference and Technology): Development of a Text-Integration Intervention for Middle School Students with Comprehension Difficulties (Marcia A. Barnes, Principal Investigator); and R305F100013, Understanding Malleable Cognitive Processes and Integrated Comprehension Interventions for Grades

7–12 (Sharon Vaughn, Principal Investigator).

The contents of this book are solely the responsibility of the authors and do not necessarily represent the official views of any of the sources of grant support.

Contents

Purchasers of this book can download and print the text and references of Chapter 2, "History of the Field," from the first edition of *Learning Disabilities* at *www.guilford.com/fletcher-materials* for personal use.

CHAPTER 1

Introduction

Since the federal designation of learning disabilities (LDs) as a "handicapping condition" in 1968 in the United States, the proportion of children identified with LDs increased steadily until the past decade. At its peak, students with LDs represented almost one-half of all children receiving special education services (U.S. Department of Education, 1999). But from 2002 to 2011, the number of children in special education with LDs declined about 2% per year, or a total of 18%, although the number of students identified for special education declined only 3% (National Center for Learning Disabilities, 2014). These figures have stabilized through 2016 to about 35% of children served in special education. Although autism and "other health impaired" (OHI) are now the fastest growing eligibility categories, partly because of the explicit inclusion of children with attention-deficit/hyperactivity disorder (ADHD) in the OHI category, students with LDs are still the largest group, representing about 4.6% of all students in the U.S. public education system (National Center for Education Statistics, 2017).

While there was relatively little research on LDs at the time that the U.S. federal special education legislation was initially enacted in 1975, significant progress has been made in understanding and treating LDs involving reading, mathematics, and written expression since then. As we noted in the first edition of this book (Fletcher, Lyon, Fuchs, & Barnes, 2007), major advances had been made in classification and definition issues, cognitive processes, neurobiological correlates involving the brain and genetics, assessment practices, and intervention. Lyon and Weiser (2013) provided additional coverage of advances from 2007 to 2011, including an analysis of the scientific quality of these advances. Much of this progress was in areas related to word reading, or dyslexia (see Chapter 6), especially in younger children because of a research emphasis on early identification and prevention.

Since 2007, the word-reading area has expanded across the lifespan and considerable progress has been made in domains related to reading comprehension, math, and written expression (Lyon & Weiser, 2013). The advances in intervention are especially promising. Although research has shown that reading and math disabilities are preventable in many children, it is now apparent that there are both preventative and remedial interventions in all the five domains of LDs reviewed later in this book (word reading, reading comprehension, math computations, math problem solving, and written expression). Service delivery models based on response to intervention (RTI), now more generally termed "multitiered systems of support" (MTSS), have emerged as schoolwide approaches to instruction and intervention. These approaches are also sources of controversy, especially when the identification of students with LDs is considered (Reynolds & Shaywitz, 2009).

Knowledge about neurobiological factors underlying reading, math, and writing disabilities has been consolidated and more is known about the intrinsic link of genetic factors that put the brain at risk for LDs. Environmental factors that provide the context through which LDs are expressed, such as instruction and the home literacy and language environment, can increase or reduce risk for these LDs. Knowledge of neurobiological correlates is not to the point where it can or should affect instruction, but is important for informing theory and understanding of LDs. The impact on instruction, especially the need for *explicit* approaches for children who are struggling, is very apparent when neuroscience research is evaluated. The neural systems that mediate reading and math skills develop through instruction and experience, which must be explicit for many children if these systems are going to emerge.

In the first edition of this book, we observed that a comprehensive model had emerged for word-level reading difficulties (dyslexia), the most common LD, which is grounded in reading development theory and accounts for neurobiological and environmental factors in addition to the effects of intervention (Pennington, 2009; Elliott & Grigorenko, 2014). We reiterate that the same theory of reading development that explains how children acquire reading skills explains why some fail, unifying the research on LDs in reading and the normative development of reading ability, and making accounts of LDs more compelling. This appears true for other LDs. The defining attributes of LDs (e.g., low achievement, inadequate instructional response) appear normally distributed in the population and there is little evidence of qualitative variation that would suggest categories, much less where LDs begin in relation to typical development. Such decisions are often resource-driven.

Despite this scientific progress, the construct of LDs and the many definitions that serve as conceptual frameworks for their identification and treatment continue to be misunderstood. The field continues to be plagued by pervasive disagreements about the definition of LDs, diagnostic criteria, assessment practices, treatment procedures, and educational policies. The translation of scientific progress into classrooms remains difficult (Chapter 11), and anecdotes and older belief systems continue to prevail. If anything, there is less emphasis on the use of science as a basis in 2018 than there was in 2007, a heady time for scientists investigating LDs.

WHY A SECOND EDITION?

In writing a second edition, we aimed to continue to integrate the disparate sources of information into a more coherent account of LDs, beginning with an evidence-based approach to definition and classification (Chapter 3) and the implications of what we describe as a hybrid approach for assessment and identification (Chapter 4). With an adequate classification, it becomes possible to comprehensively discuss research on the nature, types, causes, and treatment of LDs (Chapters 5–10), thus beginning to integrate science and practice (Chapter 11).

This second edition also addresses the horizontal integration of knowledge on LDs, providing less depth within different domains of knowledge in favor of the connections across these domains and the boundaries across disciplines. Because science has advanced, there is a need to revise and update this account. In addition, because of the difficulties with implementation of this scientific knowledge, we hope to provide a clearer set of principles about how to implement scientific knowledge in relation to instruction (Chapter 5), with an emphasis on examining the converging evidence in support of different instructional practices in Chapters 6–10. Instead of detailed, systematic reviews of the literature, we tried to focus even more on general principles that have emerged and to provide more concrete, practical guidelines to facilitate intervention. Hence, we have altered the book by dropping the chapter on history with the exception of recent updates in Chapter 2, which is a new chapter addressing issues related to the validity of the LD construct. The history chapter from the first edition is available online (see the box at the end of the table of contents). Thinking about the simple question of whether LDs represent "real" entities is important as policymakers among others struggle with resource issues and ideologies that interfere with implementation of the intensive interventions needed by many individuals with LDs. Chapter 3 is an updated chapter that focuses on identification issues, illustrating how problems identifying individual people with LDs underlie any attempt to categorize inherently normal distributions (i.e., achievement, instructional response, cognitive functions) regardless of the assessment method employed. In addition, Chapter 3 updates the research on identification methods, specifically questioning the reliability and validity of approaches that focus on assessment of students in isolation of instructional response.

Chapter 4 updates assessment approaches, especially in the context of MTSS methods of service delivery. For clarification, we will refer to RTI when we discuss identification methods and to MTSS as a broader service delivery model consistent with the reauthorization of the U.S. Elementary and Secondary Education Act in 2015, also called the Every Student Succeeds Act (*www.ed.gov/esea*). Chapter 5 is new, focusing on principles of instructional design for people with LDs. Chapters 6–9 focus on LDs involving word reading and spelling (dyslexia), specific reading comprehension, math computations (dyscalculia) and problem solving, and written expression. All four chapters have been extensively updated, reflecting the amount of research on LDs in the past decade. Whenever possible, we refer to meta-analyses (quantitative syntheses) of research and use individual studies to illustrate interesting findings and effective interventions.

We no longer discuss reading fluency as a separate LD, but instead focus on the more general issue of automaticity in reading, math, and writing in the new Chapter 10. The final chapter (Chapter 11) discusses issues related to the difficulties with implementation and scaling of scientific knowledge from contemporary and historical perspectives, with an eye toward lessons learned over the past decade.

AN OVERARCHING FRAMEWORK

Figure 1.1, which was introduced in the first edition, presents a framework for understanding the different sources of variability that influence academic outcomes in children with LDs. The framework encompasses three levels of analysis that underlie an integrated account of LDs and is anchored in a hypothetical classification of LDs based on variations in academic skills. Evidence suggests five major prototypes of LDs involving word recognition (and spelling), reading comprehension, mathematics computations, mathematics problem solving, and written expression. These domains have been selected both because of their prominence in current definitions of LDs, and because most children and adults are identified as having LDs manifest unexpected underachievement or atypical development in one or more of these areas.

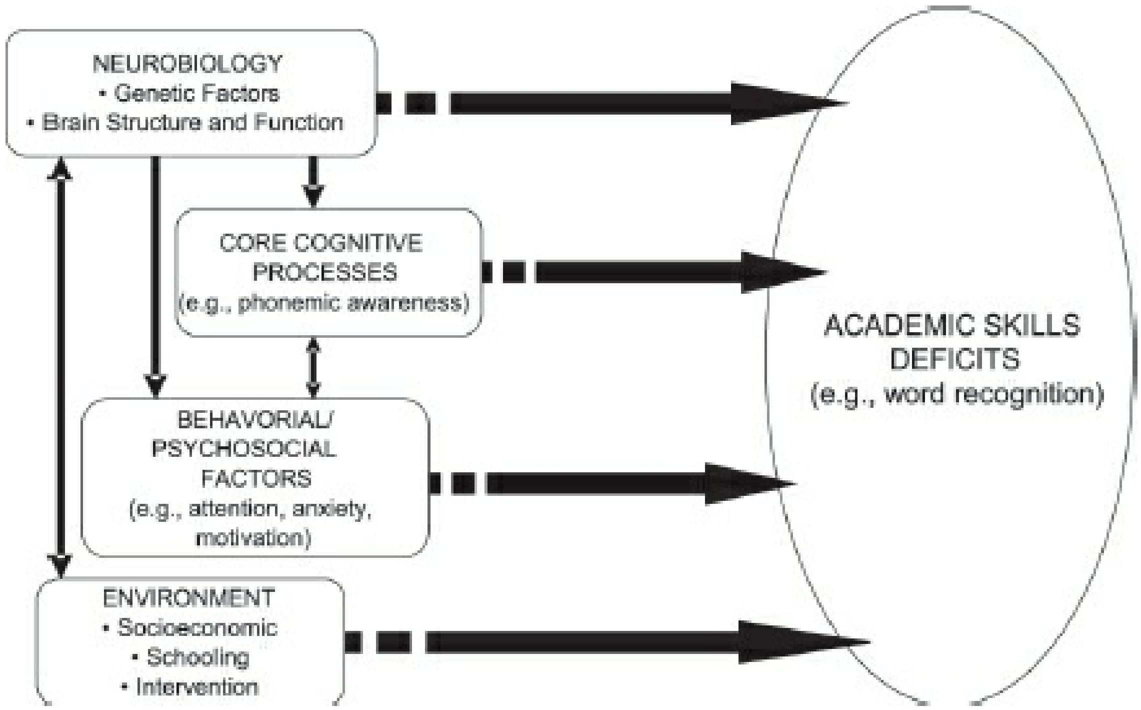

FIGURE 1.1. Framework representing different sources of variability that influence academic outcomes, the primary manifestations of the disability, in children with LDs.

For each LD, the primary manifestation of the disorder represents specific academic skill deficits in the five domains of LDs. By referring to these domains as "disabilities," we use historically established language, but would add that what makes LDs a disability rather than a disorder or a deviation from normal development is (1) the severity of underachievement, which is unexpected because the individual has not responded adequately to instruction that is effective for most individuals; and (2) the evidence of adaptive impairment, such as poor school achievement. Thus, disability determination is always a two-pronged determination based on the existence of a problem and evidence of adaptive impairment, the latter representing the weakest part of most definitions of LDs (see Chapter 2).

The second level of analysis involves person-level characteristics, including core cognitive processes (e.g., phonological awareness and vocabulary) that are correlated with the academic skill deficits (e.g., word recognition skills and reading comprehension) in addition to academic strengths. Reading, math, and writing are also complex cognitive skills that represent the manifestations of other cognitive skills, but separating academic and core cognitive skills is useful for assessment and intervention purposes.

Academic strengths and weaknesses are also influenced by a second set of person-level characteristics encompassed in the psychosocial domain, such as motivation, social skills, and behavioral problems involving anxiety, depression, and/or inattention that interfere with performance in academic domains. The arrow between core cognitive processes and behavioral/psychosocial factors is bidirectional because cognitive difficulties can also lead to problems with, for example, attention and social skills, which can in turn influence academic abilities. Neither type of person-level characteristics (i.e., cognitive and behavioral/psychosocial factors) should be considered diagnostic of LDs, although the psychosocial/contextual component and the possibility of other co-occurring disorders must be evaluated in order to plan intervention. The need to evaluate cognitive characteristics in isolation of academic skills is controversial and we argue for direct assessments of academic skills and psychosocial components because of the absence of evidence that assessment of cognitive skills adds value to intervention (Chapter 3) and the lack of evidence that interventions based on cognitive skills generalize to academic skills (Fletcher & Miciak, 2017; Mann, 1979).

The third level of analysis involves neurobiological and environmental factors. Neurobiological factors include genetic and neural sources of variability that impact academic skill deficits either indirectly through their influence on person-level characteristics or directly on attainment of the academic skills. Environmental factors are contextual and include the social and economic circumstances in which a person develops and functions, as well as schooling influences, such as the quality of the school and different interventions. The arrow linking neurobiological and environmental factors is bidirectional, indicating the synergistic influence of these domains. Although the idea that neurobiological factors lead to LDs is not new, it is important to recognize that instruction and experience reorganize the neural systems involved in LDs and influence the expression of biological factors. In an integrated account of LDs, all three levels of analysis must be considered. As in the first edition, we focus on the relations of academic skills with core cognitive processes, neurobiological factors, and intervention.

Historically, research on LDs has emphasized the second (and third) levels of the framework as opposed to the first level of analysis. Although

Figure 1.1 includes multiple levels of analysis, a strong classification is based on a parsimonious set of markers that identify members into the different parts of the classification. Our discussion of academic skill deficits attempts to identify these markers, which should predict the cognitive and neurobiological factors. There are important relations with the psychosocial and environmental variables that are essential for understanding the impact of intervention. Thus, adequate identification of valid markers and the effectiveness of interventions require a focus on achievement, instructional response, and other factors that impact the development of academic skills. These latter factors are typically used to exclude people from LD classifications. However, without a focus on these factors, many children will be identified as LD for whom the explanation of the disability is *poor instruction* and not *unexpected underachievement.*

The strengths and weaknesses in cognitive skills that some view as essential to the nature of LDs (e.g., phonological awareness, working memory) can be accounted for simply by assessing the achievement domains (e.g., word recognition). Over the past decade, little evidence has emerged showing that cognitive skill assessments contribute significant value-added information to predictions of academic outcomes (Stuebing et al., 2015) or to treatment planning (Kearns & Fuchs, 2013), although working memory and oral language remain viable candidates (Peng & Fuchs, 2016; L. S. Fuchs et al., 2014b). This does not mean that cognitive skills are not related to LDs or that research might one day identify a role for assessment and intervention with cognitive skills, but it has yet to emerge (Schneider & Kaufman, 2017). Regardless, routine assessment of cognitive skills is not indicated, just as the impressive research base on neuroimaging does not suggest a need for brain scans of each child suspected of LDs. The neural correlates are predicted by the tasks used to elicit brain activation (word reading, math calculation, etc.), which should also predict the correlated cognitive processes, again demonstrating the major role of levels of achievement in the prediction and identification of LDs. The ability to make these predictions and simplify classification, identification, and assessment processes signal the emergence of an evidence-based approach for classifying LDs, with simple decision rules focused on direct assessment of key academic skills that leads to the rapid provision of effective interventions, which is the goal of identification.

From our perspective, the future of LDs is tied to the scientific process, and the field must embrace the evolving process of scientific research and move away from poorly verified clinical intuition and slick marketing in order to provide a solid foundation for practice (Chapter 11). In many respects, this is more of a problem today than in 2007 and we are concerned that the field is regressing vis-à-vis a reemergence of reliance on untested assumptions and superstition in identification, intervention, and remediation practices. Clinical experience is a fertile ground for hypothesis generation, but the inferences that emerge from experience must be empirically verified, particularly in identification practices and intervention. The issue remains: For whom do different factors converge to cause LDs, and how do different components of intervention relate to the various expressions of LDs?

CAVEATS

This edition has similar caveats to the first edition. We present a particular approach to understanding LDs, which is based on a classification with its roots in academic achievement and which we use to account for the heterogeneity of LDs. Academic deficits are necessary, but not sufficient, for a classification of LDs. However, without achievement as an anchor, it is difficult to validate the construct of LDs.

Accordingly, we do not review research on students broadly defined with LDs when the specific form of academic impairment is not indicated, unless that approach predominates in the instructional literature. In the absence of this type of specification, the samples included in such studies are too heterogeneous to determine valid relations with specific forms of LDs. Likewise, we do not review research suggesting that LDs involving social or executive functions should be separately identified because we do not feel that such approaches to identification result in effective classifications of LDs. Although we recognize that other approaches to defining "verbal" and "nonverbal" LDs have represented major contributions to the field (e.g., Johnson & Myklebust, 1967; Rourke, 1989), we do not explicitly organize our approach around this dichotomy for definition and classification. The reader is encouraged to examine these approaches, such as the approach to the definition of "verbal" and "nonverbal" LDs developed by Rourke and colleagues (see *www.nld-bprourke.ca/index.html*) and addressed most recently by Cornoldi, Mammarella, and Fine (2016). There are major issues regarding the hypothesis of nonverbal LDs (Pennington, 2009; Spreen, 2011). These include specific diagnostic criteria, the fact that academic problems are not considered a defining characteristic, whether the characteristics are better accounted for by classifications stemming from ADHD or autism spectrum disorder, and the role of social skills. Etiological hypotheses involving differences in hemispheric distribution of white matter or problems involving the right hemisphere have not found consistent support. Renaming nonverbal LDs as right-hemisphere LDs or as visuospatial LDs seems to confuse the behavioral description with hypotheses about etiology. More research would be useful, but it is not a focus of our book and does not fit into our framework for understanding LDs.

Given the enormous volume and complexity of the literature on topics associated with treatment of and instruction for LDs, our review of relevant research is selective rather than exhaustive. It was not possible to systematically address research related to ADHD or to social and emotional difficulties—areas of development that are clearly problematic for many students with LDs. These influences are usually *comorbid,* representing frequently co-occurring difficulties as opposed to qualitatively disparate disorders. In terms of Figure 1.1, we do not provide an extensive discussion of the psychosocial and behavioral factors or a broad assessment of environmental factors (e.g., poverty) that impact the development of children with LDs (for a review, see Phillips & Lonigan, 2005). This is partly because there is little evidence that the phenotypic manifestations of academic difficulties vary by putative cause. We focus instead on intervention.

In our analysis of the literature, most psychosocial and environmental influences contribute to the severity of academic achievement problems, but do not produce qualitative variation; hence the importance of instructional response in operationalizing unexpected underachievement. Although various theoretical and conceptual models related to treatment are implicit in our review of interventions, as are specific intervention methods, we do not view the work emanating from these different sources and perspectives as necessarily contradictory and do not discuss these models in detail. Rather, thoughtful integration of these models is resulting in more efficacious interventions for individuals with different types of LDs. Academic therapies that involve substantial exposure to reading, mathematics, and writing are most effective; other approaches to interventions that teach cognitive or motor processes, train the brain, or focus on aspects of the disorder (e.g., vision) that are not directly tied to the academic skill do not result in improved outcomes for students with LDs. Further, the literature is replete with claims for instructional and treatment methods that are based on subjective, nonreplicated clinical reports, testimonial information, and anecdotal statements on groups broadly defined with LDs. We have limited our discussion to empirical research.

Finally, we attempted to review research conducted internationally, but our focus on history and policy is narrowly focused on the U.S. We do not have sufficient access to policy formulations in other countries and sometimes lack access to the many excellent studies completed by our international colleagues, especially in the intervention area.

Even with these stipulations, the range of research covered in this book is broad, and there is wide variation in the quality of the studies and syntheses we have selected for discussion. We generally tried to select the strongest possible studies and syntheses for review. As we show in Chapter 2, the scientific basis for LDs continues to evolve and has expanded since the first edition of this book in 2007. LDs are unique among developmental disorders not only in the dramatic growth of knowledge across different domains, but also in the extent of vertical, cross-disciplinary integration that has occurred, especially for word-level disorders (Elliott & Grigorenko, 2014; Lyon & Weiser, 2013; Shavelson & Towne, 2002). In the future, we believe that this type of cross-disciplinary integration is essential to the development of a comprehensive model encompassing all forms of LDs, and we offer this second edition in anticipation of continued development of an integrated understanding of LDs.

CHAPTER 2

Are Learning Disabilities Real?

Constructs like LDs are often questioned because there is no "gold standard" indicating what is or is not an LD. Indeed, a recent book (Elliot & Grigorenko, 2014) was entitled *The Dyslexia Debate* and was widely interpreted as suggesting that dyslexia does not exist. In a similar vein, LDs are often characterized as "mild" disabilities, and some question whether LDs are in fact conditions that meet criteria for a disability. The description of "mild" is difficult to reconcile with the adaptive consequences of being a poor reader or of having inadequate mathematics skills in our society. In this chapter, we provide an affirming "yes" to the question of whether LDs represent a "real" construct. We also provide a conceptual framework for understanding disorders like LDs, where the defining attributes exist along a continuum and are noncategorical (i.e., *dimensional*), unlike medical conditions like mumps and measles or life and death (Ellis, 1984).

We believe the evidence supports the validity of the construct of LDs, and that it has evolved as a scientific construct with an evidence base that should guide practice. We acknowledge that this evidence base is often not used as a basis for decision making in education, but argue that it should be used, especially in translating science into practice (see Chapter 11). Presently many approaches to identifying and treating LDs are not strongly evidence-based but have their roots in historical conceptions, anecdotes, unsystematic observation, and approaches for which the evidence base has been studied and found inadequate. The lack of attention to empirical evidence has hampered the field, much to the detriment of the children and adults with these types of academically based disabilities.

Most questions about whether LDs exist actually address uncertainty about how to define them. The ensuing controversy about definition is misconstrued as an argument about whether LDs represent true disabilities. To reiterate, there is no gold standard for any definition of LD, which is also the case for many other "disorders," such as ADHD, obesity, or hypertension (Ellis, 1984; Hinshaw & Scheffler, 2014). Rather, we use different types of measures to "indicate" the construct of LD. As we discuss in Chapter 3, these measures have inherent unreliability when it comes to identifying the extent to which a person displays the indicators of the construct, which occur on a continuum of severity. This does not mean that the indicators are not real or that the construct is not real; obesity and hypertension, which like LD rely on indicators that occur on a continuum of severity, are also real (Ellis, 1984). It simply means that valid measurement is nonnegotiable and essential.

Elliott and Grigorenko (2014) attracted considerable media attention for putatively questioning whether dyslexia existed. In fact, even a cursory reading shows that the authors did not really question whether dyslexia existed. Rather, they questioned whether the *term* had any specific utility because "dyslexia" was used in so many different ways and proposed purposes that the label was questionably meaningful, a longtime issue in the field. In particular, Elliott and Grigorenko noted that there was little indication that providing the label of "dyslexia" was associated with specific approaches to intervention. In Chapter 6, we suggest that many children with word-level reading difficulties benefit from interventions targeted at their specific reading and spelling weaknesses, regardless of whether the dyslexia term is or should be applied to the child. Our recommendation is that the use of the term "dyslexia" be referenced (in part) to the nature of the academic difficulties, a conspicuous problem in reading and spelling isolated words. This approach can reduce confusion of what to do when children have word-level difficulties (see Chapter 6), which is more important than the label.

The issue of whether LDs exist can be empirically addressed. In this chapter, we do so by providing a brief historical context to help explain why there is confusion—individuals with LDs are phenotypically heterogeneous, meaning that what people see is a blend of academic and behavioral difficulties that are variable. We discuss critical issues related to the construct of LDs, including the idea of LDs as an unobservable construct that are only identified by how they are measured; the measured attributes are dimensions that vary normally in the population (like weight and blood pressure) and become a problem with adaptation when they are on the extreme end of the distribution. They are heritable, have a basis in brain structure and function, and need intervention when the condition interferes with some form of adaptation. We then discuss evidence of the evolution of LDs as a scientific concept with a firm but changing evidence base that can guide research and practice.

HISTORICAL PERSPECTIVES AND U.S. PUBLIC POLICY

There are many reviews of the history of the concept of LDs (e.g., Doris, 1993), including Chapter 2 in the previous edition of this book (Fletcher et al., 2007; see the box at the end of the table of contents).

LDs originated in the concept of intrinsic behavior problems that originated in the brain, not the environment. These notions gave rise to the concept of minimal brain injury (Strauss & Lehtinen, 1947) and minimal brain dysfunction (MBD) in the 1960s (Clements, 1966). With the advent of DSM-III (American Psychiatric Association, 1980), the concept of MBD largely disappeared because the group identified with MBD was extremely heterogeneous. Instead, academic skills disorders and ADHD were separately defined, thus separating LDs and behavior disorders. Kirk (1963) and his colleagues formally introduced LDs as an educational entity. The essential tenets were that children with LDs (1) demonstrated learning difficulties that were "unexpected" given the children's strengths in other areas; (2) had different learning characteristics than children diagnosed with intellectual disabilities or emotional disturbance; (3) manifested learning characteristics that resulted from intrinsic (i.e., neurobiological) rather than environmental factors; and (4) required specialized educational interventions. No mention was made of intelligence, just of the absence of intellectual disability.

As with MBD, definitions of LD and dyslexia were difficult to operationalize and typically led to groups that were extremely heterogeneous

(Benton, 1975). The definitions specified no inclusionary criteria and were largely definitions by exclusion (Rutter, 1982). Genetic, cognitive neuroscience, and intervention research made little progress, partly because of the heterogeneity of the groups and the variation in selection criteria across labs (Doehring, 1978).

Why Are LDs Difficult to Define?

Three major issues make LDs difficult to define. As we noted in the first edition (Fletcher et al., 2007), LD represents an *unobservable latent construct* that does not exist apart from attempts to measure it. As such, LD has the same status as other unobservable constructs, such as IQ, achievement, or ADHD. The second involves the *dimensional* nature of LDs (i.e., the attributes representing LDs exist on a continuum and do not represent discrete categories; Ellis, 1984). The third issue is the problem of *comorbidity* with other developmental disorders (Pennington, 2009).

LDs Are an Unobservable Construct

LDs are a latent construct and not directly observable. Identification of a group of children whose academic underachievement is unexpected historically required ensuring the absence of other circumstances known to produce low achievement (sensory disorder, mental retardation, emotional disturbances, economic disadvantage, linguistic diversity, inadequate instruction), which leaves a very heterogeneous group. To remedy this problem, many efforts at definition and identification have been attempts to measure the attributes of *unexpected underachievement,* which epitomizes the LD construct. The primary approach to identification has been through cognitive discrepancy models in which the measurement of unevenness in academic or cognitive development is a marker for the "unexpectedness" of LDs, along with the exclusion of other causes of underachievement that would be "expected" to produce underachievement. Thus, children must be tested to identify discrepancies that would indicate unexpectedness and the latent construct of LDs.

A general problem that emerges with any form of testing is that the measures are imperfect indicators of the underlying construct. This is a problem with any approach to identification of LDs that involves psychometric tests. If different tests are used, different people will be identified with LDs because of differences in how the constructs are operationalized in the tests. This problem is magnified by slight amounts of unreliability in the measurements of the key academic, cognitive, and instructional attributes (see Chapter 3). We can observe what is measured, such as reading, math, cognitive processes, or instructional response. Each of these observable measures is intended to indicate, albeit imperfectly, the latent construct of LDs. The measurement is imperfect because no single measure captures all the components of the construct and each measurement contains a certain amount of error. The critical issue is the effect of these imperfect measurements on the reliability and validity of the overarching classification that is the basis for identifying LDs.

The Attributes of LDs Are Dimensional

The second issue is the dimensional nature of the attributes of LDs. As we observed above, most of the research on LDs, particularly that affecting reading, shows that the defining attributes occur along a continuum of severity rather than presenting as an explicit dichotomous category delineated by clear cut points on the achievement distribution. Indeed, the psychometric markers of LDs, such as achievement test scores, appear normally distributed in most population-based studies (Lewis, Hitch, & Walker, 1994; Rodgers, 1983; Shalev, Auerbach, Manor, & Gross-Tsur, 2000; S. E. Shaywitz, Escobar, B. A. Shaywitz, Fletcher, & Makuch, 1992; Silva, McGee, & Williams, 1985).

This conclusion is not without controversy. Some studies of children with LDs in reading have suggested that the distribution of achievement test scores is not normal and have identified a natural cut point where a separate distribution of nondyslexic poor readers can be identified (Miles & Haslum, 1986; Rutter & Yule, 1975; Wood & Grigorenko, 2001). In the studies summarized by Rutter and Yule (1975), the separate distribution, or "hump," has been attributed to an inadequate ceiling on the reading test (van der Wissell & Zegers, 1985) and to the inclusion of a large number of children with brain injuries who had IQ scores in the intellectually deficient range (Fletcher et al., 1994). However, most of the research generated surrounding the distribution of achievement scores in samples with LDs supports Stanovich's (1988) contention that people with LDs fall along a spectrum of impairment, that is, students with severe LDs do not differ *qualitatively* from students who land at the milder end of the spectrum. Findings supporting the dimensional nature of LDs are consistent with studies applying methods from behavioral genetics, which have not identified qualitatively different genetic constellations associated with the heritability of reading and math disorders (Fisher & DeFries, 2002; Grigorenko, 2005; Plomin & Kovas, 2005). As these are dimensional traits that exist on a continuum, there would be no expectation of natural cut points that differentiate individuals with LDs from those who are underachievers, but not identified with LDs; the distribution is simply a continuum of severity (S. E. Shaywitz et al., 1992).

If we evaluated the average performances of groups with and without LDs, as is done in empirical research, the dimensional nature of LDs (and the imperfection of measurements of the construct) would not be a major problem because the errors of measurement would be reflected in the variability around the mean. However, in public policy and educational applications it is necessary to identify individuals who have or do not have LDs. We rarely talk of degrees of LDs except in terms of severity, which is also a dimensional concept. The need to identify individuals for access to resources makes it necessary to categorize inherently normal distributions. Even with this need, the potential unreliability associated with these decisions must be recognized.

Comorbidity

Comorbidity refers to the co-occurrence of the attributes of two different disorders in the same person. It is well known that many children with dyslexia also have problems with math and/or ADHD. Sometimes they have accompanying speech and language disorders (Pennington & Bishop, 2009). In these instances, it is usually not the case that one problem causes another, although they may be linked. Rather, the individual actually meets diagnostic criteria for more than one disorder.

In retrospect, people who formulated early concepts of MBD were struggling with the fact that children with problems in reading or behavior often had overlapping difficulties. They also showed variable differences on cognitive, motor, and perceptual tasks that are still identified as special or pathognomic signs of LDs and targets for treatment, despite decades of evidence disputing whether LDs have any pathognomic signs and even clearer evidence that treating problems with perception, motor coordination, left–right reversals, and other "special signs" do not lead to improvement in academic skills (Mann, 1979) or ADHD behavior (Nigg, 2009; Hinshaw & Scheffler, 2014).

Exact determinations of comorbidity of LDs with other disorders vary considerably across studies and are ultimately arbitrary because any prevalence estimates depend on where the cut point is set for identification of the disorder. A major determinant is whether the individual is identified in the schools or in a clinic; the latter is associated with much higher rates of comorbidity diagnoses. However, estimates are that approximately 4–5% of the population experience comorbid word-level reading disability (RD) and ADHD (Carroll, Maughan, Goodman, & Meltzer, 2005; Pastor & Reuben, 2008), so that 25–50% of children identified as having word-level RD are also identified with LD (Pennington, 2009). About 20% of children with ADHD are identified with an RD and likely even more with math and writing problems, but these estimates are not reliably available (Carroll et al., 2005). Altogether, children with RDs are about four times more likely to present with ADHD behavior than children without an RD (Carroll et al., 2005). In many children, it is inattention rather than hyperactive-impulsive behavior that accounts for the common link with RD (Willcutt et al., 2010a; 2010b), although this is hardly an exclusive association. In terms of math and written expression, most people with reading problems also have writing problems; estimates of the co-occurrence of reading and math disability range from 30 to 70%, presumably because of shared cognitive liabilities (Willcutt et al., 2013).

Some researchers trying to understand comorbid relations of reading LDs and ADHD created an early framework suggesting that poor attention caused poor reading (Stanovich, 1986). Another early alternative hypothesized that poor reading leads to poor attention due to inability to fully engage in the classroom (Hinshaw, 1992). However, most of the current research is consistent with a *correlated liabilities* hypothesis, which predicts that some attributes are associated with ADHD and LDs in isolation, but that the different disorders share common weaknesses (Willcutt et al., 2010b). Interestingly, two recent reading intervention studies found that treatment for reading problems directly leads to improved reading, which in turn leads to improved teacher ratings of attention (Roberts et al., 2015; Miller et al., 2014). The hypothesis that inattention causes poor reading would predict that the reading intervention would have little effect on attention or that an intervention that improved reading would need to directly target attention skills, which in turn would affect reading. The intervention results described above do not support these predictions, finding instead that attention and reading improved in tandem.

More direct support for the correlated liabilities hypothesis comes from studies comparing cognitive performance in RD, math disability, and ADHD. Figure 2.1 compares cognitive processes in children impaired in word recognition with and without ADHD, showing that the two types of disorders are distinct and separable (Pennington et al., 2009; Willcutt et al., 2013; Wood, Felton, Flowers, & Naylor, 1991). LDs involving word recognition are consistently associated with deficits in phonological awareness regardless of the presence or absence of ADHD, whereas the effects of ADHD on cognitive functioning are variable, with primary deficits noted in processing speed, working memory, and other executive functions (Barkley, 2015; Pennington et al., 2009). Furthermore, ADHD appears relatively unrelated to phonological awareness tasks (Pennington, 2009). A child who meets the criteria for both an LD in reading and ADHD shows characteristics of both, but the impairments are more severe than those of a child with only one of the two disorders. This suggests that certain skills are impaired both by LD and by ADHD, so that when both disorders are present, these skills are doubly weakened. What these subgroups share most often are difficulties in processing speed for symbolic material (e.g., McGrath et al., 2011).

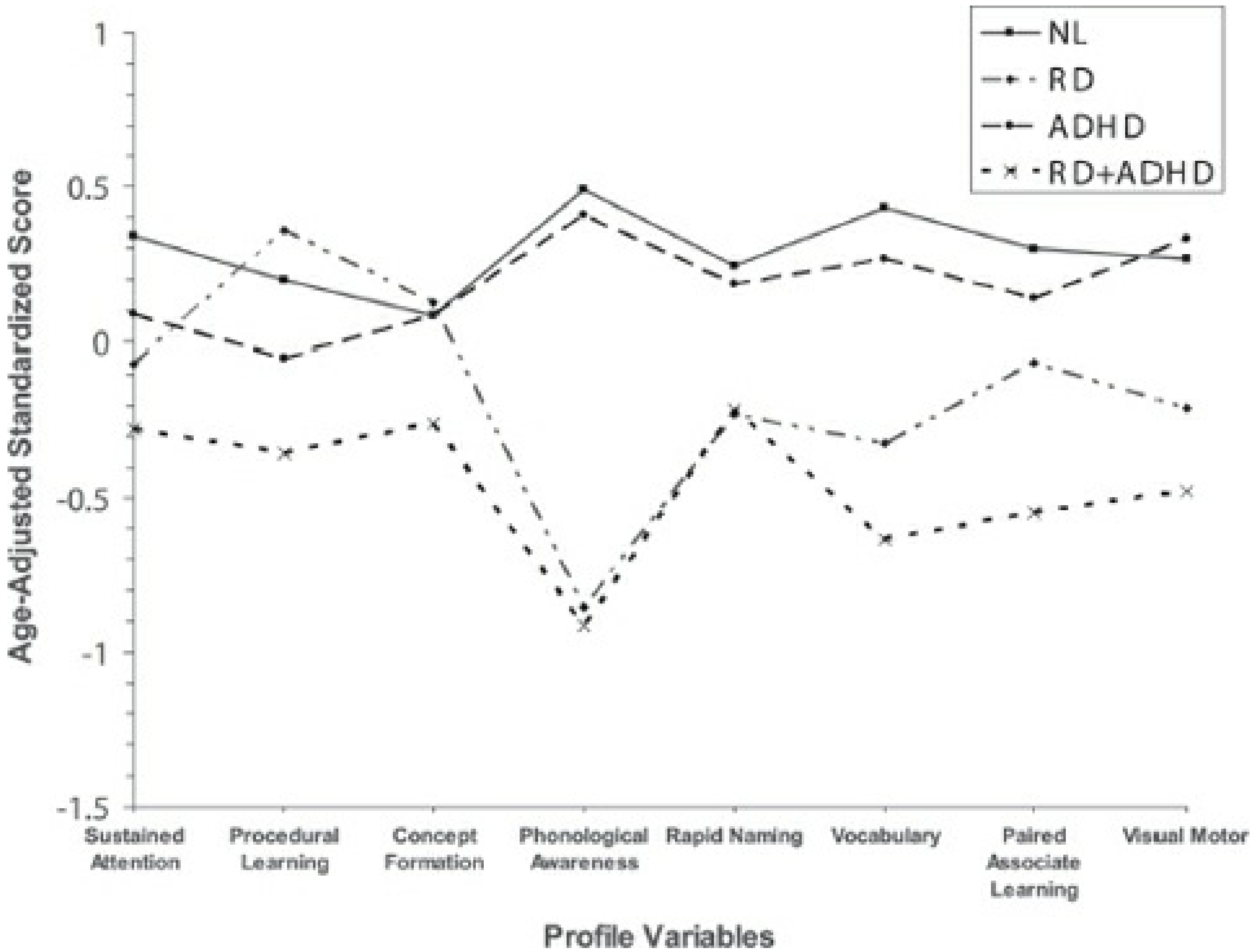

FIGURE 2.1. Profiles of cognitive performance by children with only reading disability (RD), only attention-deficit/hyperactivity disorder (ADHD), both RD and ADHD (RD + ADHD), and typically achieving children (NL). ADHD results in more severe RD, but the shape differences are not significant between the two reading-impaired groups. From Fletcher (2005, p. 310). Copyright © 2005 PRO-ED. Reprinted by permission.

In studies examining the comorbidity of math disabilities and ADHD (see Figure 2.2), the groups overlap more than groups with RDs and ADHD. This likely reflects the role of executive functions (strategy use, procedural learning) and working memory in both math disabilities and ADHD. The behavioral phenotypes of the disorders share deficits in working memory, processing speed, and verbal comprehension, but each disorder also has unique correlates (Willcutt et al., 2013). The disorders are separable on dimensions involving attention and behavior, with individuals who meet criteria for both disorders showing characteristics of both disorders. When children are identified with written language difficulties, ADHD is common

(Barkley, 2015), as are word-level reading problems. In most instances, these appear to be comorbid associations; a child with disabilities involving ADHD and a domain-specific LD appears like a child with ADHD through the behavioral lens, and like a child with LDs through the cognitive lens. However, when both lenses are considered simultaneously, the cognitive and academic deficits invariably appear more severe than the behavioral ones (Figures 2.1 and 2.2).

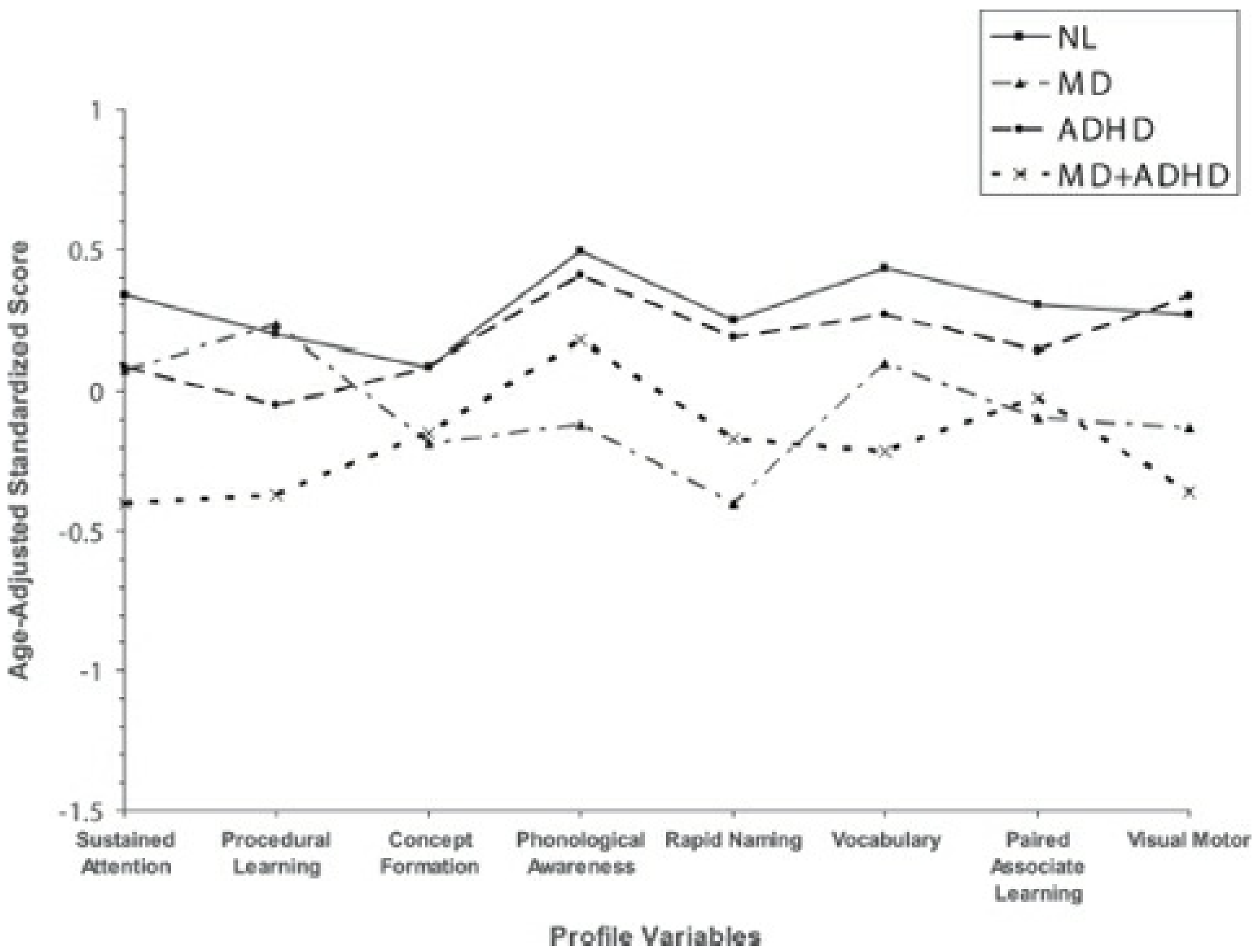

FIGURE 2.2. Profiles of cognitive performance by children with only math disability (MD), only attention-deficit/hyperactivity disorder (ADHD), both MD and ADHD (MD + ADHD), and typically achieving children (NL). ADHD results in more severe MD, but shape differences are not significant between the two math-impaired groups. From Fletcher (2005, p. 311). Copyright © 2005 PRO-ED. Reprinted by permission.

In a large study, Willcutt et al. (2013) compared cognitive performance in groups defined with only RD, only math disability, both an RD and a math disability, and a non-LD comparison group. All groups defined with LDs performed lower than the comparison group on most measures, with greater impairment in the group with both a reading and a math disability. Weaknesses in processing speed, working memory, and language comprehension were shared across all groups with LDs. However, the group with only a reading LD had weaknesses in phonological awareness and rapid naming. In contrast, only problems with set shifting were uniquely associated with math LDs. In another study making the same comparisons, Cirino, Fuchs, Elias, Powell, and Schumacher (2015) found that the group with both reading LD and math LD had the same weaknesses as the group with only reading or math LD, but they were more severe. Moll, Gobel, and Snowling (2015) compared verbal, visual–verbal, and visual number processing in children with only reading LD, only math LD, both reading and math LD, and typically developing children. Children with only RD were impaired only on verbal number tasks; children with only math LD were impaired across number tasks; and children with comorbid reading and math LD had deficits characteristic of both the other groups. They suggested that number processing in reading LD represented a phonological deficit, while math LD was associated with a more basic numerosity problem. These results support the correlated liabilities model of comorbidity because reading and math LDs have unique correlates, but share cognitive difficulties with processing speed, working memory, and language comprehension.

A final source of understanding of comorbidity comes from behavioral genetics research. These studies, which cut across potential domains of comorbidity, show that there are shared and unique genetic influences on the heritability of reading, math, and attention disorders. The shared influences have been articulated in the *continuity hypothesis* (Plomin & Kovas, 2005), which indicates that different characteristics of LDs and ADHD are associated with some of the same "generalist" genes: (1) the same genes influence high and low levels of academic abilities; (2) many of the genes associated with one aspect of LDs (e.g., phonological processing) also influence other aspects of this LD (e.g., vocabulary); and (3) some of the genes that influence one LD (e.g., RD) overlap with those that influence other LDs (e.g., mathematics disability) and ADHD.

We discuss these genetic issues in more detail in Chapter 6. It is important to remember that these correlates represent dimensional attributes of these domains and are correlated. The key to dealing with comorbidity in research and practice is to ensure that individuals are broadly assessed across domains so that the shared and unique components of academic and behavioral domains can be specified, especially if the goal is to develop an effective intervention program.

U.S. Public Policy

The difficulties with classification and definition have made policy formulations more difficult. Whereas researchers struggle with these fundamental issues, policymakers want approaches that are not complex and serve as vehicles for supporting services and allocating resources. It is interesting to examine U.S. public policy as it has evolved over the past 40 years to reflect the complexity of LDs.

Statutory Definition

Despite problems with definitions, through advocacy the concepts underlying emerging frameworks for LDs were eventually represented in U.S. public policy in 1968, forming what is still *the current statutory* definition of LDs in special education legislation with the adoption of Public Law 94–142 (Education of All Handicapped Children Act) in 1975:

> The term “specific learning disability” means a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which may manifest itself in an imperfect ability to listen, speak, read, write, spell, or to do mathematical calculations. The term includes such conditions as perceptual handicaps, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. The term does not include children who have learning disabilities, which are primarily the result of visual, hearing, or motor handicaps, or mental retardation, or emotional disturbance, or of environmental, cultural, or economic disadvantage. (U.S. Office of Education, 1968, p. 34)

Regulatory Definition

The statutory definition did not provide criteria for defining LDs as an entity.

In 1977, the U.S. Office of Education (now the U.S. Department of Education) provided the first regulatory definition of LDs, which was remarkable because it moved the underlying classification model from a neurological framework focusing on special signs indicative of presumed neurological dysfunction (e.g., perceptual–motor problems, letter and number reversals) to a psychometric framework focusing on cognitive discrepancies:

> [A child must exhibit] severe discrepancy between achievement and intellectual ability in one or more of the areas: (1) oral expression; (2) listening comprehension; (3) written expression; (4) basic reading skill; (5) reading comprehension; (6) mathematics calculation; or (7) mathematic reasoning. The child may not be identified as having a specific learning disability if the discrepancy between ability and achievement is primarily the result of: (1) a visual, hearing, or motor handicap; (2) mental retardation; (3) emotional disturbance; or (4) environmental, cultural, or economic disadvantage. (U.S. Office of Education, 1977, p. G1082)

The use of IQ–achievement discrepancy as an inclusionary marker for LDs had a profound impact on how LDs were conceptualized. There was some research at the time validating an IQ–achievement discrepancy method (Rutter & Yule, 1975), but these findings have not stood up over time (see Chapter 3). However, researchers, practitioners, and the public continued to assume that such a discrepancy was a marker for specific types of LDs that were unexpected and categorically distinct from other forms of underachievement. The impact of IQ–achievement discrepancy was clearly apparent in the regulations concerning LD identification in the 1992 and 1997 reauthorizations of the Individuals with Disabilities Education Act (IDEA), the name of the general special education statute that followed in subsequent reauthorizations of Public Law 94-142. The statute maintained the definition of LDs formulated in the 1968 legislation, and the regulations maintained the 1977 procedures until the 2004 reauthorization of IDEA.

IDEA 2004

In the most recent revision of IDEA (U.S. Department of Education, 2004), the regulatory definition of LDs was revised for the first time in 40 years. This occurred because the U.S. Congress passed statutes that permitted alterations of the 1977 regulations, indicating specifically that (1) states could not require districts to use IQ tests for the identification of students for special education in the LDs category, and (2) states had to permit districts to implement identification models that incorporated response to scientifically based instruction. In addition, the statute indicated that children could not be identified for special education if poor achievement was due to lack of appropriate instruction in reading or math, or to limited proficiency in English:

> A State must adopt . . . criteria for determining whether a child has a specific learning disability. . . . In addition, the criteria adopted by the State:
>
> - Must not require the use of a severe discrepancy between intellectual ability and achievement for determining whether a child has a specific learning disability. . . .
> - Must permit the use of a process based on the child's response to scientific, research-based intervention; and
> - May permit the use of other alternative research-based procedures for determining whether a child has a specific learning disability. (U.S. Department of Education, 2006, p. 46786)

In response to the statute, the Office of Special Education and Rehabilitative Services (OSERS) within the U.S. Department of Education (2006) published federal regulations for the revision of rules for the identification of LDs. The revision was partly a response to the converging scientific evidence bearing on the limited value of IQ–achievement discrepancies in identifying LDs (see Chapter 3). At the same time, it underscored the value of RTI in the identification process, formally operationalizing the assessment of instructional quality and the student's response as one part of the identification process. These components effectively shifted the concept of unexpected underachievement from a cognitive discrepancy to an instructional discrepancy, although approaches based on cognitive discrepancies are still permitted despite lack of evidence of their validity (Chapter 3). This summary is from the 2006 regulations:

> A child has a specific learning disability . . . if:
>
> - The child does not achieve adequately for the child's age or meet State-approved grade-level standards in one or more of the following areas, when provided with learning experiences and instruction appropriate for the child's age or State-approved grade-level standards: Oral expression. Listening comprehension, Written expression, Basic reading skills, Reading fluency skills,

Reading comprehension, Mathematics calculation, Mathematics problem-solving; or

- The child does not make sufficient progress to meet age or State-approved grade-level standards in one or more of the areas identified in 34 CFR 300.309(a)(1) when using a process based on the child's response to scientific, research-based intervention; or the child exhibits a pattern of strengths and weaknesses in performance, achievement, or both, relative to age, State-approved grade-level standards, or intellectual development, that is determined by the group to be relevant to the identification of a specific learning disability, using appropriate assessments . . . and the group determines that its findings . . . are not primarily the result of: A visual, hearing, or motor disability; Mental retardation; Emotional disturbance; Cultural factors; Environmental or economic disadvantage; or Limited English proficiency.

To ensure that underachievement in a child suspected of having a specific learning disability is not due to lack of appropriate instruction in reading or math, the group must consider, as part of the evaluation . . . :

- Data that demonstrate that prior to, or as a part of, the referral process, the child was provided appropriate instruction in regular education settings, delivered by qualified personnel; and
- Data-based documentation of repeated assessments of achievement at reasonable intervals, reflecting formal assessment of student progress during instruction, which was provided to the child's parents. (U.S Department of Education, 2006, pp. 46786–46787)

Although a number of advocacy and practitioner groups have questioned specific provisions of the regulations, what is encouraging is that most of these groups acknowledged the critical importance of using research to guide policies and practices concerning students with LDs, which is clearly reflected in the IDEA 2004 statutes and 2006 regulations. Equally significant in the new statute and regulations is the more explicit recognition that LDs should not be identified in the absence of evidence of appropriate instruction. Thus, the IDEA 2004 statute moved toward the accumulating research base on LDs by reducing the focus on IQ tests and emphasizing the critical role of instruction both for preventing LDs and for their identification.

DSM-5

DSM-5, the latest edition of the *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association, 2013a; see Tannock, 2013, for a summary of changes affecting LDs and ADHD), continues this emphasis and change in conceptual frameworks:

> Specific learning disorder is diagnosed through a clinical review of the individual's developmental, medical, educational, and family history, reports of test scores and teacher observations, and response to academic interventions. The diagnosis requires persistent difficulties in reading, writing, arithmetic, or mathematical reasoning skills during formal years of schooling. Symptoms may include inaccurate or slow and effortful reading, poor written expression that lacks clarity, difficulties remembering number facts, or inaccurate mathematical reasoning.
>
> Current academic skills must be well below the average range of scores in culturally and linguistically appropriate tests of reading, writing, or mathematics. The individual's difficulties must not be better explained by developmental, neurological, sensory (vision or hearing), or motor disorders and must significantly interfere with academic achievement, occupational performance, or activities of daily living. (American Psychiatric Association, 2013b)

DSM-5 explicitly recognizes that the attributes of LD (and ADHD) are on a continuum, but maintains an approach that is essentially categorical. The use of IQ–achievement discrepancy criteria were explicitly rejected because of lack of evidence of validity, although a threshold for low IQ is recommended to differentiate LD from an intellectual disability (essentially an IQ score greater than two standard deviations below the mean). DSM-5 has a category for communication disorders, into which it placed difficulties with speaking and listening. This is different from the U.S. IDEA definition of LD (see above), and appropriate because such disorders should be covered under "specific language impairment" (SLI) in IDEA. Although there is some overlap between SLI and learning disabilities, it is far from complete (Bishop & Snowling, 2004).

DSM-5 identified different types of LD in reading (word-reading accuracy, reading fluency, and reading comprehension), written language (spelling accuracy, grammar and punctuation accuracy, organization of written expression), and mathematics (basic number sense, accuracy and fluency in recalling number facts, calculation accuracy and fluency, and math reasoning). Within these categories of academic skills deficits, there are four primary criteria for identification: (1) *persistence* despite the provision of adequate treatment for at least 6 months; (2) *low achievement,* with scores below the mean for age on a norm-referenced academic achievement test (with no specified threshold, although recommendations for a range of at least one to one and a half standard deviations are implied); (3) *age of onset,* with the problem manifesting during early years of schooling; and (4) *exclusions* of cases in which there is evidence that another condition (e.g., intellectual disability, sensory problem, other mental or neurological disorder, psychosocial adversity, lack of educational opportunity) provides a better explanation for the presence of persistent low achievement.

Some of the controversy about DSM-5 involved its failure to use the term "dyslexia," although by covering problems with the accuracy and fluency of single word-reading skills, it addresses dyslexia in all but name. Altogether, DSM-5 is more strongly aligned with current scientific evidence than previous formal definitions.

Summary: Historical Perspectives

U.S. public policy has evolved to reflect the current state of LD research, especially in the 2004 reauthorization of IDEA and DSM-5. In contrast with earlier attempts at describing LDs, these recent definitions are more tightly aligned with empirical findings, removing exclusive reliance on IQ–achievement discrepancy methods. In addition, these definitions recognize that there is heterogeneity in the academic presentations of LDs, manifesting in subgroups of LDs. Finally, the definitions are more explicit about links with intervention response.

THE REALITY OF LDs

It may seem odd to address an issue like the reality of LDs. However, it is an issue that has always been raised, usually by referring to them as "mild" disabilities or simply referring to people who struggle because of LDs "lazy" or "unmotivated," or by suggesting that someone with an LD has to be at the bottom of the achievement distribution. There are three sources of data that specifically address whether LDs are real. The first, which should be a focus of more research, is the impact of LDs on adaptation. The second source is classification research (see Chapter 3), which treats LDs as a hypothetical construct that is operationalized and then validated. The third is whether the construct has generated an empirical base of research.

Do LDs Interfere with Adaptation?

Education policy requires that in order for a problem with academic skills to be eligible for special education, there must be evidence of "educational need." This criterion typically means that the person has poor grades, can't pass state accountability tests, and generally needs accommodations or interventions that extend beyond what can be provided in general education. Consensus-based classifications like DSM-5 also require evidence that the academic problem leads to adaptive impairment with grades, social functions, or other domains. There is clear evidence that LDs interfere with adaptation on both short-term and long-term bases. In an ideal world, we would use information on adaptation to help set thresholds for issues like identification and intervention intensity. However, insufficient focus on long-term academic outcomes in relation to adaptation at school, home, and society (vocational success, social adjustment) is a limitation of research on LDs.

When adaptation is examined, it is clear that identification with LDs is associated with impairments in multiple domains. Figure 2.3 is based on data from Willcutt et al. (2007) and compares a group of typically developing students with groups of students with word-level reading disability, ADHD, and both ADHD and RD on a variety of adaptation-related outcomes: rates of student retention, school- and parent-identified academic impairment, evidence of social and/or occupational impairment, and involvement with the juvenile justice system. The participants were 8- to 18-year-olds when originally recruited for a longitudinal study in which all twin pairs in the Colorado Range region were asked for permission to review their records. If there was evidence of a reading problem, both twins were recruited and assessed in multiple domains.

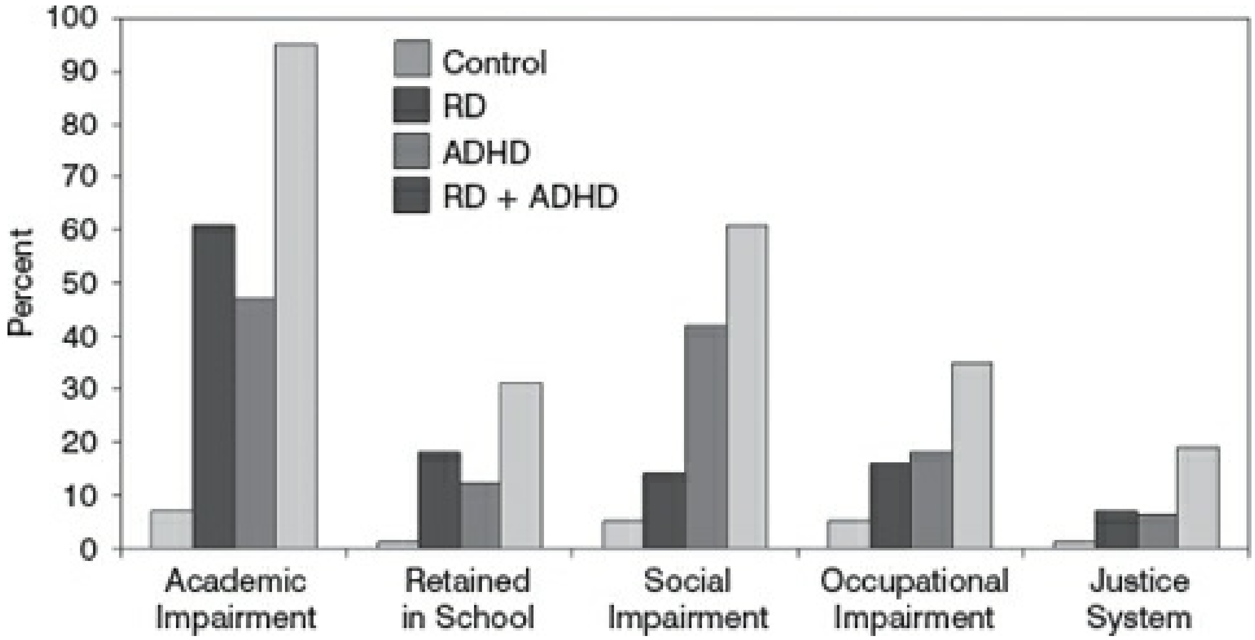

FIGURE 2.3. Academic impairment and social outcomes in groups with RD, ADHD, and RD + ADHD. Adapted from Willcutt et al. (2007). Courtesy Eric Willcutt.

The data presented in Figure 2.3 are from a 5-year follow-up of a portion of this large sample. All three groups with RD and/or ADHD show significantly higher rates of impaired adaptation than the comparison group, with the group with both RD and ADHD showing rates of academic impairment and school retention that are slightly higher than in the group with RD alone. Occupational impairment is also higher in RD alone and ADHD alone, with ADHD alone leading to more social impairment. A comorbid presentation leads to high adaptation difficulties across all comparisons. Additional assessments in Willcutt et al. (2007) showed higher rates of mood and behavior difficulties, as well as substance abuse in all groups relative to the comparison group. These results provide strong evidence that RDs in isolation and comorbid with ADHD lead to difficulties with adaptation.

Another approach to assessing adaptation issues is to survey individuals affected by LDs, which the National Center for Learning Disabilities (NCLD; 2014) has done three times. On the most recent survey, conducted in 2012, parents identified clear adaptive impairments with negative effects on overall adjustment at school and home, and for long-term vocational outcomes. The NCLD report also examined publically available data, finding that LDs were associated with lower grades, failure on high-stakes tests, retention (one-third of school-identified LD cases), lower graduation rates, and higher rates of school dropouts. Post-high school, people identified with LDs were more likely to have high rates of involvement with the criminal justice system, with lower rates of college participation and completion. They form the largest disability group seeking vocational services. This NCLD report echoed the findings of Willcutt et al. (2007) and established that LDs were associated with adaptive impairments that can be significant.

The NCLD report identified several public misconceptions about LDs, such as a strong association of IQ and LD, causal attributions to excessive TV watching, and common beliefs that eye glasses and medication remediate LDs. Other data were cited showing that many people associate LDs with intellectual disabilities, sensory impairments, the home environment, and laziness. Despite progress in research, policy, and practice, the general public still has a weak understanding of LD as a construct.

Classification Approaches

Another strategy for establishing the validity of the concept of LDs is to approach them from an empirical classification perspective (Chapter 3). Classifications, definitions, and identification are not the same. Classifications are systems that permit a larger set of entities to be partitioned into smaller, more homogeneous subgroups based on similarities and dissimilarities in attributes thought to define different aspects of the phenomenon of interest. The process of designating entities as belonging to subgroups represents an operationalization of the definitions emerging from the classification. Identification (or diagnosis) occurs when the operational definitions are used to determine membership in one or more subgroups. This process occurs in biology when plants and animals are assigned to species; in medicine when diseases are organized into categories based on etiology, symptoms, and treatment; and in LDs when a determination is made that a child's difficulties in school represent LD as opposed to a behavior problem, oral language problem, or intellectual disability. Even deciding that a child needs academic intervention is a decision that reflects an underlying classification (children who need or do not need intervention; Morris, 1988).

Many of the issues involving different methods for identifying children with LDs reflect confusion about the relations of classification, definition, and identification. The relation is inherently hierarchical, in that the definitions derived from a classification yield criteria for identifying members of the subgroups. This hierarchy is depicted in Figure 2.4, which uses the concrete examples of LDs, ADHD, and intellectual disabilities. Definitions of LDs originate from an overarching classification of childhood disorders (as in DSM-5) that differentiate LDs from intellectual disabilities and various behavior disorders, such as ADHD. This *classification* yields *definitions* and resultant criteria based on attributes that distinguish LDs from intellectual disabilities and ADHD. These criteria can be used to *identify* children as members of different subgroups within the classification model.

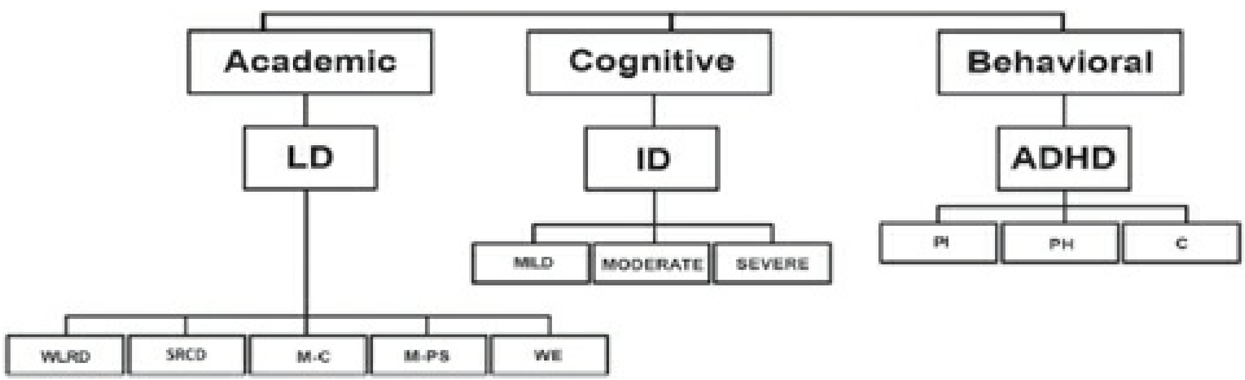

FIGURE 2.4. Classification of learning disabilities (LD), intellectual disabilities (ID), and ADHD. The diagram also shows major subgroups under LD (word-level reading disability [WLRD], specific reading comprehension disability [SRCD], math calculations [M-C], math problem solving [M-PS], and written expression [WE]); ID (mild, moderate, and severe levels); and ADHD (predominantly inattentive presentation [PI], predominantly hyperactive–impulsive presentation [PH], and combined presentation [C]). Courtesy Whitney Roper.

Although this classification terminology describes groupings, we are really referring to decisions about how individuals are related on correlated attributes that are dimensional and that defines and overlaps across the subgroups. For LDs, the decisions are arbitrary and are subject to measurement error that leads to identification issues, as we explained above in discussing the unreliability of assessing unobservable latent constructs. However, these problems do not subjugate our capacity to demonstrate the validity of the construct of LD, even though it is unobservable. Thus, it is critical to formally assess the validity and reliability of the subgroupings. The fact that subgroups can be created does not necessarily mean that valid classifications exist. Rather, the subgroups making up a valid classification can be differentiated according to variables not used to establish the subgroups (Skinner, 1981), demonstrating external validity of the classification. Internal validity and reliability hinge on evidence that the classification is not dependent on the method of classification (i.e., changing the measurement methods used to create subgroups does not change the essential nature of the subgroups), can be replicated in other samples, and permits identification of the majority of cases of interest. Reliable and valid classifications facilitate communication within scientific and professional communities, prediction of impairment severity, and treatment planning (Blashfield, 1993).

Certain types of classifications may be more useful or appropriate for some purposes than others. Classification may be needed to identify children as needing intervention; as having LD or as being typically achieving; as having LD as opposed to an intellectual disability or ADHD; and, within LDs, as being reading- rather than math-impaired. As LDs are hypothesized to represent a subgroup of people with *unexpected* underachievement, LD is differentiated from *expected* underachievement due to emotional disturbance, economic disadvantage, linguistic diversity, and inadequate instruction (Kavale & Forness, 2000). These types of classification represent hypotheses that should be evaluated for the reliability of the hypothetical model and for validity by reference to variables that are different from those used to establish the classification and assign individuals to subgroups.

It can be demonstrated that different academic subgroups of LD can be reliably differentiated on attributes not used to define them. Consider, for example, Figure 2.5. This figure displays cognitive profiles for three groups of students in grades 2 and 3 who participated in a classification study by the Yale Center for Learning and Attention Disorders (S. E. Shaywitz, 2004). These children represented groups with isolated problems in the domains of word recognition and math, along with a comparison group of typically achieving children. The subgroups of students with word recognition and math computational problems were identified according to several different approaches to definition, including discrepancies relative to Verbal IQ, Performance IQ, or Full Scale IQ, as well as a low-achievement definition that simply required performance below the 26th percentile on either word recognition or math calculations coupled with an IQ score of at least 80, but with no requirement of a discrepancy.

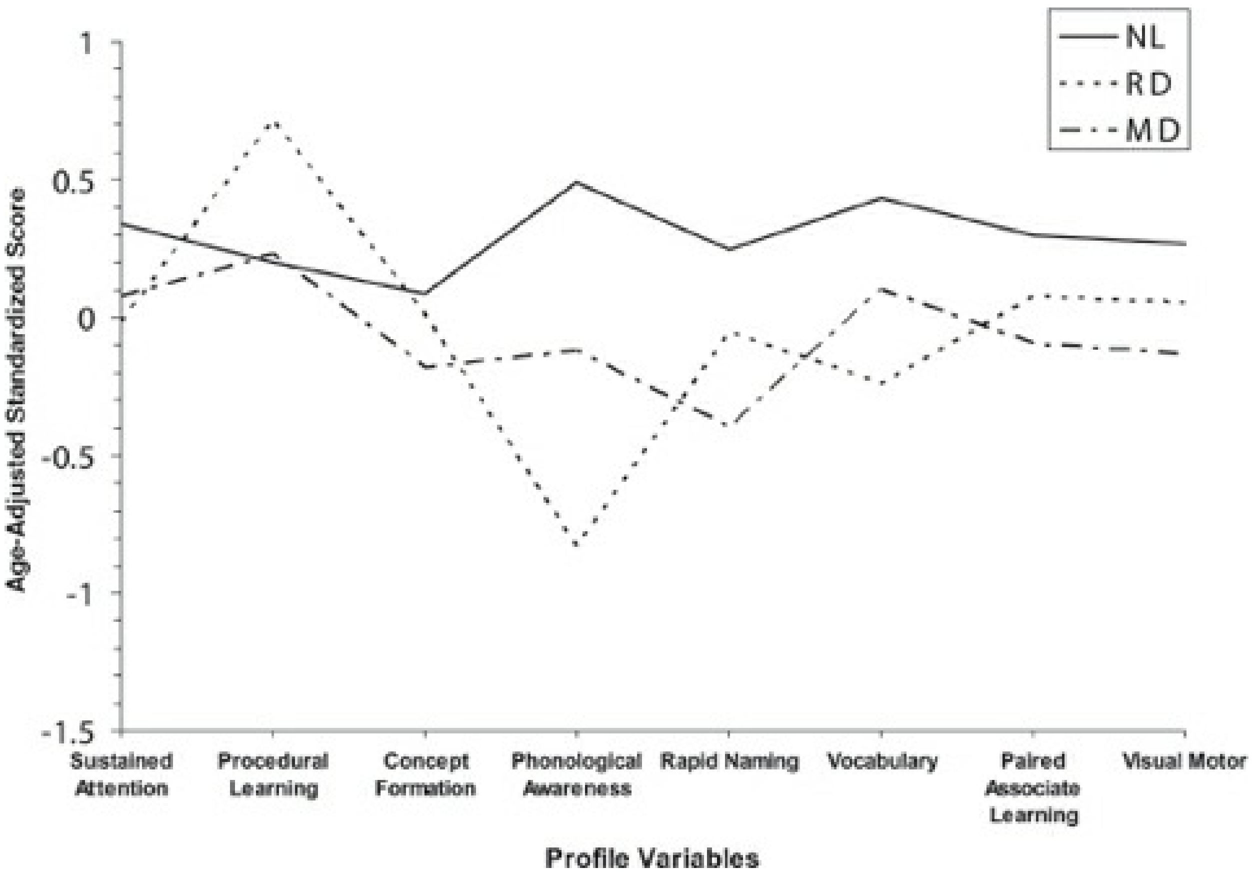

FIGURE 2.5. Profiles across different cognitive tests for children who are impaired only in reading (RD) and only in math (MD) relative to typical achievers (NL). The groups differ in shape and elevation, suggesting three distinct groups.

To validate the underlying hypothetical classification of LDs into reading versus math subgroups, the children received assessments of cognitive skills that were not used to identify the LD subgroups. These measures included assessments of problem solving, concept formation, phonological awareness, rapid naming, vocabulary development, verbal learning, and visual motor skills. As Figure 2.5 shows, the three groups were distinct in their patterns and levels of performance, indicating that the implicit classification of LDs in reading versus math subgroups is supported, along with clear evidence that children defined with LDs in reading and math domains differed from typically achieving students. As will be seen in subsequent chapters, subgroups similarly defined differ in both the neural correlates of reading and math performance and the heritability of reading and math disorders. These achievement subgroups, which by definition include children who meet either low achievement or IQ–achievement discrepancy criteria, also differ in response to instruction: Effective interventions are specific to the academic domain, so that teaching math to children whose problem is in reading (and vice versa) is ineffective (see Morris et al., 2012, for an empirical demonstration).

Scientific Maturation of the Field of LDs

A third approach to demonstrating the reality of LDs is to ask about the state of empirical science in the field. Given the significant debates in education about what constitutes “science” and, to a lesser extent, a “scientific field,” we refer to *science* as the pursuit of knowledge based on observable phenomena capable of replication and validation, resulting in a body of reliable knowledge that can be logically and rationally explained. The field of LDs actually constitutes a subfield of education science, which is a subfield of the social and behavioral sciences. Much of the research in LDs over the past decade integrates scientific principles and methods from several subdisciplines including education, cognitive psychology, neuroscience, and genetics. Scientific research in LDs has varied considerably over the past century with respect to the questions asked, the designs and methods used, and the interpretation of the data. The question central to demonstrating the reality of LDs is: Has the last decade provided the field with a more consistent application of scientific principles in the study of LD?

To address the question of LD as a scientific construct requires not simply an examination of the quantity and quality of research on topics relevant to LD but also assessment of the extent to which an appreciation for evidence can be inferred in general ways from education policies in education (and particularly special education) that strongly recommend assessment and instructional practices grounded in well-defined converging research outcomes (see Chapter 11). Such an evidence-based culture requires that practitioners in the field be educated to make decisions about identifying, selecting, and implementing effective practices on the basis of trustworthy research. Trustworthy research is characterized by studies that pose relevant questions objectively, seek knowledge using appropriate research designs and methods, promote replication of findings, ensure that samples being studied represent the population in question, and consider the conditions under which the new knowledge can be implemented. Unfortunately, as we discuss in Chapter 11, the scientific basis exists, but has not been used on a consistent basis, and the idea that decision making in education should rely upon scientific principles has no real consensus. Therefore, we focus on the scientific basis in the remainder of this chapter and return to its use in Chapter 11.

What Is Scientifically Based Research?

The National Research Council (NRC), a branch of the National Academy of Sciences, published a report in 2002 titled *Science and Education* (Shavelson & Towne, 2002). The report, commissioned by the U.S. Department of Education (USDOE), stated that, in order for studies to be deemed scientifically based, they must: (1) pose significant questions that can be investigated empirically; (2) link research to theory; (3) use methods that permit direct investigation of the question; (4) provide a coherent and explicit chain of reasoning; (5) replicate and generalize across studies; and (6) disclose research data and methods to encourage professional scrutiny and critique.

“Evidenced-based education” (EBE) and “evidence-based practices” (EBP) are terms that have entered the education lexicon in the past decade, along with the ubiquitous “scientific-based research” (SBR). The terms have roots in medicine, which embraced them over the past three decades. Sackett, Rosenberg, Gray, Haynes, and Richardson (1996) were early users of the term “evidence-based” in medicine. They described “evidence-based medicine” (EBM) as the explicit utilization of the best evidence from research in clinical decision making. In a later publication, Sackett, Straus, Richardson, Rosenberg, and Haynes (2000) emphasized that EBM integrates the most robust research evidence with clinical expertise and patient values. According to Hood (2003), this concept of the use of evidence in identifying effective treatments emphasizes a decision-making process where judgments about what is best for each patient are made on a case-by-case basis using the best evidence available.

For educators, the term "evidence-based practices" represents a broader and more practical concept than "scientific-based research" because it incorporates, as it does in medicine, practitioner decision making. Both EBP and SBR share the common goal of ensuring that the practices we implement are valid, but EBP places a greater emphasis on the role of the clinician/practitioner in customizing the extent to which SBR and EBP are combined. This does not mean, however, that EBP reflects a less rigorous process for identifying and implementing effective programs in schools. As Hood (2003) explained:

> An evidence-based practice (EBP) is any practice that has been established as being effective through scientific research that conforms to some set of explicit criteria . . . [including] (1) the practice has been standardized through manuals, guidelines, or certified training in the practice, (2) the practice has been evaluated through controlled research . . . , (3) objective measures were employed that demonstrated valued outcomes, and (4) these outcomes have been replicated by . . . research. (p. 14)

Is Research on LDs Scientifically Based?

A noteworthy advance in research on LDs during the past 10 years since the publication of the first edition (Fletcher et al., 2007) has been the increased application of robust experimental designs and methods appropriate to the specific research questions posed—a critical requirement for meaningful data that was not always common in education research (Shavelson & Towne, 2002). This advance has led to significant improvements in isolating specific cognitive, linguistic, neurobiological, genetic, and instructional factors and their interrelationships that characterize different types of LDs. Moreover, much of this research has been conducted within a multidisciplinary longitudinal context, thus allowing for an examination of the developmental course of well-defined LD and how the trajectory of that course can be influenced by intervention efforts, genetics, and neurobiology. The more frequent application of randomized controlled trials (RCTs) and regression discontinuity designs (RDDs) that permit strong inferences about causality to determine the effectiveness of programs and practices, which we encouraged in the first edition of this book, underscores the increasing quality of research in LD.

Lyon and Weisner (2013) reviewed research from the decade 2001 to 2011 and found that the research literature reveals that as a field, scientific research on LDs is aggressively working to elevate research standards. With the increase in rigorous application of scientific methods, LDs research is comparable to other areas of behavioral and social sciences research in its ability to produce sound findings, allowing a better understanding of individual differences in complex learning domains and contexts. The research was characterized by increased application of robust research designs and experimental methods to LD research questions, as well as a significant increase in the number of multidisciplinary studies involving close collaboration among researchers in special education, educational psychology, cognitive psychology, developmental neuroscience, neuroradiology, genetics, psychiatry, and classification science.

The impact of research on LDs over the past 10 years shows that the quality and impact of recent research has significantly enhanced our understanding of etiology, phenotypic characteristics, and instruction outcomes. The increase in the quality and impact of research has made possible the actual use of the research evidence in forging federal and state education policies relevant to LD. A common language has also emerged to describe the scientific levels of evidence as seen in terms such as SBR and EBP. The remainder of this book will pick up where Lyon and Weisner (2013) left off, expanding the research review in the same domains with continued and increased impact.

CONCLUSIONS: LDs ARE REAL

In this chapter, we have made a case for the fact that LDs are real entities. They are hard to define because LD as a construct is not observable independently of how we measure it. Additionally, the attributes measured to define LDs are dimensional and correlated, leading to comorbidity and overlap across domains and with attributes of ADHD. Despite these challenges, clear evidence has emerged showing that LDs interfere with adaptation on a short-term and long-term basis. Empirical approaches to classification support the validity of the concept, especially in relation to different domains of reading and math, and to ADHD. There is a substantial evidence base that can be used to support decision making in multiple areas related to LDs. Altogether, there is clear evidence that LDs are real. Much of the controversy is how to define them, a topic we turn to in Chapter 3.

CHAPTER 3

Classification and Definition of Learning Disabilities

The Problem of Identification

The evidence base underlying classification, definition, and identification issues in LDs can be subjected to decision-making processes that have a basis in scientific research for determining optimal approaches. Many of the issues are not fully resolved, but progress has been made and, in many instances, there is a research-informed consensus for making informed decisions. Often this consensus is about what *not to do,* with active debate around what *should be done.* On the surface, identification seems straightforward. The attributes of LDs that indicate that a person is a member of a class of all people with LDs and is not a member of other classes of people that do not fully share these attributes need to be defined. If the attributes are known, they can be measured. Criteria can be established and individuals can then be reliably identified into classes of people simply classified as having LDs and not having LDs.

Classification depends on a theory of what constitutes LDs and specifically what represents unexpected underachievement. Historically, unexpected underachievement has been defined from neurological, cognitive discrepancy, and instructional frameworks. Each of these frameworks leads to an operational definition of the critical attributes of LDs that can be assessed at an observable level for identification purposes.

This relation of observable and latent constructs is shown in Figure 3.1, which is based on an instructional framework and uses three potential indicators of LDs: (1) low achievement; (2) unexpected underachievement (e.g., cognitive discrepancy, instructional response); and (3) exclusionary factors, which involve contextual factors influencing learning and performance (see Figure 1.1) and other disorders that preclude LDs. In Figure 3.1, the underlying conceptual model for unexpected underachievement is inadequate response to instruction. In a method based on neurological dysfunction, it would be a special sign of brain dysfunction, such as finger agnosia or perceptual–motor difficulties. A cognitive discrepancy method would specify differences in aptitude and achievement, or differences in two cognitive domains. Most methods would include low achievement, which is necessary, but not sufficient, for identification of LDs. These different approaches to measuring unexpected underachievement are hypotheses about how to classify and define LDs. As hypotheses, they need to be tested.

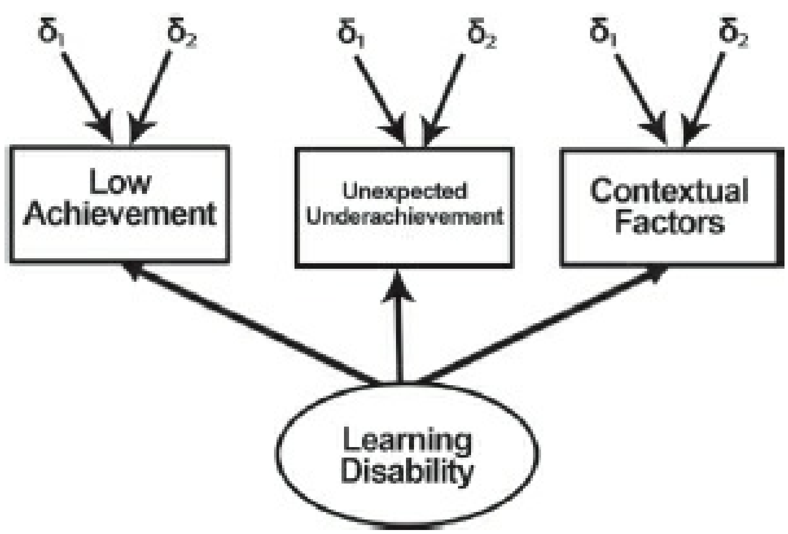

FIGURE 3.1. Model for the relation of the latent construct of LDs and observable attributes from an instructional model involving low achievement and inadequate instructional response. The two observable indicators can be measured, but are imperfect because of measurement error. Multiple indicators typically increase the reliability of indicating a latent, unobservable construct. Courtesy Whitney Roper.

Since we can only measure the observable attributes, identification will always have some inherent unreliability even if the measure itself is high in reliability. All measurements that involve human behavior have error, so that any indicator of an attribute will have a certain degree of unreliability that will affect the precision by which individuals with these attributes can be identified. Using multiple indicators helps address the unreliability issue, but unreliability can't be fully resolved and needs to be taken into account when identifying individuals with LDs.

ATTRIBUTES OF LDs

Inclusionary Criteria

LDs are traditionally defined by indicators that are both inclusionary and exclusionary. Inclusionary criteria indicate the presence of LD, such as an indicator of unexpected underachievement. Exclusionary criteria indicate the absence of LD. An example is that a person's low achievement is not due to an intellectual disability. Most controversy surrounds inclusionary criteria because they represent the core attributes of the concept of LDs. This controversy is not about whether the core concept of LDs is unexpected underachievement. What varies across classification hypotheses is the conceptual framework for indicating unexpected underachievement. In the next section, we review different hypotheses about which indicators of unexpected underachievement are most useful for identification and classification of cases.

Neurological Hypotheses

Historically, the earliest conceptions of LD were neurological, representing the idea of LDs as disorders of constitutional origin. As we discussed in Chapter 2, these conceptions produced heterogeneous groups of children with problems ranging from hyperactivity to poor academic skills, often reflecting comorbidity. From this classification hypothesis, identification was based on the presence of signs of neurological dysfunction, which might be reflected in hyperactivity, clumsiness, sensory–motor difficulties, or language problems (Benton, 1975). The presentation of these "symptoms" was believed to relate to the integrity of the central nervous system. Interventions based on these classifications, such as perceptual–motor or auditory/visual modality training, were not strongly related to academic outcomes (Mann, 1979;

Vellutino, 1979). Because neurological origin was assessed indirectly by behavioral measures (and thus was an inference that was difficult to support), these hypotheses have receded in favor of alternative approaches to indicating unexpected underachievement.

Cognitive Discrepancy Hypotheses

Cognitive discrepancy hypotheses are more contemporary conceptions of LD that focus on unevenness in cognitive abilities. The most prominent of these hypotheses is the aptitude–achievement discrepancy hypothesis, commonly operationalized as a discrepancy between measured IQ and academic achievement. Within this model, the IQ score, for identification purposes, must exceed the achievement score, with the numerical magnitude of this "gap" varying depending on different policies adopted by districts and states. A "gap" of significant magnitude indicates that a child has a discrepancy consistent with LD. The absence of a gap indicates a "slow learner" who is achieving at the limits of his or her aptitude. The IQ–achievement discrepancy was the central part of U.S. federal regulations for identification from 1977 to 2004, and is still permitted under IDEA 2004. However, there is little support for the validity of this hypothesis in classification and identification processes (Fletcher et al., 1994; Siegel, 1992; Stanovich & Siegel, 1994; Stuebing et al., 2002, 2009).

Other cognitive discrepancy approaches represent intraindividual difference methods that focus on patterns of strengths and weaknesses in cognitive skills as a core inclusionary attribute indicative of unexpected underachievement. It is well-established that LDs are associated with specific impairments in cognitive processes and that there is variability in the cognitive strengths and weaknesses displayed by individuals with LDs. In definitions based on this classification hypothesis, LD is identified when there are strengths in some cognitive functions and weaknesses in other cognitive functions related to academic achievement. For example, an individual with low achievement in the word-reading domain might display strengths in visual–spatial skills and poor phonological processing skills. Thus, there is a weakness in an academic domain with a corresponding weakness in a domain correlated with poor word reading and a strength in a cognitive domain presumably unrelated to reading. However, the emerging research that addresses the reliability and validity of classification models based on patterns of strengths and weaknesses in cognitive skills continues to show little support for the use of these methods as a component of identification or intervention.

Instructional Discrepancy Hypotheses

The most recent classification hypothesis uses data from service delivery frameworks based on MTSS or, when identification is involved, RTI. An MTSS framework is primarily an approach to service delivery in schools, with the goal of improving academic and behavioral outcomes for all children (Vaughn & Fuchs, 2003; Fletcher & Vaughn, 2009). In classification based on the RTI identification methods, there are two core attributes in the underlying classification that are inclusionary: inadequate instructional response and low achievement. The core indicator of unexpected underachievement is inadequate instructional response, representing an attribute that can only be assessed in relation to efforts to teach the person. Thus, *intractability* to intervention is the measureable indicator of unexpected underachievement so that underachievement occurs despite adequate instruction. As with all identification methods for LDs, there are problems with the reliability of identification of individuals with LD. There is growing evidence of the validity of a classification incorporating inadequate instructional response and low achievement (Miciak, Fletcher, & Stuebing, 2015a).

Exclusionary Criteria

Certain conditions are represented as “exclusions” for LDs because they may represent other *primary* causes of low achievement, which means that low achievement would be expected, not unexpected. The exclusions could represent another disorder (e.g., sensory disorders, intellectual disabilities, behavioral difficulties that interfere with motivation or effort) or contextual factors like economic disadvantage, minority language status, and poor instruction that are often associated with low achievement.

Definition by Exclusion

Early definitions often based identification of LD solely on the presence of low achievement and the absence of exclusionary conditions. In a sense, these approaches defined LD by “what it is not.” Defining a disorder by exclusion is not a satisfactory approach to classification because it does not produce a conceptual model of what LD might represent. It has not been a useful approach to defining LD because the group of children emerging as having LD is very heterogeneous if only poor academic achievement and exclusionary criteria are applied (Rutter, 1982). However, a classification hypothesis based on simple low achievement with or without the exclusionary conditions should be considered because it is often the de facto method in research.

Stipulating that LDs are not due to intellectual disabilities, sensory disorders, or linguistic diversity is reasonable, as children with these characteristics have different intervention needs. A person whose primary language is a minority language should not be identified with LDs unless it can be demonstrated that the difficulties producing the reading or math problem are a pervasive characteristic across languages (see Chapter 4). There are also issues with distinctions between intellectual disabilities and LDs that make the precise demarcation unclear, but information beyond IQ tests is essential for identifying cognitive impairment (Schalock et al., 2010).

Other exclusions stemmed from policy decisions that involved the need to avoid the mixing of special education and compensatory education funds, as well as the existence of other eligibility categories in IDEA to support children with special needs (e.g., intellectual disabilities, emotional disturbance). The original exclusionary criteria were not meant to preclude children from placement, but to better classify each child’s difficulties—on the assumption that when economic disadvantage, emotional disturbance, and inadequate instruction are the primary causes of underachievement, different interventions are needed.

For the other conditions considered exclusionary of LDs, determining which are “primarily” the cause of underachievement has proven a difficult proposition. The cognitive correlates of academic difficulties in children with achievement deficiencies attributed to emotional disturbance, inadequate instruction, and economic disadvantage do not appear to be different according to these putative causes. Moreover, the intervention needs and mechanisms whereby interventions work do not appear to vary according to these factors (Fletcher, Denton, & Francis, 2005a; Lyon et al., 2001). As such, these distinctions are not strongly related to the types of intervention programs that are likely to be effective. Of particular concern is the idea that inadequate instruction precludes identification of LDs, when in fact it may cause people to manifest the attributes of LDs. Later in this section, we examine specifically exclusion due to socioeconomic disadvantage and lack of opportunity for learning.

Emotional and Behavioral Difficulties

Most definitions of LDs exclude individuals whose poor achievement is due primarily to emotional and behavioral difficulties. This assessment is difficult to make because of comorbidity (Chapter 2). Determining which disorder is primary is difficult, as those who struggle may develop behavioral difficulties that are secondary to lack of success in school. Researchers have also reported that children with reading disabilities present with co-occurring social–emotional difficulties. In some clinical studies, these difficulties appear to be secondary to difficulties in learning to read. For example, of the 93 adults in a clinic population with LDs, the majority of whom displayed reading problems, 36% had received counseling or psychotherapy for low self-esteem, social isolation, anxiety, depression, and frustration (Johnson & Blalock, 1987). Likewise, others (Bruck, 1987) have reported that many of the emotional problems displayed by readers with LDs reflect adjustment difficulties resulting from labeling or academic failure. Large-scale clinical trials show that improving reading and math instruction in programs that provide positive behavioral support reduces subsequent behavioral difficulties in first graders followed into middle school. The most significant path is from achievement to behavior, so poor achievement clearly leads to behavioral difficulties (Kellam, Rebok, Mayer, Ialongo, & Kalodner, 1994).

Despite these studies of highly selected populations, meta-analyses of the relations of LDs and social skills found little evidence for specific deficits in children broadly defined as having LDs (Zeleke, 2004) or for the effectiveness of interventions addressing these problems (Kavale & Mostert, 2004) unless such a student had low self-esteem before the study began (Elbaum & Vaughn, 2003). Many of the studies analyzed in meta-analysis did not adequately control for other factors related to social skills, such as ADHD and socioeconomic status (SES). The common failure to specify the subgrouping of LDs into reading versus math disabilities is unfortunate, as there is evidence that children with math disabilities are more impaired than those with reading disorders, especially if other nonverbal processing skills are also impaired (Rourke, 1989, 1993). Other studies find that reading problems are associated with higher rates of internalizing and externalizing psychopathology, even in nonclinical samples (Willcutt et al., 2007). However, comorbid associations of reading disabilities with ADHD contributed to these relations; even comorbid reading and math disorders have higher rates of psychopathology, and comorbid disorders are also more severe. In a sample of children who responded adequately and inadequately to reading instruction in grade 1, Grills, Fletcher, Vaughn, Denton, and Taylor (2013) found higher rates of anxiety in children who had not responded adequately to instruction.

Altogether, these findings illustrate the significant need to identify and intervene early with those children who are at risk for academic failure, given the substantial social and emotional consequences that can occur if the disabilities are not remediated. The empirical evidence does not support the idea of excluding individuals from identification with LDs if they show evidence for emotional, behavioral, or social difficulties.

Economic Disadvantage

Although most current definitions of LDs state that the academic deficits encompassed by the disorders cannot be attributed to economic disadvantage and cultural factors (including race or ethnicity), limited information exists regarding how race, ethnicity, and cultural background might influence school learning in general and the expression of different types of LDs in particular. For example, Wood et al. (1991) conducted a longitudinal study of specific LDs (in reading) within a random sample of 485 children selected in the first grade and followed through the third grade (55% European American, 45% African American). Wood et al. (1991) found that the effects of race were important and complicated. At the first-grade level, once a child's age and level of vocabulary development were known, race did not provide any additional predictive power to forecasting first-grade reading scores. By the end of the third grade, race was a significant predictive factor even when the most powerful predictors—first-grade reading scores—were also in the prediction equation. By the end of the third grade, African American children were having significantly greater difficulties in learning to read. The effect is likely due to economic disadvantage.

In support of these findings, Ritchie and Bates (2013) examined the role of SES by analyzing data from the National Child Development Study, a longitudinal study of 18,588 infants born in the United Kingdom in 1958 and followed for almost 50 years. In examining the effects of reading and math achievement at age 7 on SES at age 42, better reading and math skills at age 7 had positive effects on SES at age 42. The effects were apparent even when SES at birth, IQ, academic motivation, and duration of education were included in the prediction. The later variables also had positive associations, but this study supports a reciprocal effect of SES and achievement.

In an intervention study addressing 6.5- to 8.5-year-old children significantly impaired in word reading, Morris et al. (2012) stratified their sample for race, SES, and IQ. They found that these variables were not associated with the amount of growth during the intervention or to long-term outcomes after one school year. There were no interactions with program type, indicating that the interventions worked similarly across levels of IQ, race, and SES.

In a functional neuroimaging study, Noble, Wolmetz, Ochs, Farah, and McCandliss (2006) recruited children of similar phonological processing skills who varied in SES. They found that activation in brain areas that mediate phonological processing were reduced in children with lower levels of SES, but no indication of an interaction. In French-speaking children with and without dyslexia who varied in SES, Monzalvo, Fluss, Billard, Dehaene, and Dehaene-Lambertz (2012) found that the effects of poor reading during reading and listening tasks were largely independent of SES, but brain activation was less reduced in children who were lower in SES and poor readers. Thus, SES was associated with quantitative, but not qualitative, differences in degree of activation: the same areas of the brain were involved, but activation was more reduced in association with SES.

As these studies show, many of the conditions that are excluded as potential influences on LDs interfere with the development of cognitive and language skills that lead to the academic deficits that in turn lead to LDs (Phillips & Lonigan, 2005). Parents with reading problems, for example, may find it difficult to establish adequate home literacy practices because of the cumulative effects of their reading difficulties (Wadsworth, Olson, Pennington, & DeFries, 2000). Children who grow up in economically disadvantaged environments have reduced linguistic input in the home and are behind in language development when they enter school (Hart & Risley, 1995). This delay interferes with the development of reading and math skills. Moreover, interventions that address the early development of these skills seem to promote academic success in evaluative studies of Title I programs provided to economically disadvantaged schools, as well as in intervention studies in which programs that incorporate explicit phonics instruction have been shown to be advantageous for economically disadvantaged children (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; National Insitute of Child Health and Human Development, 2000). Thus, the mechanisms and practices that promote reading success in advantaged populations appear to be similar to those that promote reading success or failure in disadvantaged populations.

There is little evidence that the phenotypic representation of LDs in reading varies according to SES. Children at all SES levels appear to have reading problems predominantly (but not exclusively) because of word-level difficulties apparent in the beginning stages of reading development (Cirino et al., 2013; Foorman et al., 1998; Wood et al., 1991). As Kavale (1988) and Lyon et al. (2001) have pointed out, the basis for excluding economically disadvantaged children from the LD category has more to do with how children are served than with empirical evidence demonstrating that characteristics of reading failure are different in groups with LDs as opposed to those who are economically disadvantaged.

Inadequate Instruction

Exclusion based on the opportunity to learn and the provision of appropriate instruction in general education makes sense if there has been no systematic effort to teach a child, but this notion is often expanded to include children whose instruction has not been adequate. Although children's failure to respond to appropriate instruction is a very strong indication of a disability, the cognitive problems associated with their LDs parallel those exhibited by children who have not had adequate instruction. The two types of children are equally disabled. Of the different exclusionary criteria for LDs, instructional factors are the least frequently examined but perhaps the most important. The opportunity-to-learn exclusion presumes that the field has a good understanding of what constitutes adequate instruction. In methods based on instructional discrepancies, consideration of the students' response to high-quality intervention is inclusionary (L. S. Fuchs & D. Fuchs, 1998). Why use the complex identification criteria and expensive due-process procedures of special education before an attempt is made to provide intervention early in a child's development?

Summary: Inclusion and Exclusion Criteria

The history of classification and definition of LD is reflected in the movement from neurological to cognitive discrepancy to instructional discrepancy classification hypotheses. Each hypothesis is different because of how unexpected underachievement is conceptualized and therefore measured. There is presently considerable tension among these frameworks because the approaches to identification and the type of individual considered as representative of LDs are different, which also has consequences for how professionals practice, the type of assessments used to identify LDs (Chapter 4), and how schools operate in balancing identification and intervention.

Despite the differences, these classification hypotheses share features, including a focus on ability or learning discrepancies, psychometric models and cut points for operationalization, and relatively narrow views of LD driven by the attributes of interest. In addition, they overlap in considering the exclusionary criteria. For a method in which cognitive discrepancies indicate the presence of LD, inadequate instruction and opportunity to learn are regarded as exclusionary criteria; in contrast, methods in which identification is in the context of RTI apply instructional response as inclusionary and do not regard cognitive discrepancies as informative.

Exclusionary criteria are often invoked automatically because of their presence in the U.S. statutory definition of LD. Definition by exclusion has not proven fruitful (Rutter, 1982). Many exclusions are more likely comorbid associations that need to be specified, especially in relation to instruction (see Chapter 4). Other exclusions are really different disorders with different intervention needs (e.g., intellectual disabilities, sensory disorders). Exactly how to consider LDs in relation to poverty or minority language status is not clear, although these certainly represent contextual factors that need to be considered. In the remainder of Chapter 3, we review the reliability and validity of different methods because issues involving identification are universal across methods and have not been adequately acknowledged by the field despite many years of investigation. In reviewing the reliability issues and the universality of identification problems with psychometric methods as presently implemented, we emphasize that our concerns about identification do not mean that we reject or have concerns about the validity of the concept of LD (see Chapter 2).

RELIABILITY OF IDENTIFICATION

Agreement across Identification Methods

What would constitute evidence for the reliability of methods for identifying LDs? Typically, reliability would be indicated by the existence of assessment methods that showed strong internal consistency and test–retest reliability. In addition, different methods should converge on which students meet criteria for LDs. At the outset, the problems do not involve an inability to reliably measure the core attributes important for different identification methods. We have highly reliable measures of aptitude, especially intelligence, academic achievement in the five domains of LDs, and methods for assessing instructional response. The problem is that different methods, even within the same conceptual model, identify different individuals with LDs. This problem is due both to the nature of the attributes of LDs, and because of psychometric factors that amplify these problems, including the slight measurement error associated with the tests used to indicate LDs. In addition, there are issues related to setting cut points, or thresholds that determine the presence and absence of LDs, or the significance of a discrepancy. Remember that these attributes of LDs are dimensional; placing firm cut points on a normal distribution also contributes to the unreliability of individual decisions.

Instructional Response

The issue of agreement has been raised most recently in the context of identification methods based on RTI, where there is often low agreement across methods based on assessments of instructional response (Barth et al., 2008; Fletcher et al., 2014; Speece & Case, 2001). Low agreement is the basis for many of the strongest criticisms of methods based on RTI (Reynolds & Shaywitz, 2009), but such problems are not unique to this framework. Any psychometric approach based on cut points will not identify the same students as inadequate responders, whether the discrepancy is based on the assessment of instructional response, low achievement, or some type of cognitive discrepancy (Francis et al., 2005).

To illustrate, Fletcher et al. (2014) compared different methods for assessing instructional response. Their identification of inadequate responders based on assessments of final status at the end of an intervention or on indices incorporating growth (slope) showed low agreement concerning which students were inadequate responders.

Given the low agreement across measures, Fletcher et al. (2014) performed a statistical simulation of agreement between the two highly reliable norm-referenced assessments of decoding and fluency, respectively: the Basic Reading composite from the Woodcock–Johnson Psychoeducational Test Battery III (WJ-III; Woodcock, McGrew, & Mather, 2001) and the Test of Word Reading Efficiency—Second Edition (TOWRE-2; Torgesen, Wagner, & Rashotte, 2011). This simulation involved creating large, normally distributed databases of "cases" based on the correlations of the measures and the thresholds for subdividing the resultant distributions to indicate LDs. If the tests were perfectly correlated and perfectly reliable, the agreement would be 1.0. However, the measures are not perfectly correlated nor are they perfectly reliable. In the empirical sample above, the two measures were highly correlated (.94). The reliabilities published by test developers are .98 for the WJ-III and .90 for the TOWRE-2.

Simulating two normally distributed, perfectly reliable variables with a correlation of .94 and a cut point at the 25th percentile yielded a chance-corrected agreement of .76, which is on the lower end of levels where agreement is considered "excellent." The slight reduction in the "perfect" correlation had a large impact on agreement. Similar problems occurred if other aspects of the simulation were manipulated. If the two measures were perfectly correlated, but used the published reliabilities, the agreement was coincidentally 0.76. If the observed correlation was 0.94 and the reliabilities of 0.98 and 0.90 were used, the simulation yielded agreement of 0.67. If differences in the normative samples of the two tests were taken into account, the agreement fell below 0.40. Because identification is often based on tests from different assessment batteries, this reduction is alarming.

Low Achievement

The simulation in Fletcher et al. (2014) is easily extrapolated to any approach to identifying LD based on a firm threshold. Note that the simulation was based on two reliable norm-referenced achievement tests. Any test could be substituted and what would vary are the reliability and the correlations. Lower reliability and lower correlations would reduce agreement, often to chance levels (Macmann, Barnett, Lombard, Belton-Kocher, & Sharpe, 1989).

Cognitive Discrepancy

The problems with identification are magnified if identification is based on a difference between two different measured attributes. Some problems are well known, including the lower reliability of a difference score (Rogosa, 1995) and the need to take into account the correlation of any two tests (e.g., IQ and achievement) in estimating a discrepancy (Macmann & Barnett, 1985).

Because the measures are correlated, simple comparisons of aptitude and achievement measures are associated with *regression to the mean* and will underidentify people with lower aptitude as "not-LD" and overidentify people with higher aptitude as LD; the failure to take into account regression to the mean and the pervasiveness of the influence of any discrepancy in cognitive skills as an indicator of LD has fueled controversy over children who appear both gifted and LD because of these discrepancies, which often are artifacts of regression to the mean (see Chapter 4).

In an early study, Macmann and Barnett (1985) observed that three factors impacted the reliability of an aptitude–achievement discrepancy: the reliability of the difference, where difference scores are generally less reliable than single test scores; selection of specific tests because of their correlations; and the location of the cut point designating presence and absence of LDs. They then simulated aptitude–achievement discrepancies at different reliabilities, intercorrelations, and cut points. When examining identification rates across different methods, the highest levels of agreement were only 50–60%; in general, one in four observations identified with LDs in the simulation were likely artifacts of measurement error, with the selection of specific observations dependent on the tests and cut points selected. Macmann and Barnett concluded that "the results of any psychometric classification procedure may be extremely tenuous, especially when the relatively limited degree of generalization across the different measures of the same construct is considered" (p. 372).

Francis et al. (2005) evaluated the stability of identifications based on aptitude–achievement discrepancy models using simulated data and actual data from the Connecticut Longitudinal Study (S. E. Shaywitz et al., 1999). This study focused on stability over time in relation to specific assumptions about the reliabilities and intercorrelations of the tests, using composite IQ and reading scores and a cut point at the 25th percentile. Simulating the effects of two assessments or modeling actual changes in the stability of classification between grades 3 and 5 in the Connecticut sample, Francis et al. (2005) reported that over 30% of children identified with LDs or as without LDs based on an IQ–achievement discrepancy changed groups just by virtue of a repeated assessment. Children close to the cut point are very similar and small amounts of measurement error significantly influence the identification of individuals.

Intraindividual Differences

Hale et al. (2010) suggested that contemporary approaches to cognitive discrepancy hypotheses demonstrate stronger reliability because the measures are improved. However, there are major issues concerning the reliability of discrepancy methods based on proposed patterns of strengths and weaknesses (PSW) in cognitive processing. In each of these approaches, LDs are indicated by a pattern of cognitive strengths and weaknesses in relation to specific academic weaknesses (Flanagan, Ortiz, & Alfonso, 2007; Hale & Fiorello, 2004; Naglieri, 2010).

Figure 3.2 shows how a PSW method works. There is a cognitive strength, a cognitive weakness, and the cognitive weakness is linked to the academic weakness. For example, if Figure 2.5 referred to an individual child who was a prototype of the group with a word-level RD, there would be a strength in problem solving and a weakness in phonological awareness; it is well established that weaknesses in phonological awareness are related to problems in reading and spelling single words. This person would be identified with a specific LD in reading; if there was no strength in problem solving, the person might have achievement difficulties, but would not be identified with LDs, and is often referred to as a "slow learner."

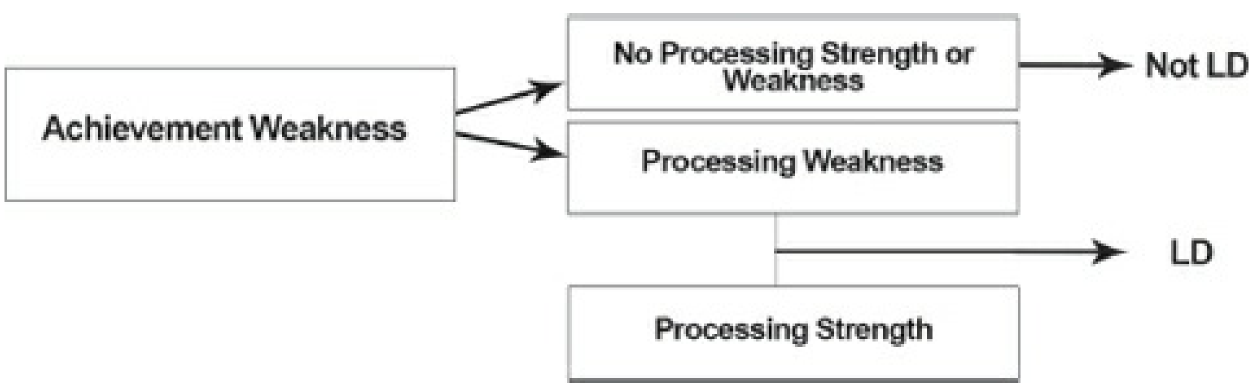

FIGURE 3.2. The relation of cognitive and academic strengths and weaknesses in a PSW identification method. Courtesy Whitney Roper.

Three PSW approaches have been proposed: (1) the concordance–

discordance method (C-DM; Hale & Fiorello, 2004), (2) the cross-battery assessment method (XBA; Flanagan et al., 2007), and (3) the discrepancy–consistency method (D-CM; Naglieri, 2010). These methods differ in how low achievement and the PSW profile are defined, and in how exclusionary factors are considered. For example, the C-DM is an ipsative (within-person) approach in which cognitive scores are used to identify a PSW profile. In contrast, in the XBA, strengths and weaknesses emerge from normative comparisons of the cognitive tests to create the PSW profile. The D-CM uses both ipsative and normative comparisons. The methods also differ in their theoretical orientation, with the XBA method selecting tests based on the Cattell–Horn–Carroll (CHC) theory of intelligence. The D-CM approach is based on the Planning, Attention, Simultaneous, and Successive (PASS) factors of intelligence measured by the Cognitive Assessment System (Naglieri & Das, 1997). The C-DM emphasizes flexibility across measures and theoretical orientations (Hale & Fiorelli, 2004).

Each of the methods uses additional criteria for determining presence and absence of LDs, such as exclusionary criteria. All three emphasize flexibility in applying the psychometric components. But our focus is on the psychometric components of the methods because of the impact on reliability. Not surprisingly, these cognitive discrepancy methods are influenced by the same psychometric issues influencing any method based on discrepancy scores or profile analysis: the reliabilities and intercorrelations of the tests and the cut points for identification.

The influence of these psychometric issues were clearly apparent in a simulation of identification by the three PSW methods (Stuebing, Fletcher, Branum-Martin, & Francis, 2012). Latent data were generated based on multiple reliabilities, intercorrelations, and cut points founded on the assumptions of each of these three methods. Observed data were generated and the concordance in identifications as LD and not-LD was assessed between simulated latent and observed levels. The results showed that all three methods were consistently biased toward not-LD decisions; only a small percentage of the simulated population (1–2%) met LD criteria. The three methods all had excellent specificity and negative predictive values, which indicate that decisions concerning the absence of LD were often accurate. However, moderate-to-low sensitivity and very low positive predictive values were observed because false positive rates for identification of LD were high. False positives in a method oriented toward the identification of cognitive strengths and weaknesses to promote alignment of cognitive processing and intervention could result in those identified as LD receiving an intervention not correctly aligned with their cognitive profile.

Even in actual data, the same problems with agreement can be observed. Kranzler, Floyd, Benson, Zaboski, and Thibodaux (2016) used the normative data from the WJ-III (Woodcock et al., 2001) to create classification decisions based on the XBA approach. Like Stuebing et al. (2012), they found a low base rate of children were identified with LDs. Accuracy was very high for "not-LD" decisions, but not for "yes-LD" decisions, reflecting a high false positive rate.

Miciak, Fletcher, Stuebing, Vaughn, and Tolar (2014a) used a sample of 139 adolescents demonstrating inadequate RTI to identify participants as meeting or not meeting PSW LD identification criteria based on the C-DM and XBA methods. As in the simulation discussed above, both approaches identified a low percentage of participants as LD (range 24–66%) despite the sample demonstrating inadequate RTI. Agreement was poor (range = −.04 to −.31), suggesting that the two approaches are not interchangeable.

Miciak, Taylor, Denton, and Fletcher (2015b) investigated the reliability of LD identification decisions using a single method (C-DM) across different achievement tests, following Macmann and Barnett (1985). Criteria based on the C-DM method were applied to assessment data from second graders who showed inadequate instructional response. The measures were equivalent at the construct level, but utilized different achievement tests (e.g., decoding represented two different subtests of the WJ-III). The two batteries identified a similarly low number of participants with LDs, but agreement was poor (.29). In addition, despite the high correlation of the two indicators of each achievement construct, the two assessments showed low agreement for identifying the academic domain representing underachievement.

Why Is Reliability of Identification a Universal Issue?

The issue of low agreement is a universal concern when identifying LDs using psychometric tests and firm cut points, whether the cut point is on a single achievement distribution or the bivariate distribution of a cognitive and achievement measure. The problem of agreement is inherent in attempts to create groups based on cut points of normally distributed variables that are not perfectly reliable, are correlated, sometimes measure different constructs, and have different normative samples. They reflect, in part, the effort to treat the attributes of LDs as categorical (yes or no) indicators, when in fact the attributes are continuous, normally distributed attributes that vary in degree, not kind (see Chapter 2). In addition, this problem reflects the facts that the attributes are correlated and not perfectly measured. *From a measurement perspective, there is no justification for policies that set firm thresholds for identification, which are common in U.S. state and district policies. Such approaches are inherently flawed because they do not take into account the measurement error and correlation of the tests and the continuous nature of the attributes of LD. Firm cut points are therefore inherently unfair in identifying individual people with LD, especially when identification is tied to access to civil rights and services.*

Thresholds and Cut Points

These are strong statements that address fundamentals of public policy, which admittedly tends to be resource-driven, with little attention to empirical realities. Many efforts to identify individuals with LD rely on setting fixed cut points, so that any person scoring below this threshold is considered to possess the attribute of LD. However, with dimensions, any threshold is somewhat arbitrary. Although few would agree that thresholds should be in the average range (> 25th percentile), exactly where in the subaverage range a threshold should be set varies considerably.

Correlation of Indices

The measures used to indicate the attributes of LDs (IQ, achievement, instructional response) are not independent. Because they are usually moderately correlated, the impact of unreliability and measurement error are magnified if multiple tests are used that do not correct for the correlation of measures. This leads directly to regression to the mean: when individuals are chosen because of low performance on one test, they will, on average, score closer to the mean on the second test.

Examples

Figure 3.3 shows the influence of firm thresholds and correlated variables, epitomized in an IQ–achievement discrepancy method. In Figure 3.3 (Fletcher et al., 2005a), the regression line that would differentiate those with and those without LDs is steeper for Verbal IQ than for Performance IQ because of the higher population correlation of reading (.69) and Verbal IQ than for reading and Performance IQ (.40). The difference in slopes and in measures shifts individuals at the edges of the regression cut point on one IQ measure to either a discrepant or low-achieving subgroup when the other IQ measure is used. Because the correlation of IQ and reading is lower, effect sizes would be larger for Performance IQ than for Verbal IQ (see Fletcher et al., 1994). Nonetheless, collapsing across IQ-discrepancy and low achievement definitions, 80% of the sample is consistently identified as LD, simply shifting from one LD subgroup to another. Changing the IQ measure moves the observations left or right across the cut point, but does not move them up or down because the achievement measure is the same. These shifts are displayed in Figure 3.3 by a line that connects pairs of observations. An observation that does not change in the identified group has the same symbol connected by a faint horizontal line; observations that change groups have two different symbols that are connected by a dark horizontal line. As Figure 3.3 shows, observations with IQ scores that are most different and that are located near the cut point are most likely to shift, reflecting both measurement error and differences between how the construct of aptitude is assessed by Verbal and Performance IQ.

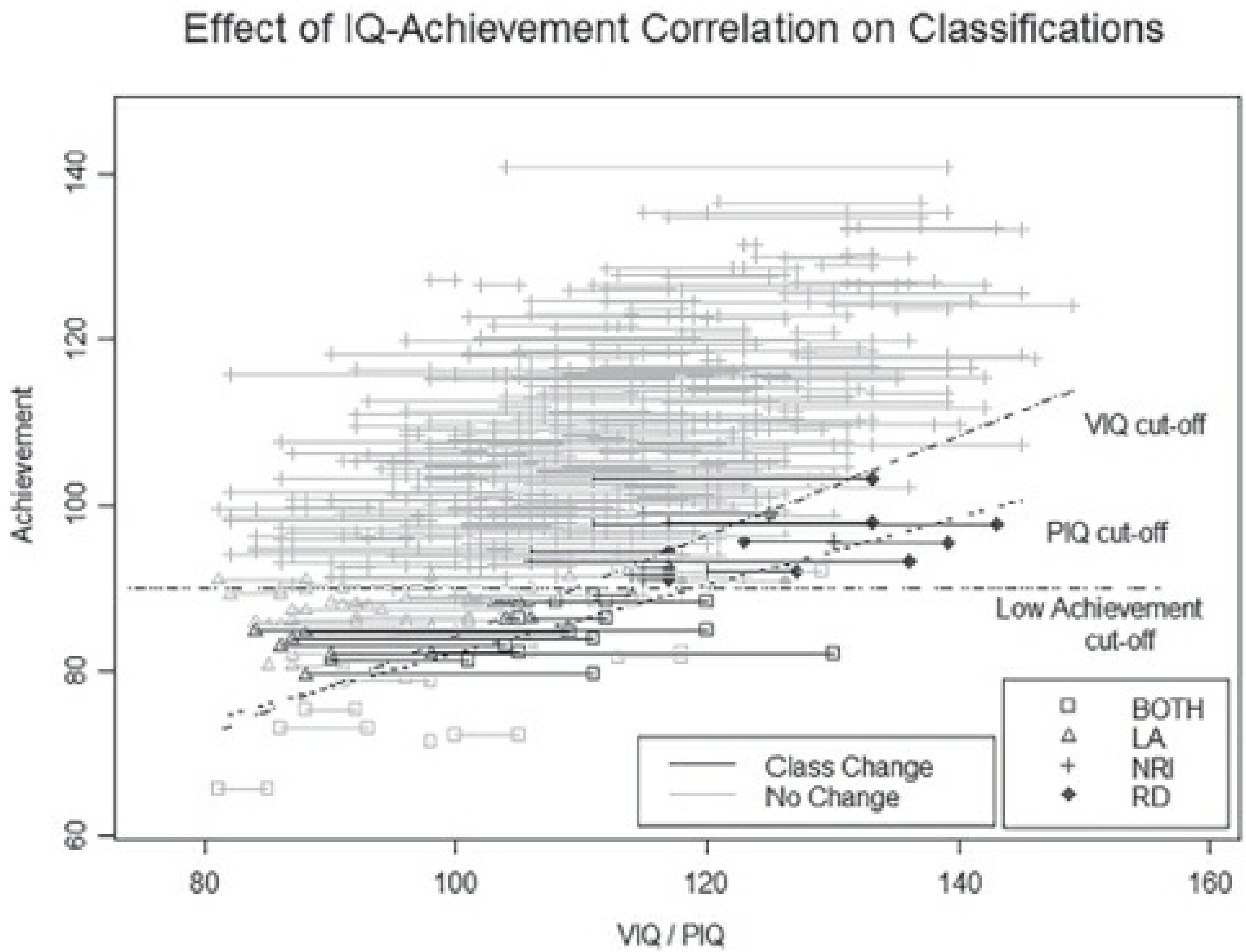

FIGURE 3.3. Regression lines based on the population correlations of the WJ-III Basic Reading Skills score with Performance IQ (PIQ) and Verbal IQ (VIQ). Higher correlations are reflected in steeper slopes, so that different decisions about group membership are made because of slight shifts in slopes. Individual observations are connected and show significant movement around the cut points demarcating those who meet both low-achievement and discrepancy (BOTH) definitions, only low-achievement (LA) definitions, only discrepancy (RD) definitions, and not reading impaired (NRI). From Fletcher, Denton, and Francis (2005a, p. 548). Copyright © 2005 PRO-ED. Reprinted by permission.

Figure 3.4 shows these relations in a different manner. Here we have a cut point for levels of intelligence and for low achievement. In addition, a regression line demonstrating a cut point for an IQ–achievement discrepancy is shown. The levels of IQ are arbitrary, reflecting decisions researchers make about the level of IQ to exclude children from a study of LDs. These decisions do influence decisions about the severity of the reading problem because IQ and achievement are correlated. If we put a line for an IQ cut point at 70, it would correspond to the traditional threshold for an intellectual disability, but would not change the message about cut points. The low achievement cut point of 90 is often used in research, but is at the 25th percentile and more liberal than many schools might use. The curved line is the 1.5 standard error regression threshold that reflects an IQ–achievement discrepancy. This line could represent any effort to measure a two-test discrepancy because the measures are correlated and we assume that the method accounts for the correlation of the measures. The numbers represent actual children referred for a study of LDs and the space they occupy in this bivariate, normal space. In fact, the cloud-like appearance of the dispersion represents a normal bivariate distribution in two-dimensional space, highlighting the dimensionality of these attributes. Children who are demarcated with a "6" meet both an IQ–achievement and a low achievement definition. Children with a "7" meet an IQ–achievement discrepancy definition, but not a low achievement definition because their reading scores are above 90. Children with a "2" do not meet any definition of LDs. A "3" meets the low achievement definition, but not the IQ–achievement discrepancy definition. A "1" is excluded because of the IQ cut point at 80. Again, note the close proximity of many children to the cut point. An adjustment of the low achievement threshold to 85 would reclassify several children identified as "3" into the "2" category. More importantly, what is the difference between a 2 and a 3 when they are clustered around the cut point; or a 1 and a 3? Because of measurement error, giving the same tests twice or different, highly correlated measures of IQ and achievement from different assessments would shift over half the sample. In research, decisions about these cut points directly influences the size of a group difference depending on the correlation of the identification and outcome variables.

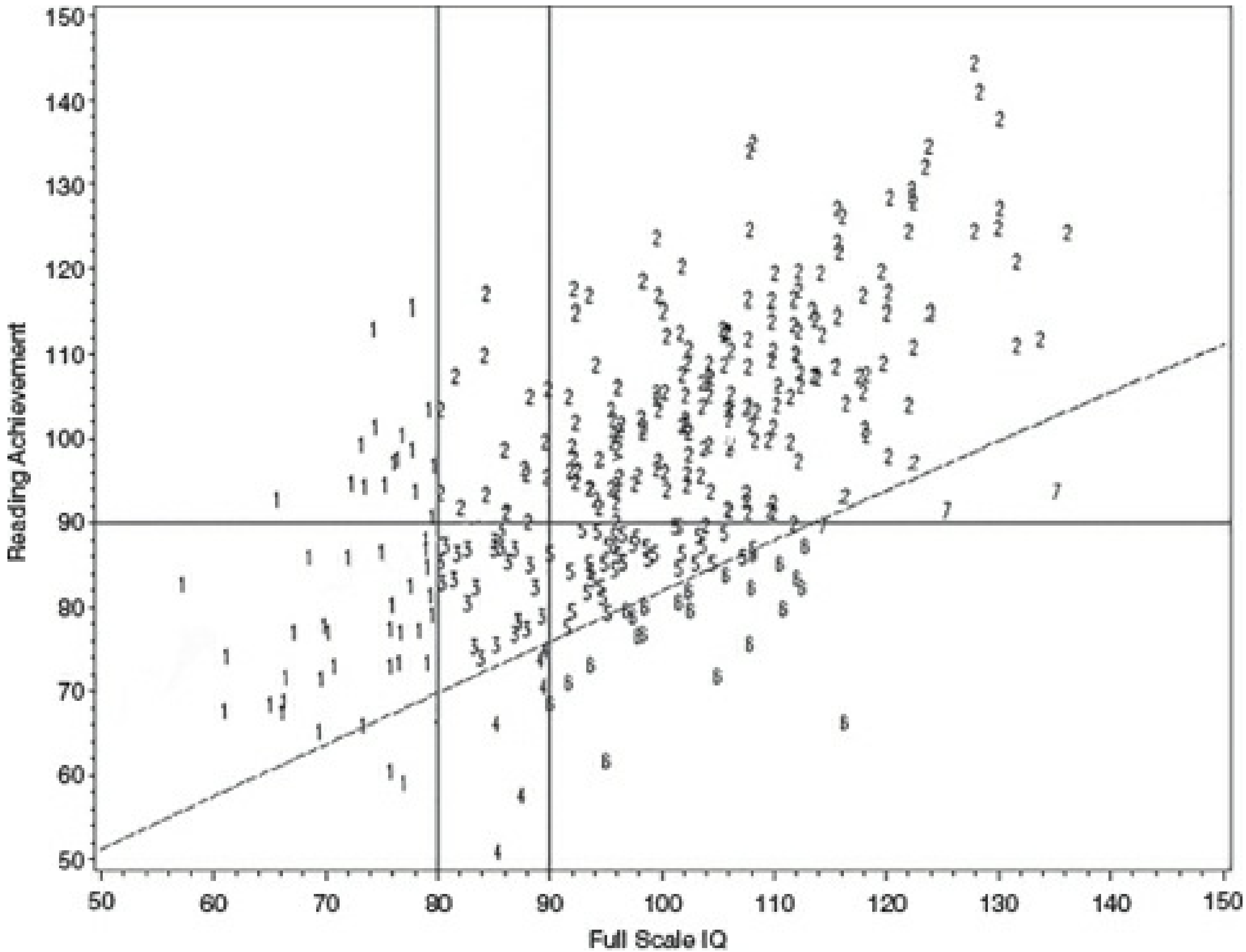

FIGURE 3.4. Individual children recruited for a study of LDs using different definitions. The scores of each child on Full Scale IQ from the Wechsler Intelligence Scale for Children—Revised (WISC-R) and the WJ-III Basic Reading Skills are displayed as numbers, reflecting clusters of children who meet and do not meet different definitions of a reading disability. Note the children who cluster near a line representing an IQ cut point, a low achievement cut point, or a 1.5 standard error cut point. Courtesy Karla Stuebing and Whitney Roper.

Solutions

There are solutions to these problems. For example, as when assessing for the presence of intellectual disabilities (Schalock et al., 2010), eligibility could be expressed as a range of scores based on the standard error of measurement of a test, that is, as a confidence interval. Another approach is to propose multiple inclusionary criteria or to assess the construct of interest using two or more assessments, with membership in the class (i.e., eligibility) requiring scores below the threshold on multiple indicators. A third would be to move toward different types of psychometric approaches, such as a Bayesian approach that captures multiple indicators that might include more than psychometric test scores. For example, gender and family history could be included and membership in the LD class would be expressed as a probability figure instead of an absolute yes-or-no decision. This approach could be tied to intervention response (Spencer et al., 2014b). Finally, a fourth approach would be to recognize that assessments of initial status for individual people are fraught with errors and to reconceptualize class membership (eligibility) as a recursive process in which the system constantly revaluates the decision-making process (Macmann et al., 1989). In many respects, this is accomplished in methods based on RTI because of the emphasis on screening, progress monitoring, and ongoing evaluations of intervention response.

VALIDITY OF METHODS FOR CLASSIFYING AND DEFINING LD

In this section we focus on studies that represent explicit tests of classification hypotheses Note that this review of validity is based on research that occurs at the group level and thus does not directly involve the issue of individual identification decisions. This means that the method is based on a hypothesis about the underlying inclusionary attributes. Groups include people who meet and who do not meet definitions of LD based on the hypothesis. As we discussed in Chapter 2, if the hypothesis is valid, it should be possible to differentiate the groups based on comparisons on variables not used to define the groups (Morris, 1988). Some may argue that these types of group comparisons do not do justice to methods based on intraindividual differences, where each case is unique, but scientific research always permits disconfirmation and exactly how classification hypotheses are falsifiable when these claims are made is not clear, especially when the proponents generate formulae for large-scale implementation. We begin with low achievement methods, then turn to cognitive discrepancy methods, and conclude with methods based on RTI and the MTSS service delivery framework.

Low-Achievement Methods

The default definition of LD is to specify simple low achievement and ensure that there are not other conditions that explain why achievement is low. We describe this approach as a "low-achievement" method because a typical operationalization would be based on absolute low achievement. For example, a reading score below the 25th percentile may indicate a LD in reading (Siegel, 1992).

A low-achievement approach does not ignore the importance of cognitive processes as factors that represent correlates of LD. The external validity of subgroups of LD based on the level and pattern of academic underachievement has long been supported (Pennington, 2009; Rourke & Finlayson, 1978). These studies, which most commonly compare children with disabilities in reading, math, and both reading and math, show that all forms of LD are not the same on a wide range of cognitive and other attributes not used to form the groups. As such, these studies support the heterogeneity of LDs and the need to tie LDs to specific domains of academic functioning. These subgroups extend to variations in reading disability, where children can be differentiated by patterns of strengths and weaknesses in word recognition, fluency, and comprehension. In fact, the strongest evidence for the validity of the concept of LD stems from the association of different cognitive processes and different achievement domains (see Figure 2.5). In a similar vein, there are neurological and genetic factors that are associated with LD that are not used for identification, but are critical for scientific understanding of LD. Low-achievement models conceptualize LD from an age-related achievement discrepancy model, and are really examples of instructional discrepancy methods.

Perhaps most importantly, there are interventions for each of these domains. Because people vary in their strengths and weaknesses across academic domains, there are clearly group-by-treatment interactions because some people with LDs need reading instruction focused on decoding, others on comprehension, and some in both domains; some people achieve adequately in reading, but not in math or written expression.

In the area of reading, the presence of group-by-treatment interactions has been dramatically demonstrated by Connor et al. (2009), who measured child attributes involving reading decoding and comprehension. They showed that helping teachers vary the amount of code-based versus meaning-based instruction according to student weaknesses in decoding versus comprehension led to better outcomes compared to classrooms in which this assessment information and assistance were not provided. Thus, whereas assessing cognitive processes for intervention purposes may not be associated with qualitatively distinct cognitive characteristics, assessment of reading components and other academic skills is justified because of the evidence for group-by-treatment interactions. Altogether, these studies epitomize how the validity of a classification can be established and support low achievement as a necessary criterion for identification of LD, representing a well-validated inclusionary criterion.

Simple low achievement departs from the original concept of "unexpected underachievement" because the group identified with LDs would include children with low achievement due to a variety of factors typically considered exclusionary. However, exclusionary criteria like those discussed above could be added to the definition, so the primary inclusionary criteria could be low achievement and unexpectedness indicated by absence of exclusionary criteria. This approach is still a definition based solely on low performance and exclusions that may not be sufficient to establish "unexpectedness."

Another problem is the level of performance that constitutes low achievement. In research studies on reading, many use the 25th percentile. In some math research, it has been proposed that performance below the 10th percentile constitutes "math LD" and the 10–25th percentile represents math low achievement, with evidence of cognitive profile differences that are difficult to untangle from the definitions and cut points (Geary, Hoard, Nugent, & Byrd-Craven, 2008b; Mazzocco & Myers, 2003). These definitions are also based on longitudinal research on the stability of intraindividual cognitive profile differences over time. As we discussed above, a problem with this research is the instability associated with cut points of any kind, partly because even highly reliable achievement tests have measurement error (see Figures 3.3 and 3.4). In addition, the attributes appear dimensional, so why differences in profiles would be expected by virtue of threshold differences (except that elevation and shape become decoupled further down the distribution) is not clear. The validity of these distinctions has not been strongly demonstrated (Tolar, Fuchs, Fletcher, Fuchs, & Hamlett, 2016), especially if level of severity is considered a marker of LD and low achievement in math.

At a practical level, the selection of a cut point implies very specific assumptions about prevalence. If we select the 25th percentile, and exclude 2% for intellectual disability, and some unknown proportion because of exclusionary criteria, the figure seems high for the number of children with LD, yet some studies report prevalence figures as high as 17.4% of the population for dyslexia (Shaywitz et al., 1992), operationalized as a word-level disorder (Chapter 6). Moreover, if we select people based on cut points across five achievement domains, there will be overlap because people may have achievement deficits in more than one domain, so the prevalence will be much higher than 25%. We can lower the cut point, but where should it be set? Any decision is potentially arbitrary in the absence of research relating thresholds to adaptational difficulties.

Cognitive Discrepancy Methods

Aptitude–Achievement Discrepancy

Methods based on aptitude–achievement discrepancies stem from Rutter and Yule (1975), which indicated that the presence of a severe discrepancy between IQ and achievement may be an indicator of a specific LD in reading. In that study, exclusionary criteria were not applied and many children identified as "backwards readers" (i.e., poor readers with no specific LD) were brain-injured, with low IQ scores (Fletcher et al., 1994). Since this study, there have been many efforts to validate this two-group hypothesis of differences in IQ-discrepant and low-achieving poor readers. In Figure 3.4, this would involve comparisons of people who would be placed into different parts of the bivariate space shown in the figure (essentially groups 1 and 6 depending on the definition).

Meta-Analyses of IQ–Achievement Discrepancy

Two meta-analyses of the cognitive and achievement correlates of LD in reading have been completed. Hoskyn and Swanson (2000) identified 69 studies conducted from 1975 to 1996, coding 19 that met stringent IQ and achievement criteria. Effect sizes were computed to compare groups of students with higher IQ and poor reading achievement (IQ-discrepant) and students with both lower IQ and poor reading achievement (low achievement; LA). They reported negligible to small differences on several measures of reading and phonological processing, but larger differences (IQ-discrepant > LA) on measures of vocabulary and syntax.

Stuebing et al. (2002) independently identified 46 studies from a sample of over 300 from 1973 to 1998. These studies included measures of behavior, academic achievement, and cognitive abilities. From these studies, effect sizes were computed for cognitive, behavioral, and achievement domains. The effect sizes estimates were negligible for behavior and achievement. A small effect size showing higher aggregated performance in the IQ-discrepancy group was found for cognitive ability. As in Hoskyn and Swanson (2000), cognitive abilities (e.g., phonological awareness, rapid naming, verbal memory, vocabulary) most closely related to reading yielded negligible effect sizes. Cognitive skills like those measured by nonverbal IQ subtests (spatial cognition, concept formation) yielded small-to-medium effect sizes, also indicating higher scores by the IQ-discrepant group. Altogether the difference across the 46 studies in cognitive ability was about 0.3 standard deviations, demonstrating substantial overlap between the groups on phonological, language, and nonphonological tasks. Stuebing et al. (2002) also reported that variation in effect sizes across studies could be modeled by the scores on the IQ and reading tasks used to define the groups (i.e., sampling variation across studies) and the correlation of these definitional variables with the tasks used to compare the two groups. Thus, variation in effect sizes largely reflected differences in how groups are formed.

These meta-analyses concur in questioning the role of discrepancies in IQ and achievement as an indicator of LDs. However, they had different interpretations of the role of IQ for identifying LDs. Stuebing et al. (2002) questioned the relevance of IQ for identification, while Hoskyn and Swanson (2000) observed that IQ was related to different indicators of LDs. Swanson (2013) summarized several additional analyses showing that variations in IQ mediated instructional outcomes, especially at lower levels of reading. He concluded that “variations in IQ and reading cannot be ignored when predicting treatment outcomes and are therefore a critical ingredient to the identification process" (p. 638). Our view is that in order for any variable to be important for identification, it must add value to direct assessments of academic skills; that an outcome is correlated with or even mediates outcome is not sufficient evidence to justify assessments for identification purposes. As we see below, evidence for a role of IQ or other cognitive skills in identification or prediction of treatment outcomes is limited, but remains an important area for investigation.

Prognosis and Long-Term Development

For prognosis, null results are apparent in multiple longitudinal studies (Flowers, Meyer, Lovato, Wood, & Felton, 2001; Share, McGee, & Silva, 1989; Vellutino, Scanlon, & Lyon, 2000). To illustrate, Francis, Shaywitz, Stuebing, Shaywitz, and Fletcher (1996) examined long-term development of children defined in grade 3 as meeting an IQ–achievement discrepancy or a nondiscrepant low achievement definition in reading. No differences between groups were apparent in kindergarten–grade 6. Shaywitz et al. (1999) subsequently extended the analysis through grade 12 by using the same cohort and methods (Figure 3.5).

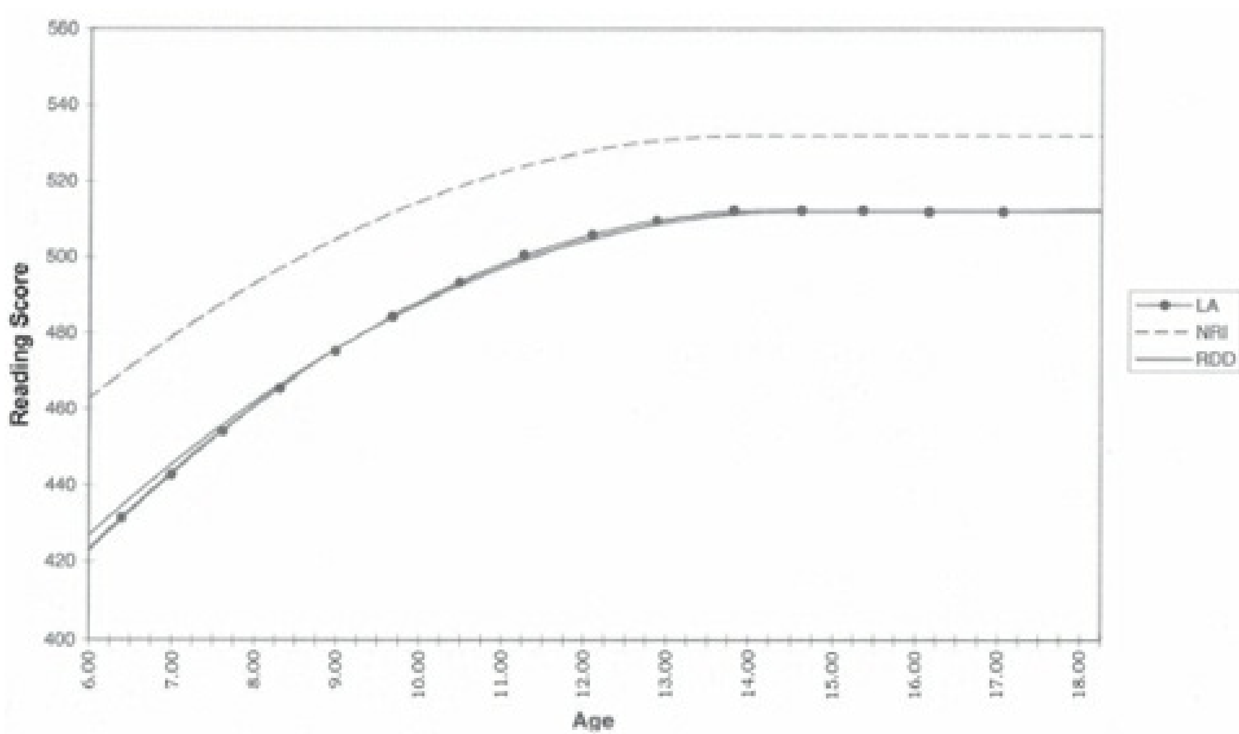

FIGURE 3.5. Growth in reading skills by children in grades 1–12 in the Connecticut Longitudinal Study based on the reading cluster of the Woodcock–Johnson. The children were identified at grade 3 as not reading impaired (NRI); reading disabled according to a 1.5 standard error discrepancy between IQ and reading achievement (RDD); or having low reading achievement with no discrepancy (25th percentile; LA). There is no difference in the long-term growth of the RDD and LA groups. From Fletcher et al. (2002, p. 193). Copyright © 2002 Lawrence Erlbaum Associates. Reprinted by permission; permission conveyed through Copyright Clearance Center, Inc.

Intervention Outcomes

Several studies have examined the outcomes of reading interventions in relation to different indices of IQ or IQ–achievement discrepancy (see Swanson, 2013), leading to a meta-analysis by Stuebing, Barth, Molfese, Weiss, and Fletcher (2009) of 22 studies that addressed the relation of different assessments of IQ and intervention response. IQ accounted for less than 1% of the unique variance in intervention outcomes. The aggregated effect sizes were not moderated by the type of IQ measure, age, or reading outcome. Simulations of the capacity of variables with effect sizes in this range to predict intervention response yielded little evidence of practical significance.

Neuroimaging Studies

Tanaka et al. (2011) compared two different samples of children identified as IQ–achievement discrepant and low achieving in reading in an fMRI paradigm. The task involved reading of real words and pseudowords. No differences were found in the activation patterns associated with word reading between the two samples. Tanaka et al. concluded that

> poor readers with discrepant or non-discrepant IQ scores exhibited similar patterns of reduced brain activation in brain regions including left parieto–temporal and occipito–temporal regions. These results converge with behavioral evidence that poor readers have similar kinds of reading difficulties in relation to phonological processing regardless of IQ. (p. 1442)

In a subsequent study, Simos, Rezaie, Papanicolaou, and Fletcher (2014) used magnetic source imaging to examine the relation of both verbal and nonverbal IQ with the brain activation patterns of children experiencing word-level reading difficulties that met or did not meet the IQ–achievement discrepancy criterion. In addition, comparisons were made to typically developing children who had higher and lower IQ scores. There was no evidence of differences in the degree of activation in reading-related brain areas based on the presence or absence of an IQ–achievement discrepancy or according to level of IQ in poor readers, although these readers were reliably differentiated from typically developing children.

In a later study, Hancock, Gabrieli, and Hoeft (2016) used a subset of the children in the Tanaka et al. (2011) study, subdividing them into readers with high IQ scores and discrepant reading relative to IQ and comparison groups of typical readers with no discrepancy, and a group of poor readers with no IQ–achievement discrepancy comparable in either IQ or reading level. Activity was comparably reduced in the IQ–achievement discrepant and nondiscrepant readers in the middle temporoparietal region relative to both comparison groups, leading the authors to suggest that the two groups share "atypicality" in this region. The study clearly warrants additional evaluation in larger samples, particularly since discrepancy was defined post hoc based on a discrepancy of about 9.2 standard score points between a vocabulary score and a word-reading accuracy measure. Many of the high IQ–achievement discrepant children had lower reading fluency scores. We return to the issue of defining "gifted" children with LDs in Chapter 4, where we observe the many problems with this concept.

Heritability Studies

The IQ–achievement discrepancy hypothesis has been addressed in behavioral genetic studies of reading disabilities. Although Pennington, Gilger, Olson, and DeFries (1992) found little evidence for differential heritability based on definition into IQ-discrepant or low-achieving poor readers, a subsequent study with a larger sample (Wadsworth et al., 2000) subdivided the twin pairs into groups with and without reading disabilities according to higher (> 100) and lower (< 100) IQ scores. The overall heritability of reading skills was 0.58, but varied according to level of IQ: the lower IQ, reading-impaired group had a heritability estimate of 0.43, whereas the estimate was 0.72 for the higher IQ, reading-impaired group. These differences in heritability are statistically significant, but are small; almost 400 twin pairs were required to detect the difference.

Alternative Approaches to Measuring Aptitude

Are other indices of IQ or assessments of listening comprehension better measures of aptitude? Some have advocated for the use of nonverbal IQ measures (e.g., Performance IQ; PIQ) because this type of measure is less confounded by language, and many students with LDs have language difficulties (Rutter & Yule, 1975). Alternatively, Hessler (1987) suggested that a verbal measure of IQ was a better aptitude assessment because difficulty in learning to read should represent a discrepancy relative to language potential. Here the distinction is essentially between students who do not learn to read despite adequate verbal skills, and those whose reading difficulties are part of a constellation of language problems. Finally, others have argued that a listening comprehension measure is a better index of aptitude for learning to read because a reading disability should represent a discrepancy between listening comprehension and reading comprehension (Spring & French, 1990).

Many of these decisions will be reflected in the correlation of aptitude and achievement. Note the differences in the regression lines in Figure 3.3 for Verbal IQ and Performance IQ with reading. Performance IQ has a lower correlation with reading compared to Verbal IQ, so more students would be identified as discrepant. But the conceptualizations of aptitude are different, so this could affect the validity of the classifications. Not surprisingly, given the psychometric problems that emerge with any effort to use discrepancy scores, there is no support for the greater validity of these approaches. Fletcher et al. (1994) and Stanovich and Siegel (1994) found only slight differences in the magnitude of effect sizes in relation to word recognition discrepancies. Like Aaron, Kuchta, and Grapenthin (1988) and Badian (1999), Fletcher et al. (1994) found small differences between discrepant and nondiscrepant poor readers based on a discrepancy between listening comprehension and reading comprehension (effect size = 0.20). In contrast, Spencer, Quinn, and Wagner (2014a) found that a discrepancy with listening comprehension contributed unique variance to a formula predicting reading disability, but the listening comprehension measure was actually a vocabulary assessment. The use of listening comprehension requires better assessment of the construct.

IQ–Achievement Discrepancy and Other Disabilities

Math LDs

In the area of math LDs, results are similar for reading LDs, with little evidence for differences in math calculation or problem-solving cognitive profiles of discrepant and nondiscrepant poor readers (Fletcher et al., 2005a; Mazzocco & Myers, 2003; Tolar et al., 2016).

Specific Reading Comprehension Disabilities

There are few studies that use IQ–achievement discrepancies to define groups with specific reading comprehension disability (SRCD) and the issue of IQ or IQ–achievement discrepancy has had little impact on research on SCRD. As we discuss in Chapter 8, more general verbal processing difficulties underlie SCRD, highlighting the difficulties that would emerge if IQ were controlled in studies of poor comprehenders. Even in typically achieving readers, Verbal IQ accounts for only a small amount of the variability in reading comprehension skills (Oakhill, Cain, & Bryant, 2003).

Speech–Language Impairments

The federal definition of LDs includes disorders of oral expression and listening comprehension. These disorders can also be represented as disorders of expressive and receptive language, which constitute a separate category in special education under IDEA. Tomblin and Zhang (1999) evaluated the role of IQ in children with oral language disabilities and found little evidence for differences in relation to IQ–achievement discrepancy in a large epidemiological sample. They concluded that "current diagnostic methods and standards for specific language impairment do not result in a group of children whose profiles of language achievement are unique" (p. 367). A consensus group convened by the National Institute of Deafness and Communication Disorders concluded that the practice of using IQ scores to identify children with these disorders was not supported by research and practice (Tager-Flusberg & Cooper, 1999).

Summary: Aptitude–Achievement Identification Methods

Table 3.1 summarizes the major issues that have emerged from research on aptitude–achievement discrepancies. Acceptance of the null hypothesis of no differences is always a difficult inference to support. Moreover, there are some studies that demonstrate statistically significant differences between IQ–achievement discrepant and low-achieving groups. However, these effects are generally small and of questionable practical significance because they do not suggest differences in treatment needs and outcomes. Moreover, there is little evidence that IQ scores are useful for planning instructional programs for children with LDs (Elliott & Resing, 2015). In a situation where weak validity accrues over multiple comparisons, the question is whether there are more fundamental problems with the classification framework underlying the hypothesis.

TABLE 3.1. What's Wrong with IQ–Achievement Discrepancy?

1. IQ–achievement discrepant and nondiscrepant low achievers do not differ practically in behavior, achievement, cognitive skills, response to instruction, and neurobiological correlates once definitional variability is controlled (Stuebing et al., 2002). The classification lacks validity.
2. IQ is a weak predictor of intervention response, especially if baseline academic skills are in the model (Stuebing et al., 2009).
3. There is little evidence of difference in brain activation profiles (Tanaka et al., 2011; Simos et al., 2014).
4. Status methods for identification may not be reliable or stable based on a single assessment or rigid cut point (Macmann & Barnett, 1985, 1997; Macmann et al., 1989; Francis et al., 2005).

Intraindividual-Difference Methods: PSW

There is little empirical research by the proponents on the validity of PSW methods (Schneider & Kaufman, 2017). Claims about validity are often restricted to single case studies or cluster analyses where poorly validated subtypes emerge. In more recent studies, Miciak et al. (2014b) compared low-achieving children with reading problems identified with LDs and as "slow learners" on achievement tests not used to define the groups. As Figure 3.6 shows, there was little difference in the shape or elevation of the achievement profiles generated by four different operationalizations of PSW methods.

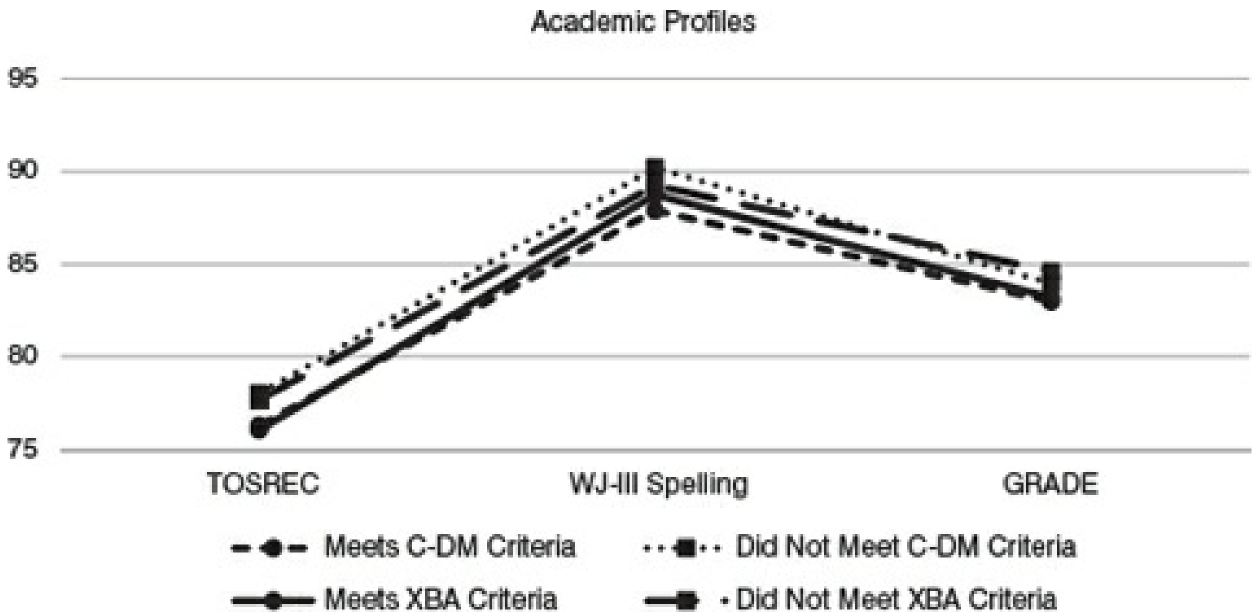

FIGURE 3.6. Comparison of achievement profiles not used to define groups of children with specific learning disability and slow learners in two PSW methods. There are no significant differences in the shape or elevation of the achievement profiles. TOSREC, Test of Silent Reading Efficiency and Comprehension; WJ, Woodcock–Johnson; GRADE, Group Reading Assessment and Diagnostic Evaluation. Data from Miciak, Fletcher, Stuebing, Vaughn, and Tolar (2014a). Courtesy Whitney Roper.

In another study, Miciak et al. (2016) used a large intervention database with extensive assessments of cognitive functions to determine if identification of a child as LD or not LD under the C-DM or the XBA improved the prediction of treatment outcomes. Figure 3.7 shows the design of the study from pretest to intervention and the application of these PSW methods. There was little evidence of value-added increments relative to baseline assessments of reading skills. Individual cognitive assessments in the absence of the application of PSW methods also did not contribute significantly to the prediction of treatment outcomes relative to baseline measures of reading. Consistent with these results, in a meta-analysis Stuebing et al. (2015) found that different cognitive measures explained extremely small amounts of growth in RTI when initial status in reading skills was taken into account.

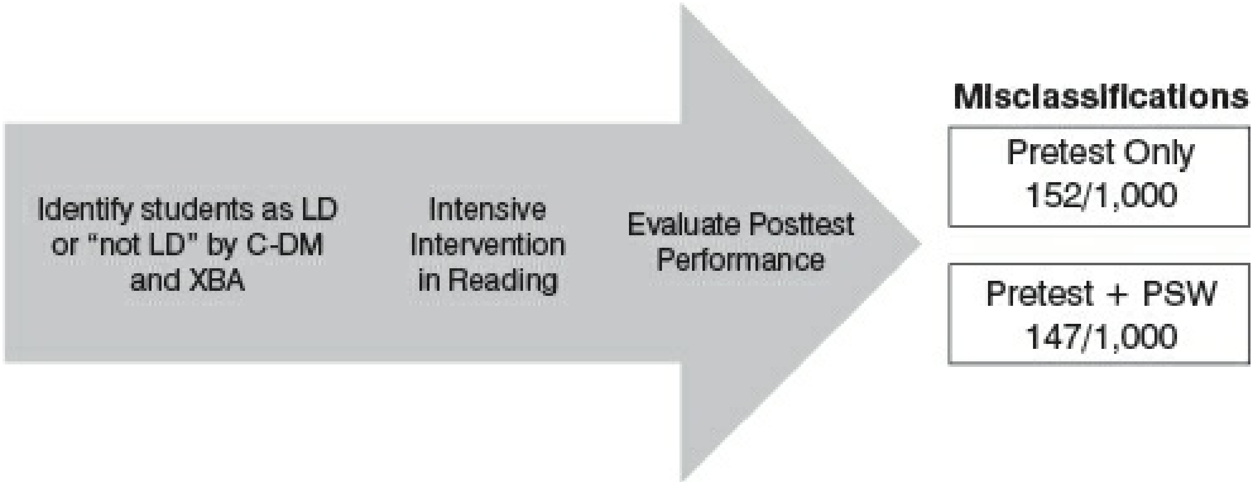

FIGURE 3.7. Design of Miciak et al. (2016), showing pretest, treatment, and outcomes. Status as reading disabled versus reading impaired, but not learning disabled based on a PSW method did not significantly increase the prediction of treatment outcomes relative to baseline pretest reading levels. Courtesy Whitney Roper.

It is ironic that PSW methods are proposed when the basic psychometric issues and shortcomings are well understood and have been documented for many years. In studies of profile analysis based on the Wechsler intelligence scales, little evidence has emerged linking LDs to patterns of strengths and weaknesses (Kavale & Forness, 1984; Watkins & Canivez, 2004). In a simulation, Macmann and Barnett (1997) evaluated differences in Verbal IQ and Performance IQ factor index scores and ipsative profile patterns on the Wechsler Intelligence Scale for Children—Third Edition (WISC-III), reporting that the reliability was poor and that practitioners should not use the results for making identification decisions. The measurement issues make any method based on cognitive discrepancies unlikely to achieve reasonable levels of reliability. Nonetheless, there are still advocates for the use of Wechsler scale performance patterns for identification of LD (Hale & Fiorello, 2004).

Finally, advocates of PSW cite interactions of cognitive process treatment and academic outcomes as evidence for the validity of these methods, arguing that the specification of a cognitive profile is necessary to understand how to design an effective intervention for the person. However, little evidence supporting the effectiveness of interventions based on cognitive-process profiles has emerged, much less for interactions of cognitive processes and interventions to influence treatment outcomes (Mann, 1979; Kearns & Fuchs, 2013; Pashler, McDaniel, Rohrer, & Bjork, 2009). In a recent meta-analysis, Burns et al. (2016) examined the role of cognitive tests in relation to intervention. Across different uses (screening, intervention design), there was an effect of cognitive tests and outcomes (g = .17), which was much smaller than the effect of reading fluency (g = .43) and phonological awareness (g = .48).

Table 3.2 summarizes the problems that have emerged in research on PSW methods. Again, there is weak validity for this general approach to cognitive discrepancy. As Elliott and Resing (2015) stated:

> Current evidence indicates that cognitive measures have limited relevance for instructional planning, and cognitive programs have yet to show sufficient cognitive gains. . . . Our energies should be devoted to the continuing development of powerful forms of academic-skills-based instruction operating within a response to intervention framework. (p. 137)

TABLE 3.2. Problems with PSW Approaches to Identification

1. The statute does not mandate that cognitive skills be assessed—just their manifestations.
2. Proponents have conducted little research on how PSW methods actually work and are related to instruction (Schneider & Kaufman, 2017).
3. PSW is predicated on a straw person view of RTI. There is no standalone RTI identification method and a comprehensive evaluation is always required regardless of the identification method.
4. Psychometric issues with discrepancy scores of any kind are well known, especially the use of rigid cut points, profile interpretations, and difference scores (Francis et al. 2005; Stuebing et al., 2012).

Methods Based on RTI

The differences between a traditional method based on cognitive discrepancy or low achievement versus a method embedded in the RTI method are presented in Figure 3.8. Identification through an RTI method that incorporates instructional response moves from a traditional refer, test, and treat model (left side) to one based on an MTSS. The MTSS service delivery framework (see Chapter 5) involves screening, introduction of increasingly intensive interventions as a series of tiers beginning in the general education classroom, progress monitoring, and repeated assessment to identify inadequate responders at each level of intervention (right side). From a classification perspective, the identification of a child considered as having an LD within a MTSS framework focuses on evidence of inadequate instructional response, which is an inherent inclusionary characteristic. Thus, children with all types of LDs share intractability to instructional programs that are effective with most children.

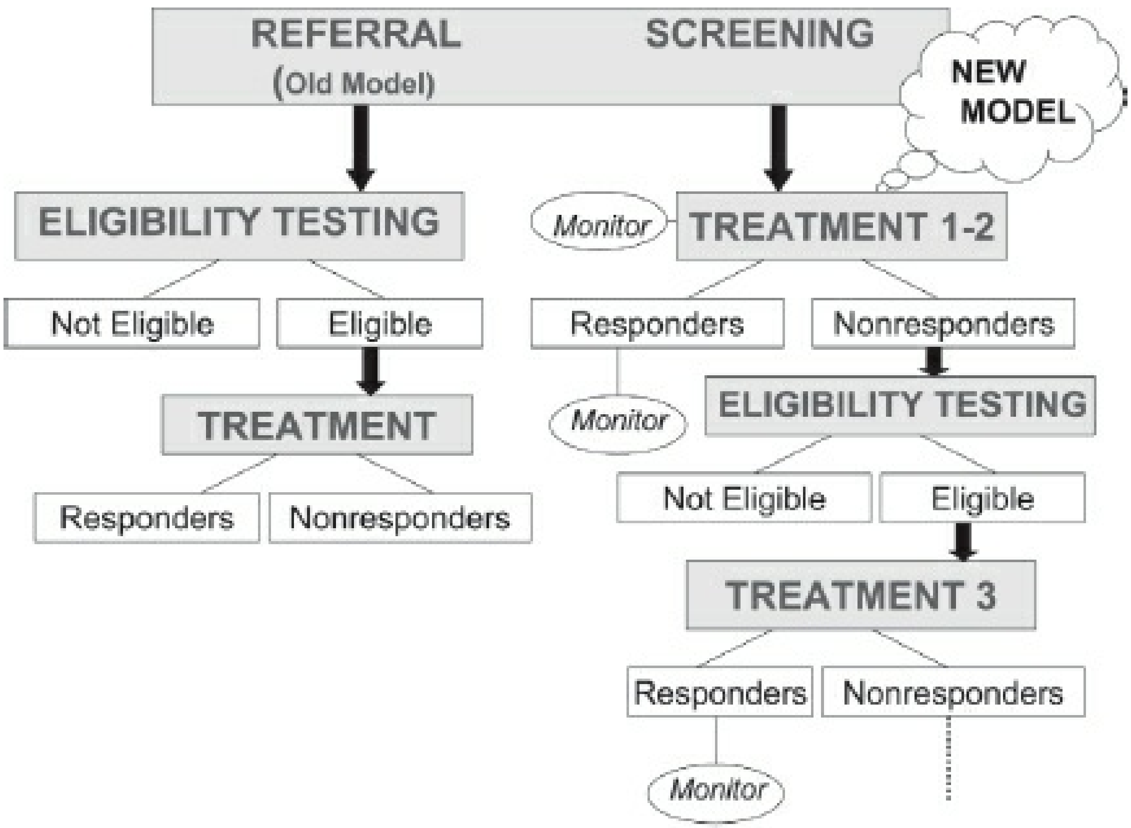

FIGURE 3.8. A comparison of a traditional referral and assessment model and a model based on RTI. On the left-hand side, the student is typically referred for an eligibility evaluation. The student is either eligible or not eligible; if eligible the student receives intervention that is evaluated every 1–3 years. In a method based on RTI, all children in a service delivery framework based on an MTSS would be screened; those at risk receive progress-monitoring assessments and immediate intervention. If there is inadequate response to different interventions, a comprehensive evaluation may occur. Courtesy of Maureen Dennis.

Empirical studies suggest that classifications based on differential intervention response consistently differentiate groups on a number of characteristics, including academic level (Al Otaiba & Fuchs, 2006; Nelson, Benner, & Gonzalez, 2003; Vellutino, Scanlon, Small, & Fanuele, 2006), cognitive characteristics (Fletcher et al., 2011; Miciak et al., 2014b), behavior

(Al Otaiba & Fuchs, 2006; Nelson et al., 2003), and brain activation patterns (see Chapter 6; Molfese, Fletcher, & Denton, 2013; Rezaie et al., 2011a, 2011b). Such data provide evidence for the validity of intervention response as a classification attribute because they show that subgroups of adequate and inadequate responders can be differentiated across attributes not utilized for group formation. The fundamental question is whether group separation between adequate and inadequate responders reflects differences in the *level* of performance or differences in the *pattern* of performance. A difference in pattern would suggest that the subgroups are qualitatively different, whereas differences in level of performance on related attributes represents differences in severity. Either pattern lends validity to the classification.

The findings are generally consistent with a continuum-of-severity hypothesis (Vellutino et al., 2006), in which achievement and achievement-related abilities lie on a continuum reflecting the severity of the achievement difficulty. To illustrate, Fletcher et al. (2011) compared cognitive attributes in typically achieving, adequately responding, and inadequately responding grade 1 children who received intervention. These students were defined as inadequate responders to Tier 2 instruction to adequate classroom instruction based on both decoding and fluency deficits and on only fluency deficits. Figure 3.9 presents the profiles across norm-referenced measures available on this sample. The measures included assessments of phonological awareness, rapid naming, expressive and receptive language, working memory, vocabulary/verbal knowledge, and nonverbal problem solving. In general, multivariate statistical tests were not significant for comparisons of the two inadequate responder groups, but both differed from the responder group. Figure 3.9 shows that elevation differences across groups tended to occur for each variable, but the shapes of the profiles are relatively similar. Measures of phonological awareness and working memory/syntactic comprehension accounted for most of the unique variance across comparisons.

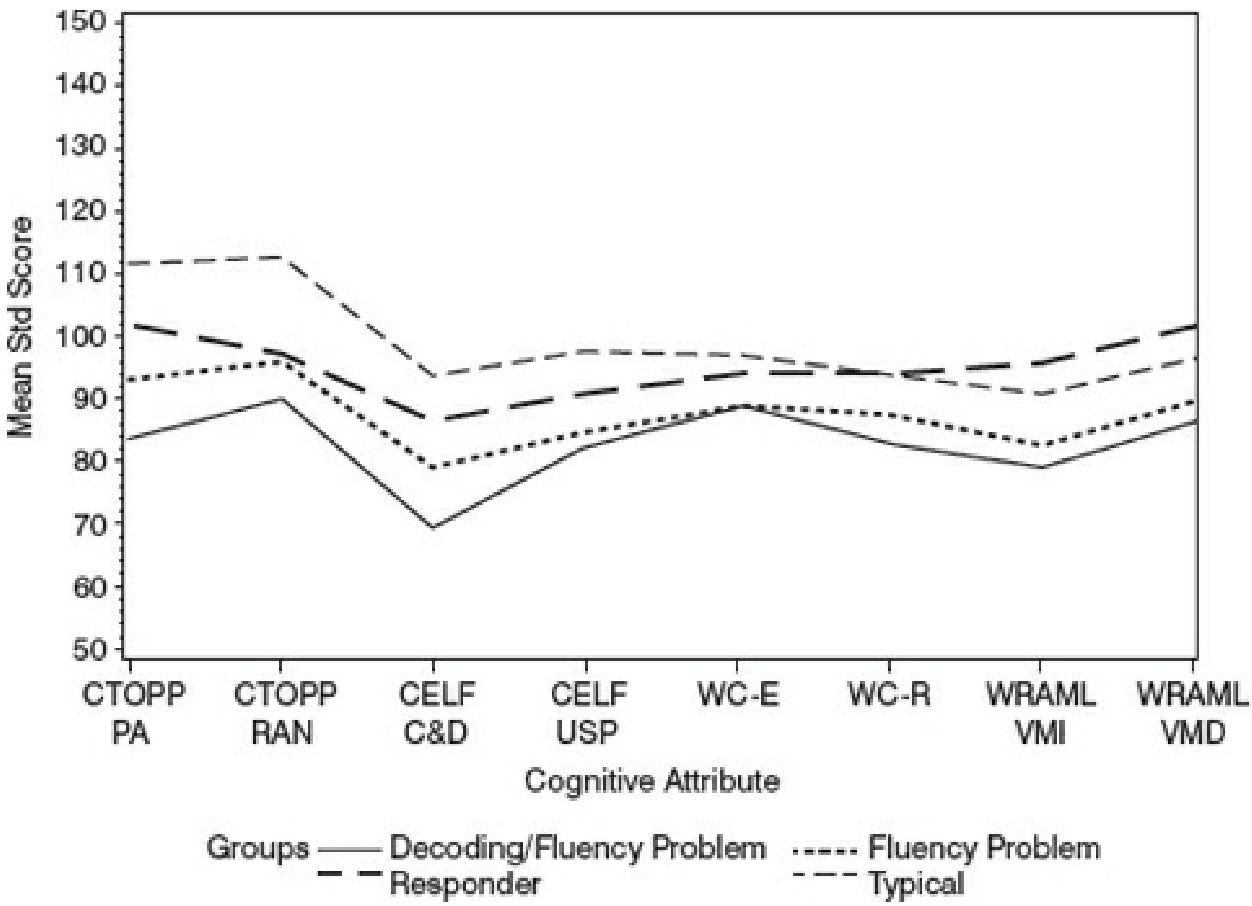

FIGURE 3.9. Cognitive profiles of inadequate responders with decoding and fluency deficits and only fluency deficits, responders, and typically developing children. The significant differences in levels of cognitive skills reflect the severity of the reading problem, with inadequate responders meeting poor decoding and fluency criteria showing the most severe reading problems and, in the figure, the lowest levels of cognitive functions. Note that the profiles differ largely in elevation, not shape. CTOPP PA and CTOPP RAN, Phonological Awareness and Rapid Automatized Naming subtests of the Comprehensive Test of Phonological Processes; CELF C&D and CELF-USP, Concepts and Following Directions and Understanding Spoken Paragraphs subtests of the Clinical Evaluation of Language Fundamentals, Fourth Edition (CELF); WC-E and WC-R, Word Classes—Expressive and Word Classes—Receptive subtests of the CELF; WRAML VMI and WRAML VMD, Verbal Memory Immediate and Verbal Memory Delayed subtests of the Wide Range Assessment of Memory and Learning, Second Edition. From Fletcher, Stuebing, Morris, and Lyon (2013, p. 46). Copyright © 2013 The Guilford Press. Reprinted by permission.

Altogether, these studies provide validity for identifying inadequate responders in the context of an RTI model by providing evidence for the effective isolation of inadequate responders as a subgroup unique to other students who struggle with academic skills at the lower end of a continuum of severity.

A HYBRID APPROACH

One solution to the difficulties posed by multiple conceptual models and unreliable identification methods is to use multiple criteria for LD identification. In addition, psychometric approaches that take into account the issues with measurement by computing confidence intervals and clearly defining costs and benefits in terms of false positive and false negative errors would be helpful to the field. It is not possible to take a single index and use it in an actuarial fashion to identify LDs. The goal of identification often being to find the "right" child reflects a system that has its origins in entitlement programs in which funds are distributed based on the presence of key attributes, without consideration of these attributes' relation to the need for specific services. In fact, the goal of LD identification should be to identify children who would benefit from intervention resources as well as civil rights protections. As such, the tendency of cognitive discrepancy models to generate high false positive errors relative to true negatives is a very conservative approach. False positive errors may be undesirable from an accounting view, but they are acceptable in a system oriented toward assessing instructional response and where a key characteristic of LDs *is* instructional response. As Macmann et al. (1989) lamented in decrying the unreliability of actuarial decisions about LD, many of these problems would be less acute if identification were oriented toward intervention and multiple criteria were used.

We support the classification approach recommended by a consensus group of researchers convened by the U.S. Department of Education Office of Special Education Programs (Bradley, Danielson, & Hallahan, 2002). It combines features of a simple low-achievement method with those of an instructional response method, the key being the idea that intractability to intervention is an inclusionary criterion that addresses the limitations of methods based solely on low achievement and/or application of exclusionary factors. Moreover, as we demonstrate in Chapter 4, this approach to classification and definition lends itself to a comprehensive evaluation that is less time-consuming and gives priority to intervention because of its focus on the assessment of academic skills and instructional response.

This group suggested three primary criteria, the first two of which are inclusionary (Bradley et al., 2002): (1) student demonstrates low achievement;

(2) student demonstrates insufficient response to effective research-based interventions; (3) exclusion factors such as intellectual disabilities, sensory deficits, serious emotional disturbances, a lack of opportunity to learn, and being language-minority children (in whom lack of proficiency in English accounts for measured achievement deficits) should be considered.

Thus, identifying children as LD, whether as part of the process stipulated in IDEA (2004), a clinic outside of school, or in research, requires the presence of low achievement and inadequate response to quality instruction as inclusionary criteria. If an achievement deficit is present and the student demonstrates intractability in response to quality instruction, this may indicate that the low achievement is unexpected. Cognitive discrepancies do not provide this assurance. In addition, low achievement and inadequate intervention response may be due to other disabilities, such as a sensory problem, intellectual disability, or another pervasive disturbance of cognition. These disorders have different identification criteria and require interventions that address a much more pervasive impairment of adaptation that contrasts with the narrow impairment in adaptive skills that characterizes LD. Contextual factors that interfere with achievement, such as limited English proficiency, comorbid behavioral problems, and economic disadvantage should also be considered (see Chapter 4).

CONCLUSIONS: DEFINITION AND CLASSIFICATION

Psychometric Issues

Because the attributes of LDs are likely to vary on a continuum and are therefore dimensional, identification may be improved by moving away from categorical decisions and considering the likelihood or probability of LD, or more importantly, the likelihood that a person would benefit from intervention. Ideally, the effect sizes of interest would be based on theory or on empirical research that links the estimated size of a discrepancy to achievement performance or to instructional response. Presently, all identification methods for LDs have issues with cut points that are usually not addressed. If criteria for LDs are defined as achievement or instructional response below the 25th percentile or a difference of one standard deviation between two cognitive attributes, they are arbitrary because selection of these specific cut points is not grounded in some set of external criteria for LDs. We would recommend that cut points be set relatively high or that confidence intervals be used around a particular cut point because false positive errors are less detrimental than false negative errors. False positives can be evaluated in the context of intervention response, but presently many children receive intervention or remediation with infrequent progress monitoring, with the potential of languishing in an intervention that is not working. Moreover, the decision process may shift toward a multiaxial, consensus method in which the judgments made by a clinician or team should be evaluated for reliability and data from this evaluation used to enhance accuracy, especially in relation to treatment outcomes. There are certainly situations where the child is not achieving because of poor school attendance, lack of motivation, or emotional difficulties; these should be assessed and treated, necessitating comprehensive evaluations of children with LDs, so that keeping them is reasonable even though reliable application on a case-by-case basis is not straightforward. The main focus should be on instructional response and determinations of the different reasons why children are not responding to adequate instruction, one of which is LDs, which may make these determinations more straightforward and treatment-related.

What Is a Slow Learner?

One common reaction to the lack of validity of classifications based on cognitive discrepancies involves children who don't meet criteria for LDs under these methods, invoking the concept of "slow learners" or the Rutter and Yule (1975) notion of "generally backwards readers." Like Elliott and Grigorenko (2014), we find little evidence supporting the referencing of achievement levels to IQ or to other cognitive discrepancies, or even for the commonly expressed idea that we should have different expectations for academic learning in children with IQ scores that are one to two standard deviations below average (i.e., 70–85). The critical issue is whether there is evidence of an intellectual disability, which then involves the level of adaptive behavior. Children with intellectual disabilities have significant problems with conceptual, social, and/or daily living skills. Adaptation difficulties in people with LDs are narrow and restricted to areas involving schooling; social skills are often an issue because of their cognitive difficulties and their problems in school, but still represent relatively narrow areas of adaptation.

Moreover, what is lurking behind concepts like the slow learner is a somewhat pernicious notion that a person's biological endowment for learning can be indexed by an IQ test or evidence of individual differences in learning. This is an assumption that has its origins in the earliest development of IQ tests and one that has been debated since their inception (Kamin, 1974). As stated by Cyril Burt (1937), "Capacity must obviously limit content. It is impossible for a pint jug to hold more than a pint of milk and it is equally impossible for a child's educational attainment to rise higher than his educable capacity" (p. 477). This view of aptitude assessment in which IQ limits a child's learning potential has been termed "milk-and-jug" thinking (Share et al., 1989) because of the unproven assumption that IQ sets an upper limit on educational outcomes. We have reviewed considerable evidence showing weak relations between IQ and other cognitive discrepancies with a variety of external variables.

Our questioning of cognitive discrepancy conceptualizations of LDs does not mean that we equate any form of low achievement with LDs. As we discuss in the next chapter, there are contextual factors and other conditions that lead to low achievement. These factors and conditions need to be identified and evaluated as part of a comprehensive evaluation and, most importantly, treated. In contrast to cognitive discrepancy models, we have reported evidence for the validity of conceptual frameworks for LDs based on inadequate instructional responses. *We stipulate that what differentiates LDs from other forms of low achievement at the inclusionary level is evidence of intractability to instruction.* This is an alternative way of thinking about LDs in that low achievement and difficulties responding to a series of increasingly intensive interventions are essential for identification. This squares the concept of LDs with typical achievement, prioritizes early intervention and treatment, and helps teachers think about people with LDs as *harder to teach, not unable to learn*. This emphasis on intractability is a fundamental change in the perception of people with LDs that focuses on how to best teach them and not on how to best diagnose them.

The identification of LDs needs to be much more fluid, especially if LDs are seen as an instructional problem. Such a change will be facilitated if classifications of LDs are based on instructional models and unexpected underachievement is viewed in part as intractability to instruction. These approaches also allow us to move away from the unreliability of individual identifications toward a recursive identification method where errors can be corrected over time and identification is dynamic, not static. As Macmann et al. (1989) stated almost 30 years ago:

> Even though the psychometric difficulties may never be completely resolved, classification systems should at least be based on a coherent psychology of helping. . . . Although there is no shortage of children who experience problems in adjustment and the acquisition of essential skills, assessments of the characteristics of these children are important to the extent that contributions are made to the design and evaluation of meaningful interventions. Assessments and classifications can be guided by principles of intervention design . . . with expected errors of judgment and measurement partly moderated through a recursive system of reflective and empirical practices. . . . The concept of sequential decisions is fundamental, permitting fallible data and resulting decisions to be evaluated over time, and modified as necessary, in an iterative fashion. (pp. 145–146)

Methods based on RTI incorporate this idea of a recursive, sequential approach to identification through an MTSS service delivery framework and potentially make identification self-correcting, viable, and tied to intervention. As we see in Chapter 4, this approach and the hybrid method we advocate suggests changes in how assessment occurs, prioritizing screening and progress monitoring as essential characteristics.

CHAPTER 4

Assessment of Learning Disabilities

The review of classification models and identification methods in Chapter 3 leads directly to a clinical assessment approach for individuals where LDs are an issue (Fletcher, Francis, Morris, & Lyon, 2005b; L. S. Fuchs & D. Fuchs, 1998). The tests and procedures selected for any assessment are derived from a classification model, which specifies the constructs that need to be measured. If the classification is based on a cognitive discrepancy, such as an aptitude–achievement discrepancy model, the primary tools would be the tests used to measure aptitude (e.g., IQ or listening comprehension) and achievement (e.g., reading, math, and written language). Alternatively, the identification method could be based on patterns of strengths and weaknesses in cognitive skills and would utilize cognitive-processing measures, or neuropsychological tests, as well as achievement measures. If the classification reflects a low-achievement model, aptitude would not be measured in favor of a focus on achievement. If a model emanates from RTI, assessments of the quality of instruction and instructional response would be required.

In the hybrid model we proposed in Chapter 3, an evaluation of LD requires an assessment of intervention response, norm-referenced assessments of achievement, and an evaluation of contextual factors and associated conditions that contribute to the achievement problems. Most important, this component of the evaluation may suggest alternative intervention needs that differ from those that directly address achievement issues through instructional methods. For example, a comorbid condition such as ADHD may be identified, which would require additional intervention methods beyond those directed solely at improving reading, math, or writing skills.

TEST AND TREAT VERSUS TREAT AND TEST

An approach to the assessment of LDs based on the hybrid model is different from the traditional test-to-diagnosis approaches that have dominated the assessment domain for many years (see Figure 3.8). In the approach to identification that we propose, LDs are not "diagnosed" on the basis of a battery of psychometric tests administered on a single occasion. Rather, LDs are identified only after a specific attempt is made to systematically instruct the person. We question whether LDs can be reliably and validly identified in the absence of intervention. We suggest that ensuring adequate opportunity to learn is a prerequisite to the identification of LDs, regardless of the setting, so that traditional test-to-diagnosis approaches can, at best, identify the person as being "at risk" for LDs. However, a single assessment will not lead to reliable identification if the approach is based on strictly applied cutoff scores or formulae (see Chapter 3).

Figure 4.1 provides a diagram of the three components of an evaluation under the hybrid method proposed in Chapter 3. The first two components are inclusionary: evidence of low achievement and inadequate instructional response. The third component includes an evaluation as needed for exclusionary conditions, such as a sensory or motor disorder or an intellectual disability; other often comorbid disorders such as ADHD and other behavioral and emotional problems; and contextual factors such as the home environment, economic disadvantage, and lack of majority language proficiency. All of these aspects of the third component are considered exclusionary of LDs; however, note that they are inclusionary of other conditions associated with low achievement and, in terms of contextual factors, must be the primary factor associated with low achievement. The primary purpose of assessment in an instructional model of LDs is not to decide who meets criteria for LDs, but to sort through different factors associated with low achievement and develop an intervention plan that may or may not involve LDs regardless of the factors believed to be primarily associated with low achievement.

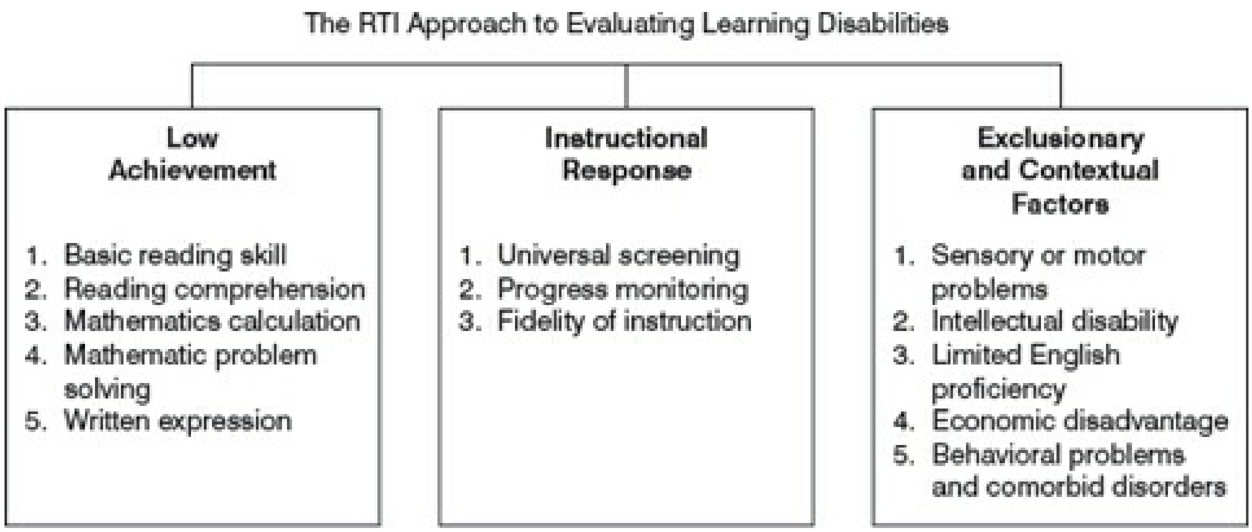

FIGURE 4.1. Rubric for identifying LDs in a hybrid model. The three components are evaluations for low achievement, inadequate instructional response, and other conditions and contextual factors associated with low achievement. Courtesy Whitney Roper.

The goal of any evaluation should be to intervene as soon as possible with a person who is struggling to achieve. In schools, screening for reading, math, and behavior problems can be done on a large scale, as advocated in the National Research Council consensus report on minority overrepresentation in special education (Donovan & Cross, 2002) and implemented in states like Texas (Foorman, Fletcher, & Francis, 2004). Screens should be very brief (< 5 minutes) and oriented to ensuring that children who are not "at risk" are accurately identified. In other words, missing an at-risk child is a more serious decision error than failing to identify a child who is at risk (false positive). Progress will be monitored and false positives can be identified. Screening is a triage system to identify those who need monitoring.

Since IDEA 2004 was passed with its support for RTI-based approaches to identification (see Chapter 2), universal screening and progress monitoring have become more widely implemented as key components of this approach. In the context of an MTSS service delivery framework (Chapter 5), those who are identified as at risk should have their progress monitored with curriculum-based measures (CBMs) and receive increasingly intensive interventions that may eventuate in identification for special education if the student responds inadequately to quality intervention and meets additional criteria (Kovaleski, VanDerHeyden, & Shapiro, 2013; Vaughn & Fuchs, 2003).

This approach, central to an MTSS framework, relies on a "treat-and-test" model to prioritize establishment of a disability over establishment of a disorder. In most eligibility systems, disability status depends on both the presence of a disorder *and evidence that the disorder interferes with adaptive functions,* so there is evidence that the disorder is *disabling*. This is apparent in the new DSM-5 as well as in IDEA 2004 (see Chapter 2), where evidence of problems with adaptation (DSM-5) or educational need (IDEA 2004) is required. Evaluating instructional response is one way to operationalize the disability component of an evaluation for LDs.

As we suggested in Chapters 2 and 3, inadequate instructional response is the marker of unexpected underachievement in a classification and identification framework emanating from RTI. This criterion is inclusionary and must be present for identification. In cognitive discrepancy frameworks, inadequate instruction is exclusionary. In an RTI method, adaptive impairment is determined first (i.e., evidence that the child does not achieve at some benchmark despite quality instruction). In the absence of an assessment of intervention response, the assessment of educational need has been somewhat subjective. This subjectivity is partly responsible for the confusion that emerges when an interdisciplinary team denies eligibility despite an identified disorder that sometimes, but not always, interferes with school performance.

This scenario often occurs when a school considers an evaluation that was conducted independently (i.e., outside of schools, but in mental health or psychoeducational clinics or other private settings). In our view, the basis for identification should still reside in the hybrid model described in Chapter 3 (Bradley et al., 2002). In clinic situations, it may be necessary to initially establish evidence of low achievement. Evidence of low achievement should first lead to concerns about the quality of the instruction the child has received. Assessments of IQ or cognitive-processing skills to "diagnose" LDs should occur only if the child presents with concerns about disorders that require such assessment (e.g., the use of IQ tests to establish an intellectual disability). Professionals who conduct assessments related to LDs should have a working knowledge of educational interventions and a relationship with professionals in or out of school who can provide intervention and measure intervention response in individuals with achievement difficulties. These interventions are tied to strengths and weaknesses in academic domains and a good assessment of these domains can help differentiate instruction (Spear-Swerling, 2015). If necessary, the intervention professional can independently evaluate progress in conjunction with more frequent assessments of learning, which are produced over the course of intervention.

At the same time, we recognize situations involving eligibility for accommodations on college assignments, college entrance examinations, and other high-stakes assessments (Lovett & Lewandowski, 2015; Mapou, 2009). In these situations, we suggest looking carefully at the individual's history to understand why a potential disability was not identified, and, if possible, looking at the history of instruction and intervention. In addition, we would use a de facto low-achievement method with assessment of contextual factors and other potential disabilities. We do not see sufficient empirical support for the value of accommodations on college entrance examinations, such as the Scholastic Aptitude Test, when based on complete cognitive and achievement assessment batteries, or evidence of an achievement deficit with a correlated processing deficit (but see Lovett & Lewandowski, 2015, for a more nuanced view of this issue, where research is really in its infancy). Not only does this approach fail to take into account the effect of intervention on the "profile," but also, as shown in Chapter 3, such assessments do not have strong reliability or validity. It is unfortunate that LDs are not always identified, but examiners should take no reassurance in comprehensive cognitive or neuropsychological assessments. In the end, it will be the history and level of achievement that drive identification, and more importantly, the need for accommodations and specialized interventions.

ASSESSMENT PRINCIPLES

Heterogeneity of LDs

Prior to discussing evaluations of intervention response, achievement, and contextual factors, the question of the relevant domains of LDs should be addressed. This question reflects long-term concerns regarding the heterogeneity of LDs—the fact that the construct of LDs can be rooted in impairment in any one of several different domains of achievement. LDs are largely domain-specific. This means that disabilities involving reading, math, and written expression are different in phenotypic characteristics and intervention needs. Although many people with LDs have impairment in more than one domain, prototypes for subgroups of people with disabilities are specific to the domains of reading and math. This heterogeneity alone complicates the proposition that LDs can be subsumed under a single overarching conceptualization. At the same time, there is evidence for domain-general factors, especially in comorbid associations with working memory, processing speed, and oral language comprehension representing the most likely cognitive processes applicable across domains. More research is, however, needed to understand potential domain-general factors. There is little evidence that assessment of these domain-specific factors facilitates identification, but it is possible they relate to intervention (L. S. Fuchs et al., 2014b; Willcutt et al., 2013).

What to Assess

Five academic prototypes epitomize the domains of a hypothetical classification of LDs we proposed in Chapter 1 involving the domains of word-level reading and reading comprehension, mathematics calculation and problem solving, and written expression (Table 4.1). However, these domains are not consistently reflected in IDEA 2004, which included eight domains of LDs. Two involved oral expression and listening comprehension, which were also addressed in the speech and language category of IDEA 2004. The reason for this duplication is that these conditions are described in the 1975 U.S. statutory definition of LDs, which includes disorders of listening and speaking (Chapter 2). Even if listening comprehension is not regarded as a component of receptive language, it closely parallels reading comprehension in children who do not show word-reading disabilities (Chapter 7). Although in our first edition we advocated for inclusion of reading fluency as a separate subgroup, as in IDEA 2004, we now believe that reading fluency problems represent more general problems with automaticity of basic academic skills. In the reading fluency domain, problems with automaticity are extensions of difficulty with the accuracy and automaticity of word-level skills. Similarly, automaticity of basic skills are factors for math fluency, and likely for written expression. In math, fluency difficulties represent a failure to automatize knowledge of basic facts and their application in online problem solving, but this is less well understood than in reading, where the problem is automatic recognition of sight words (Chapter 10). In written expression, automaticity of transcription skills is important for fluency in writing, but more research is needed. We have provided a discussion of automaticity issues in Chapter 10.

TABLE 4.1. Subgroups Forming a Hypothetical Classification of LDs

Disability type	Component academic deficits
Reading disability	Word recognition and spelling
Reading disability	Comprehension
Mathematics disability	Computations
Mathematics disability	Problem solving
Written expression disability	Handwriting, spelling, and/or compositic

Norm-Referenced Achievement

The evidence that supports the hypothesis of this classification of LDs is summarized in Chapters 6–9. For assessment purposes, these domains of potential low achievement must be evaluated as part of the hybrid model, along with assessments of automaticity. These assessments can be completed with norm-referenced achievement tests. This assessment ensures that low achievement is directly measured with tests of high reliability and validity and also provides an assessment that, with assessments of instructional response, provides multiple indicators for determining LDs. Finally, patterns across academic domains can help differentiate instruction (Spear-Swerling, 2015).

Instructional Response

Assessments of instructional response are best accomplished with progress-monitoring methods, such as CBMs. These methods are best developed for word recognition, reading fluency, math computations, writing, and spelling. They typically involve timed assessments, but a variety of procedures can be adapted (see Kovaleski et al., 2013). It is possible to assess reading comprehension with CBM measures using cloze or maze tests, but the format provides a limited assessment of comprehension, which in itself is difficult to assess because it reflects multiple underlying processes. Reading fluency CBMs, however, are strong predictors of reading comprehension. The difficulty in assessing complex skills such as reading comprehension, math problem solving, and written expression is one of the main reasons why we suggest that norm-referenced assessments of the achievement domain are important for identifying LDs. The other reason is to increase reliability of identification with multiple measures and criteria.

Progress monitoring can be done using a variety of methods, although CBMs are the most validated method. In terms of norm-referenced tests, it is possible to employ widely used norm-referenced tests if the interval between start and completion of intervention is sufficiently long. Usually, more frequent assessments with CBMs are preferred. With such tests, alternate assessment forms can be used repeatedly to model student improvement as a function of intervention, but only at sufficiently long intervals of time (usually several months). There are alternate forms for frequent CBM data collection, many of which have been reviewed by the National Center for Intensive Intervention (*www.intensiveintervention.org*). As the reviews by this center indicate, CBM procedures vary considerably in terms of reliability, number of forms, grades, and academic domains addressed. However, research substantiates that some forms of CBM provide sound information about how well students are progressing. These measures are reliable (above .90) at elementary and middle school levels (Barth et al., 2012; Espin, Wallace, Lembke, Campbell, & Long, 2010). Although passage fluency measures do not directly index comprehension, correlations with standardized measures of reading comprehension range from .50 to .90 for early grade readers (Barth et al., 2012; Shinn, 1989). Because estimates tend to vary for individual passages, it can be important to average performance across two to three passages, especially when summative decisions like postintervention performance are made.

Contextual Factors and Other Conditions

The final component involves assessment to identify contextual factors and/or other disorders sometimes considered "exclusionary" of LD. This component is operationalized using rating scales to screen for comorbid ADHD and other behavioral and emotional difficulties, interviews, history, and formal assessments related to other disorders that may be exclusionary of LDs. Standardized tests may be needed for determining intellectual disabilities and majority language proficiency, but these should be completed only when there is a question warranting such assessments. In an instructional model, the child has been in intervention and there should be hypotheses about the basis for inadequate levels of achievement that, along with inadequate instructional response, lead to the evaluation.

What Not to Assess

Assessments that do not lead to reasonable interventions do not need to be completed. *Funds not spent on assessments can be used to support intervention.* Children should be routinely screened for peripheral vision and hearing problems and should be in good physical health. There are no routine indicators for blood tests or brain scans related to LDs. Referrals to a behavioral optometrist are not indicated (Chapter 6) or to an audiologist for "auditory processing problems." There is little evidence that supports the validity of associations of these questionable diagnoses with LDs, much less the validity of the interventions that stem from them. As we discussed in Chapter 3, we do not believe the evidence supports a need to routinely assess IQ or even processing skills with established relations with academic skills (Fletcher & Miciak, 2017). As Elliott and Resing (2015) argued, IQ scores do not have clear apparent treatment implications for children with LDs. The guidance accompanying the IDEA 2004 regulations also does not encourage routine IQ or cognitive assessments for SLDs:

> The Department does not believe that an assessment of psychological or cognitive processing should be required in determining whether a child has an SLD. There is no current evidence that such assessments are necessary or sufficient for identifying SLD. Further, in many cases, these assessments have not been used to make appropriate intervention decisions. (IDEA regulations, 2006, p. 46651)

As Schneider and Kaufman (2017) stated in a defense of cognitive assessments for LDs,

> The existing evidence base that demonstrates the value of comprehensive cognitive assessments for this purpose is not nearly as strong as it needs to be. Proponents of comprehensive cognitive assessments for learning disability identification must do more to rigorously evaluate their beliefs or else concede the argument to those with better evidence. (p. 8)

There Is No Consensus, So Follow the Evidence

Not all practitioners and scientists agree with this approach to assessment. As indicated in Chapter 1, some argue that LDs extend beyond achievement domains, the most obvious example being social skills. Many individuals with LDs do have problems with social interactions. In some instances, this represents a comorbid disorder, as in the example of ADHD. In other instances, social difficulties seem to represent correlates of the same underlying processes that lead to achievement difficulties, epitomized by the hypothesis of a nonverbal LD (Rourke, 1989). Some children with LDs clearly have problems with social skills, motor skills, perceptual abilities, oral language, and other areas that do not directly involve achievement (Cornoldi et al., 2016). Consider, however, that many people with problems in these areas do not have achievement problems (Torgesen, 2002). In arguing that achievement deficits are necessary but not sufficient, we suggest that LD classifications are not viable without some type of marker that reliably indicates the presence of LDs (Stanovich, 1991).

In the next sections, we discuss the three essential components needed to evaluate and identify people with LDs. This includes the evaluation of intervention response (including the evaluation of intervention integrity), which we place first because of the treat-and-test approach we advocate. For people with inadequate instructional response, the evaluation of achievement, and the evaluation of contextual factors and associated conditions are needed. We conclude by briefly discussing issues related to cultural and linguistic sensitivity and gifted (“twice-exceptional”) people who may have LDs.

SPECIFIC ASSESSMENT PROCEDURES

Instructional Response and Progress Monitoring

For assessment of instructional effectiveness, which relies on ongoing progress monitoring to determine when modifications to intervention are needed to ensure adequate long-term progress, the assessment must be longitudinal and focus on rate of change (slope). We address this kind of ongoing use of progress-monitoring data, such as CBM, for the purpose of instructional decision making in Chapter 5.

We recommend a more simple and time-efficient approach for using CBMs in the RTI/LD identification process. L. S. Fuchs and Fuchs (1998) proposed the use of both slope and final status for identifying inadequate response in an RTI/LD identification process, suggesting that a student must demonstrate a “dual discrepancy” in which the slope and final level are both at least one standard deviation below those of peers or some type of norm-referenced standard. However, recent studies question whether slope adds information to final status *for identification purposes,* when the task is to simply identify whether the person has responded adequately or inadequately to intervention (Brown-Waeschle, Schatschneider, Maner, Ahmed, & Wagner, 2011; Fletcher et al., 2014; Tolar, Barth, Fletcher, Francis, & Vaughn, 2014). In progress monitoring for instructional decision making, slope is used to predict final performance. In the case of RTI/LD identification, the final (postintervention) performance is already known.

For classifying response as adequate or inadequate, or as part of a comprehensive evaluation for special education eligibility, the critical issue is the person’s status at the time when intervention has recently been completed. This is the same time frame in which a comprehensive evaluation should be conducted. Thus, in contrast to common practice (e.g., Kovaleski et al., 2013), we do not recommend projections of status based on regression procedures beyond the point of at which decision making will occur. The reliability of this approach has not been established, and most information will be in the final status measures administered in the comprehensive intervention. An alternative would be a reliable change index based on pretest–posttest change, although this approach has not been adequately studied. Our goal in any form of assessment is parsimony and simplicity because resources need to be directed at intervention as much as possible.

Evaluating Fidelity and Integrity of Interventions

CBM and other assessments of academic status, collected at the completion of intervention, should also be accompanied by observations of the integrity of the implementation of the intervention. This includes the nature of and the amount of time spent on supplemental instruction, especially if the child does not appear to be making progress. School psychologists are often well prepared in this area of assessment. Although a psychologist operating outside of schools may not be in a position to implement CBM or to personally evaluate the integrity with which the intervention is implemented, such assessments should be expected, especially if the referral is to a private academic therapist. A recent summary of factors related to the assessment of intervention integrity by the National Center for Learning Disabilities was provided as part of a toolkit for the identification of LDs in the context of RTI (Cortiella, Gamm, Rinaldi, & Goodman, 2014). Recommendations included documentation of the following. First, evidence-based interventions and general education classroom curriculum were used to instruct the student. Second, the intervention was appropriate for the student's instructional level. Third, the intervention has been proven efficacious with other students similar in age and level of performance. Fourth, educational professionals with appropriate training and demonstrable proficiency delivered the intervention. Fifth, implementation fidelity to the program as designed by the developers was present and demonstrated. Sixth, the intervention was delivered with sufficient time to show an effect and with sufficient intensity.

The importance of formal assessments of assessment integrity was highlighted in a study by VanDerHeyden, Witt, and Gilbertson (2007), which formally evaluated the fidelity of implementation across multiple tiers of RTI implementation. They provided checklists and observation schedules for different components of an RTI system, including screening, progress monitoring, and intervention integrity and were generally able to demonstrate relatively high, but variable, levels of fidelity in a closely monitored RTI implementation. In a review of fidelity of treatment evaluations, Keller-Margules (2012) identified the following critical components of the evaluation of implementation integrity for RTI through assessments of: (1) assessment integrity, (2) instructional and intervention integrity, and (3) procedural integrity, with the latter representing adherence to a school or district RTI plan. She recommended the use of multimethod and multi-informant methods of data collection, including direct observation using checklists and teacher/administrator reports. Forms for collection of this type of data using a model with similar components can be found in Kovaleski et al. (2013).

As this summary shows, evaluations of the integrity of instruction in an RTI approach involve the entire implementation, including the screening process, the delivery of the intervention, and the evaluation of student responsiveness. These assessment components are integral to the RTI assessment process in identifying LDs. With respect to determining instructional response, the best method is to evaluate academic status at the end of intervention, using multiple sources of data including CBM endpoints. If an individual has not achieved a final status indicating levels of achievement of sufficient foundational academic skills to succeed in the general education program, then the conclusion is that the student has been inadequately responsive to quality instruction. A comprehensive evaluation could then be required, with the likelihood that more sustained or individualized (Tier 3-like) intervention is required. The validity of this conclusion depends on whether the student was initially screened into intervention appropriately, whether the correct intervention was provided with fidelity and sufficient intensity, and whether the decision about responsiveness was based on sound data. Below, we elaborate on the prior discussion concerning methods for identifying inadequate responders.

Identifying Inadequate Responders

When CBM data are systematically collected in relation to a quality intervention, a variety of approaches have been used to establish whether the person's response is adequate. Although it is apparent that intervention response exists on a continuum and that firm cut-off scores are not likely to encompass every student of concern (Vellutino et al., 2006; Fletcher et al., 2014), specific thresholds can provide guidelines for identifying students in need. These determinations should be determined with high thresholds and confidence intervals around the desired threshold to account for the measurement error of the test. As with screening, the goal should be to avoid missing people in need because a false positive identification error is less costly to the person and it will be quickly apparent that intervention is not needed. To illustrate, Fletcher et al. (2014) used the 25th percentile on norm-referenced end-of-intervention assessments and a criterion-referenced benchmark for an oral reading fluency measure. The cut points may seem high, but missing children who need intervention (false negatives) is more serious than identifying children who do not require intervention (false positives). In a sequential MTSS service delivery framework, progress monitoring over time would reduce false positive errors at low cost.

When comparisons are made of identification decisions based on CBMs, most of the information is in the final CBM assessment. In one study, the use of slopes by themselves or both slopes and intercepts resulted in an increased likelihood of false positive identifications (Fletcher et al., 2014). In fact, identification decisions in this study, based solely on a CBM benchmark, also resulted in apparent false positive decisions. This problem was reduced if low performance was required on two different short assessments, again highlighting the need for more than one indicator for identification, reflecting issues with firm cut points identified in Chapter 3.

It is useful to examine multiple criteria. End-of-year or end-of-intervention assessments can identify students as inadequate responders when those students perform below a benchmark. This could be an age-adjusted standard score below the 25th or 30th percentile. Another common benchmark sets a criterion based on passage reading fluency. First graders, for example, should read 35–40 words per minute at the end of the year depending on the difficulty level of the text.

Final status by itself permits some students to be classified as inadequate responders, despite significant progress, because the initial level of performance was very low. Despite strong growth, some students are likely below the benchmark at the end of the intervention. For these students, focusing only on the benchmark will suggest that the intervention was ineffective, while examining the slope may suggest the need to continue intervention (with or without identification with LDs, depending on the level of performance). Similarly, examining only the slope permits some students to be identified as inadequate responders even though they meet the norm-referenced or benchmark criteria at intervention completion, suggesting positive future outcomes. If a person is below a benchmark, subsequent intervention is essential. If the slope is accelerated but the student remains below benchmark, intervention should be continued, with intensification if the student's achievement gap with respect to classmates remains severe.

Figure 4.2 shows actual progress-monitoring graphs for an adolescent reading intervention in sixth-grade children (Vaughn et al., 2010a). In this figure, slopes vary over the five assessment points used for this Tier 2 standard protocol intervention. There are fewer time points than in many interventions because the oral reading fluency alternate forms are equated and because growth in adolescents tends to be slower. It is clear that student C has not responded to the intervention. Student B is responding to the intervention, but needs more time to reach the criterion on this measure (45 words per minute). Student A has responded positively. While the slopes are different, if the goal were to determine eligibility for special education, the most relevant information is in the end-of-intervention time point.

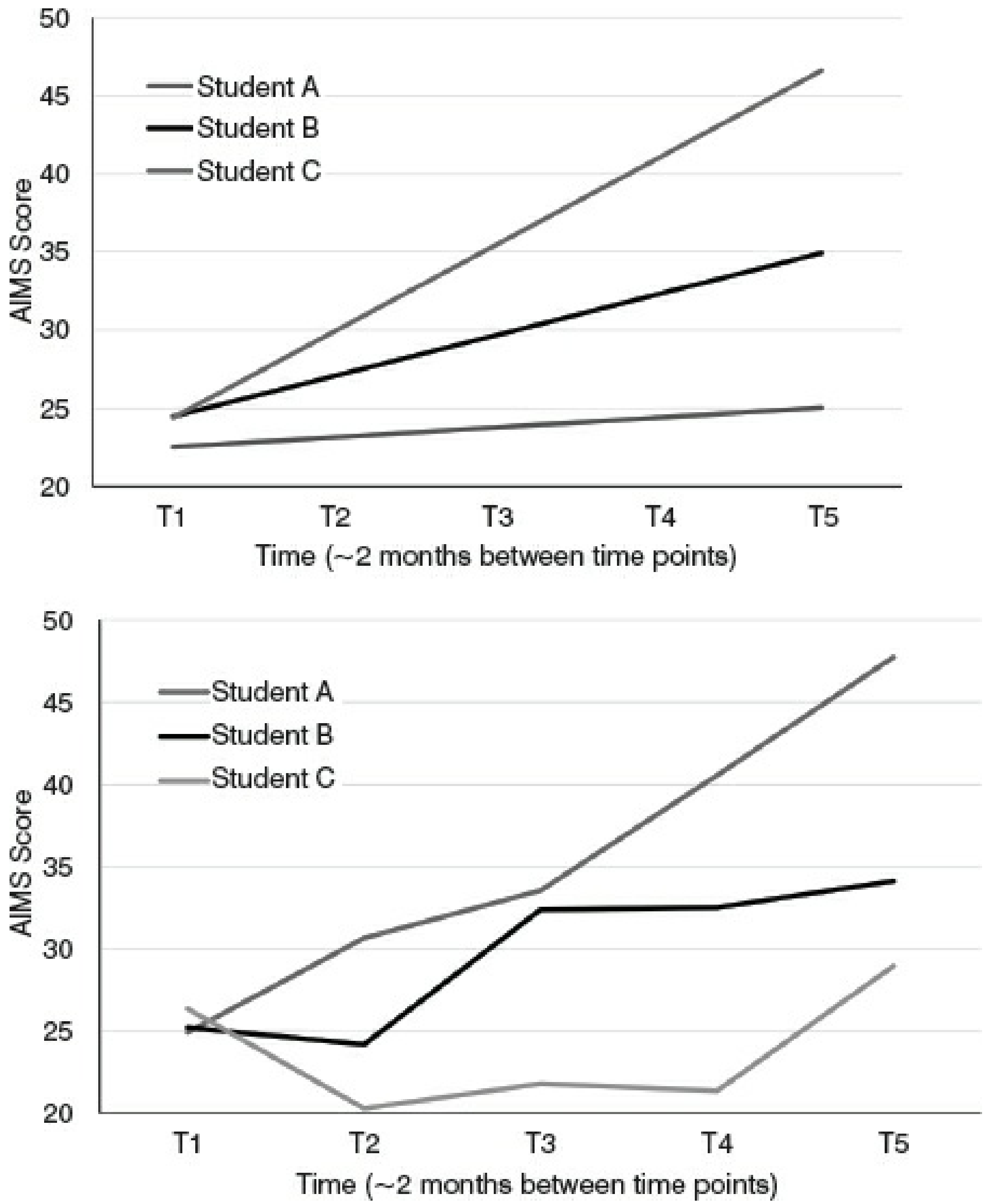

FIGURE 4.2. Individual growth curves for three adolescent students who show accelerated gains (Student A), average growth (Student B), and low average (inadequate) growth (Student C). Panel A uses equated forms and estimated growth. Panel B shows actual raw score growth, illustrating the importance of form equation and estimated growth for understanding the fluctuations in raw scores. Data from Tolar, Barth, Fletcher, Francis, and Vaughn (2014). Courtesy Tammy Tolar.

In another example from a recent evaluation of a Tier 2 math fractions intervention, L. S. Fuchs, Sterba, D. Fuchs, and Malone (2016c) found that intervention was comparably effective for students with varying levels of preintervention academic deficits. Figure 4.3 shows posttest scores on one of the fraction outcome measures. The lines show intervention versus control group differences for students whose math achievement on a nationally normed test at the start of intervention were at the 1st, 3rd, 9th, 13th, 27th, and 34th percentiles. The parallelism of the lines indicates comparable effects for students at these percentiles. In contrast, Figure 4.4 shows postintervention achievement gaps (with respect to not-at-risk classmates), expressed as effect sizes, for students who received intervention (white bars) versus those who did not (black bars) groups. Postintervention achievement gaps were substantially larger for students whose initial math achievement was more deficient. Although the intervention benefited students across the at-risk initial achievement distribution, students who began with more severe math deficiencies require more sustained intervention.

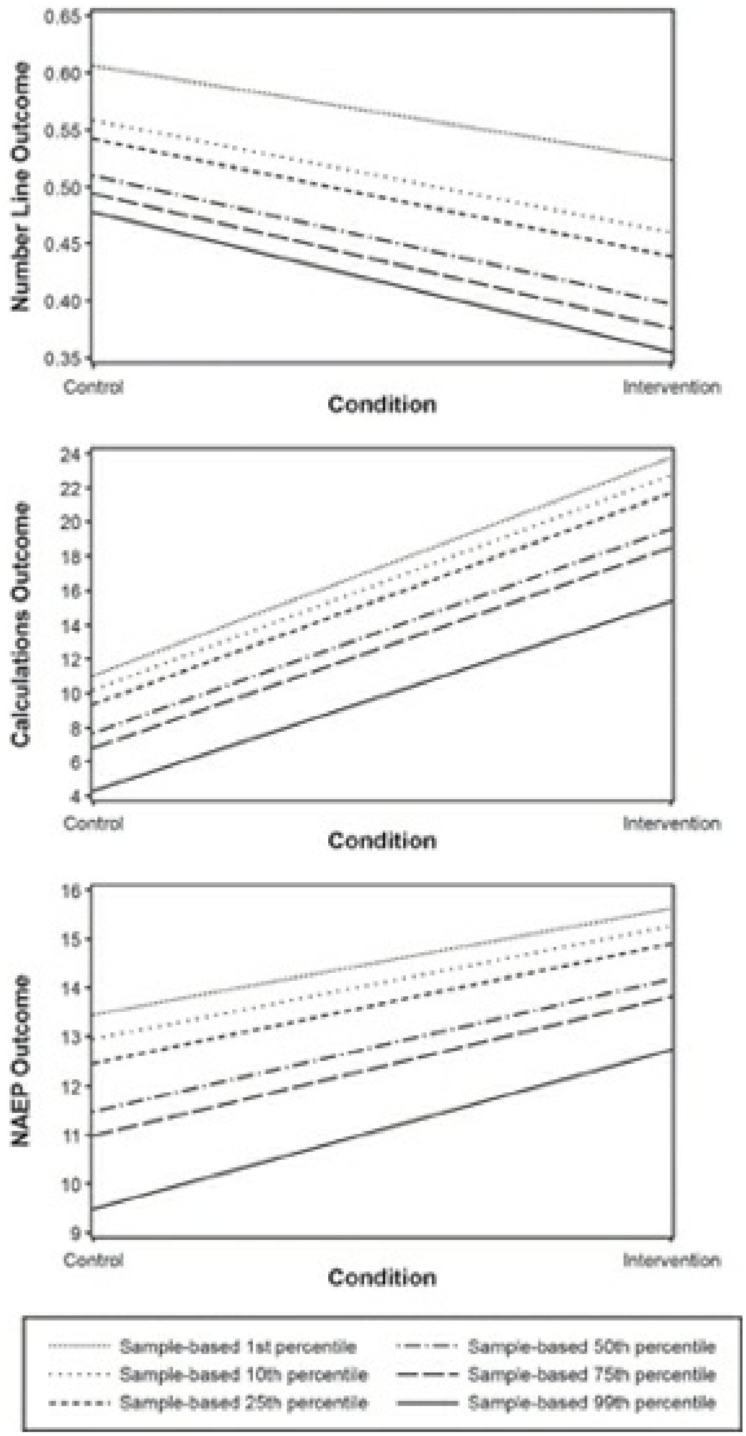

FIGURE 4.3. Posttest scores showing intervention versus control group differences for students whose math achievement at the start of intervention was at the 1st, 3rd, 9th, 13th, 27th, and 34th percentiles. The parallelism of the lines indicates comparable effects for students at these percentiles. From L. S. Fuchs, Sterba, D. Fuchs, and Malone (2016c). Copyright © 2016 Taylor & Francis. Reprinted by permission.

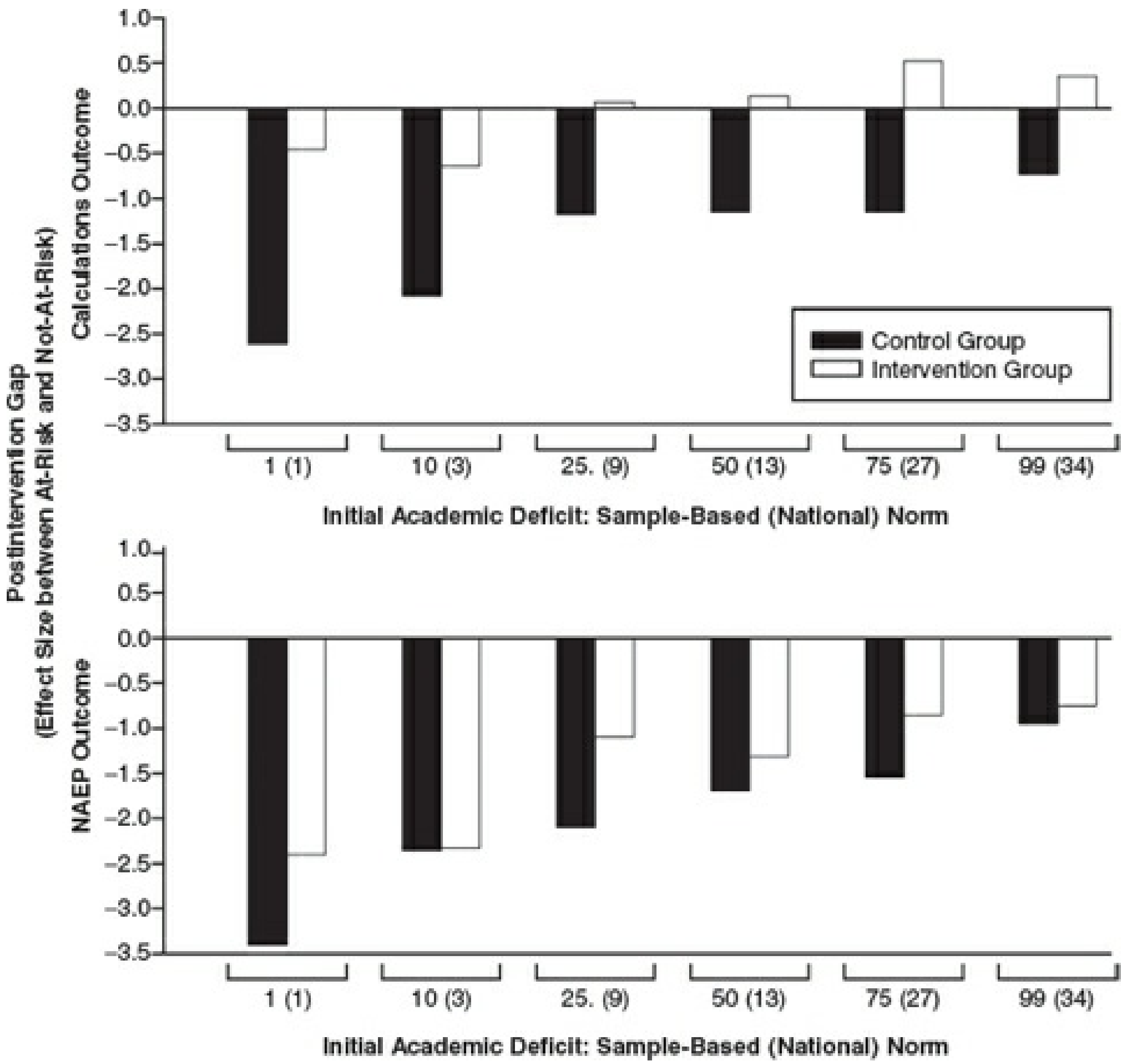

FIGURE 4.4. Postintervention achievement gaps as a function of study condition and initial academic deficit severity. From L. S. Fuchs, Sterba, D. Fuchs, and Malone (2016c). Copyright © 2016 Taylor & Francis. Reprinted by permission.

Some of these students need more intensive interventions to increase the rate of response. It is important to note that no intervention, even one that is generally effective, produces growth for all students. For individual students who do not exhibit adequate growth (improvement or slope) during an intervention (and complete intervention with sizeable achievement gaps), the necessary decision is to intensify or individualize intervention. The need for sustained or intensified/individualized intervention necessitates a comprehensive evaluation, with possibility of LD identification.

Evaluating Achievement Domains

Many norm-referenced achievement tests can be incorporated into the assessment of LDs. At a minimum, the five major achievement domains must be assessed, representing people who are primarily impaired in (1) word recognition, (2) reading comprehension, (3) mathematics computations, (4) mathematics problem solving, and (5) written expression, including spelling, handwriting, and/or composition. These patterns were established through research by Rourke and Finlayson (1978), Siegel and Ryan (1989), and Stothard and Hulme (1996); Chapters 6–9 provide extensive discussion of the evidence for these subgroups (see also Spear-Swerling, 2015). Many individuals have difficulties in multiple domains, making a complete evaluation of academic achievement necessary for anyone considered for LDs. If the specificity of LDs was indicated by impairment just in one domain, LDs would be exceedingly rare.

Fletcher et al. (2005b) proposed the use of tests from the same achievement battery because the same cohorts would be used to develop the norms. This constancy facilitates comparisons across tests. In a simulation, Fletcher et al. (2014) found that using assessments with different normative bases significantly reduced agreement across final status measures. However, more important than the battery from which these tests are selected are the constructs that are measured and the quality of the indicators of these constructs: word recognition, reading comprehension, math computations and problem solving, and written expression. In addition, assessments of the automaticity of reading and math skills should be provided.

Examining the validity studies of different norm-referenced achievement tests supports these distinctions at a construct or latent level. In a study that involved a larger sample of children with poor reading skills, and which included both norm-referenced assessments and CBM assessments, Cirino et al. (2013) found that four latent variables were indicated by different assessments of reading skills: untimed decoding accuracy (i.e., single-word decoding), timed reading accuracy (i.e., timed assessment of word list or passage reading), and a comprehension factor composed of measures of reading comprehension, listening comprehension, and vocabulary. The fourth factor was interesting because it involved CBMs that were based on timed maze or cloze procedures or timed norm-referenced assessments based on sentence reading and a verification of whether the sentence "made sense." These timed fluency measures had a comprehension component and had to be accounted for separately in the modeling process. While there were differences in the value of individual indicators for the good and poor readers, the four constructs could be identified in both samples. Both assessments of construct validity from normative samples and from specific studies with large numbers of children with achievement difficulties support the validity of these construct differentiations.

Table 4.2 maps the constructs and their assessment with the Woodcock–Johnson Achievement Battery–IV (WJ; Schrank, Mather, & McGrew, 2014), the Wechsler Individual Achievement Test–III (WIAT III; Pearson, 2009), and the Kaufman Test of Educational Achievement, Third Edition (KTEA; Kaufman & Kaufman, 2014), all frequently employed to assess achievement levels in LDs. There are also other norm-referenced assessments that can be used instead of, or to supplement, these batteries, some of which we mention below. For example, spelling can be used to screen for written expression and handwriting difficulties.

TABLE 4.2. WJ, WIAT-III, and KTEA Subtests in Relation to Component Academic Deficits

Construct	WJ	WIAT-III	KTEA
Word recognition	Word Identification	Word Reading	Letter and Word Recognition
	Word Attack	Pseudoword Decoding	Nonsense Word Decoding
Reading fluency	Word Reading Sentence Reading	Oral Reading	Silent Reading
Reading comprehension	Passage Comp	Reading Comp	Reading Comp
Math computations	Calculation	Numerical Operations	Computation
Math problem solving	Applied Problems	Problem-Solving	Concepts and Applications
Written expression	Spelling	Spelling	Spelling

Supplemental constructs			
Math fluency	Math Facts	Math Fluency	Writing Fluency
Writing fluency	Sentence Writing	Alphabet Writing	Writing Fluency
Written Expression	Writing Samples	Essay Composition	Written Expression

The evaluation of achievement levels can be conducted hierarchically, and not all tests need to be given to each person. A majority of people with significant academic problems in which LDs may eventually be a concern have difficulty with word recognition skills. This typically produces problems across the domains of reading, so that assessments beyond the core tests are usually not necessary. Isolated problems with reading comprehension and written expression occur more infrequently. In math, however, more differentiated assessments of subdomains (e.g., whole vs. rational numbers; computational skill vs. problem solving) are required. Also, if the problem is specifically math, using assessments in addition to the WJ, KTEA, or WIAT is helpful in ensuring that the deficiency is not just a matter of attention or other difficulties.

Word Recognition Accuracy

The WJ, WIAT, and KTEA include subtests requiring untimed oral reading of isolated real words and pseudowords, allowing measurement of the person's sight word knowledge and capacity for sounding out words. Most achievement batteries assess recognition of sight words typically ordered for difficulty, which is the essential component for any assessment related to LDs in this domain. These tests are typically the best single predictor of overall levels of academic achievement (Laforte, McGrew, & Schrank, 2014). Fletcher et al. (1996) found the measures of reading accuracy of real words and pseudowords were highly correlated, assessing a similar latent variable and comparable to the reading accuracy scores from other measures of word (e.g., Wide Range Achievement Test–IV [Wilkinson & Robertson, 2006]) and text reading (e.g., the Gray Oral Reading Test—Fourth Edition [GORT-IV; Bryant & Wiederholt, 2001]).

Reading Comprehension

Reading comprehension is difficult to assess with a single measure, and different comprehension tests will give different information about level of performance depending on how comprehension is assessed (Keenan, Betjemann, & Olson, 2008). This is a problem at the observed level of measurement and is closely linked with differences in normative standards. It is important to attend to the nature of the material the person reads as well as the response format. Tests assessing reading comprehension vary what the child reads (sentences, passages, genres [narrative, expository]), response format (cloze, open-ended questions, multiple choice, think-alouds), memory demands (answering questions with and without the text available), and the depth of assessment of the abstraction of meaning (vocabulary elaboration vs. knowledge, inferencing, and activation of background knowledge). If the issue is comprehension and the source contains elements beyond the child's word recognition or fluency skills, a single test is rarely adequate and multiple measures that assess reading comprehension in different ways may be needed.

To illustrate, measures like the WJ Passage Comprehension subtests are best considered screens for achievement in reading comprehension. This cloze-based assessment requires a child to read a sentence or passage and fill in a blank with a missing word. Similarly, neither the WIAT nor KTEA require reading of significant amounts of text. The problem is that some children who struggle to comprehend text in the classroom will not experience difficulties with the reading materials in the WJ, WIAT, or KTEA because the level of complexity does not parallel what children read in the classroom even though these measures all indicate a latent variable involving reading and language comprehension. A good assessment of reading comprehension requires the reading of significant amounts of complex text. For people for whom comprehension is an issue, assessments using the GORT–IV (Bryant & Wiederholt, 2001) or one of the group-based reading comprehension tests from the Iowa Assessments (Dunbar & Welch, 2012) or the Stanford Achievement Test—10th edition (Pearson, 2010), is essential. If a person has had these kinds of assessments in school, the results can be reviewed as part of the evaluation. It is important not to rely only on group tests because the person may not have exerted adequate effort or paid attention, or may have engaged in other behaviors that invalidated the test. We would use a single test as part of a standard battery and expand the evaluation in people who do not show evidence of impairment in basic reading skills.

Mathematics

In mathematics, it is important to assess performance on whole as well as rational numbers; although difficulty with whole-number skill predicts difficulty with rational numbers (Jordan et al., 2013), many students with strong whole-number performance experience problems with rational numbers (D. Fuchs, McMaster, L. S. Fuchs, & Al Otaiba, 2013). It is also important to assess computational as well as problem-solving skill. For computations, Table 4.2 identifies the Calculations subtest of the WJ, the Numerical Operations subtest of the WIAT, and the Math Computations subtest of the KTEA, representing paper-and-pencil tests of math computations that vary in items and complexity from basic arithmetic to algebra and geometry. Poor performance on these calculation tasks reliably predicts variation in cognitive skills associated with math difficulties depending on other academic strengths and weaknesses (L. S. Fuchs et al., 2008b; Rourke, 1993). The challenge is that math difficulties have multiple sources. Poor performance on these tests could reflect problems with fact retrieval and phonological memory if word recognition is comparably lower. In contrast, if word recognition is significantly higher than math performance, the problems may stem from difficulties with procedural knowledge. In any person, poor performance in mathematics can reflect attention difficulties (L. S. Fuchs et al., 2006b), especially in children with comorbid ADHD. The Arithmetic (math computations) subtest of the WRAT-4 is useful because it is timed and the problems are less organized, making it more susceptible to attention and executive function difficulties. The key in math computations is a paper-and-pencil format, which is how difficulties in math are typically manifested in children who do not have reading problems.

Table 4.2 also identifies three relatively similar measures of math problem solving. We routinely measure math problem solving because of the evidence for variations in cognitive skills related to computational difficulties versus word-problem performance and because effective interventions for problem solving are available (e.g., L. S. Fuchs et al., 2010b; see Chapter 8). These tests introduce word-problems that are difficult for children with reading difficulties, especially if they have to read the problem. But because of the role of language in reading (see Chapter 7), and word-problems (e.g., L. S. Fuchs et al., 2010b), children with reading problems often struggle with math problem solving even when the problems are read to them, which is the case for all three measures.

Written Expression

The domain for which assessment guidelines are most difficult to provide is written expression, partly because what constitutes a disorder of written expression is not well established. Does a disorder of written expression primarily involve spelling, handwriting, or composition? Problems with handwriting and spelling will constrain composition, so these domains are related (Berninger, 2004). Table 4.2 identifies spelling subtests in all three batteries, which should be assessed as it may represent the primary source of difficulty with written expression for many children, especially those with word recognition difficulties. An analysis of spelling errors can be informative in understanding whether the problem is with the phonological component of language or with the visual form of letters (i.e., orthography; Rourke, 1993). Asking people to complete spelling tasks also permits an informal assessment of handwriting.

The three assessment batteries in Table 4.2 have measures of written expression. However, the utility of these measures is not well established. From a construct view, the degree to which significant construction and writing of stories and essays is required varies (i.e., composition; Chapter 9). Handwriting and spelling skills constrain written expression for many children with word-level reading difficulties (Chapter 6) or who have motoric difficulties associated with ADHD and other disorders.

Furthermore, it is not established whether children can have isolated problems just with composition. Such problems tend to be more apparent when ADHD is involved and may reflect organizational and self-regulation difficulties (see Chapter 9). The key is writing a composition, which is required by the Essay Composition subtest of the WIAT. An alternative is the

Spontaneous Writing subtests of the Test of Written Language IV (Hammill & Larsen, 2009). We use thematic writing measures (composing stories or essays) for people who present with questions about written expression and who are not deficient in basic writing skills. We always assess spelling as a screen for these deficiencies.

Automaticity

Reading Fluency

Reading fluency measures are typically highly correlated (Barth et al., 2012), but performance can differ depending on whether they simply require timed oral reading of lists and sentences versus timed reading with a meaning component. There is not much evidence of differences because of requirements to read timed word lists versus timed passages. All three achievement batteries have timed reading fluency measures that do not include a meaning component, essentially reading word lists. In contrast, the WJ Sentence Reading Fluency subtest and the WIAT Oral Reading subtest require processing for meaning. Quick alternatives are the Test of Word-Reading Efficiency-2 (Torgesen et al., 2011), which involves oral reading of real words and pseudowords on a list, or the Test of Reading Fluency (Deno & Marston, 2001), which requires text reading. Grade-appropriate CBMs are also reasonable approaches to assessing reading fluency. The Test of Silent Reading Efficiency and Comprehension (TOSREC; Wagner, Torgesen, Rashotte, & Pearson, 2010) is a 3-minute sentence verification assessment that we have found useful in research. All of these measures are quick, efficient, and widely used. The key to assessing reading fluency is to have text read orally, so that fluency can be measured in terms of words read correctly per minute or in terms of items with a comprehension component answered correctly.

Math Fluency

As in reading, assessments of math fluency can be helpful, although there is little evidence suggestive of a specific math fluency disorder, partly because it has not been studied. All three batteries have subtests involving timed assessment of basic arithmetical computational skills that may be helpful for identifying children who lack automaticity in basic arithmetic skills, which can lead to difficulties in mastering more advanced mathematics.

Writing Fluency

Assessments of writing fluency can be informative because they can predict the quality of composition (Berninger & Hart, 1992). Thus, measuring fluency with a measure like the WJ Sentence Writing Fluency subtest, the WIAT Alphabet Writing Fluency subtest, or the KTEA Writing Fluency subtest, may be useful, particularly for screening purposes. We would add these measures for people for whom writing is an issue.

Achievement Patterns

Characteristic patterns will emerge across norm-referenced tests that can help identify the type of LD and indicate specific kinds of intervention. For each of the five types of LDs, there are interventions with evidence of efficacy that should be utilized in or out of a school setting (see Chapters 6–10). The goal is not to diagnose LDs, which is not reliable in a one-shot evaluation for the psychometric and conceptual reasons previously outlined in Chapters 2 and 3, but to identify achievement difficulties that can be addressed through intervention. If the testing professional is knowledgeable about these patterns, very specific intervention recommendations, as well as the need for other assessments, can be provided.

Table 4.3 summarizes six achievement patterns that are well established in research (Fletcher et al., 2005b) that tie directly to the hypothetical classification and the core assessments in Table 4.2 (see also Spear-Swerling, 2015). It should be understood that the cut point is deliberately set high in order to minimize false negative errors (missing people with significant problems). The cut points are not hard-and-fast decision rules, nor are the levels of discrepancy across domains firm. The patterns are the important dimension. We are not indicating that 25% of all children have an LD, only that scores below the 25th percentile are commonly associated with low performance in school, assuming that the cut point is reliably assessed. Response to validated intervention should also be assessed to determine the presence of an LD.

TABLE 4.3. Achievement Patterns Associated with Intervention

1. Word recognition and spelling < 90; math computations one-half standard deviation higher than word recognition and spelling and at least 90. This is a pattern characterized by problems with single-word decoding skills and better arithmetic ability. Reading comprehension will vary depending on how it is assessed, but is usually impaired. Children with this pattern have significant phonological processing problems and often have strengths in spatial and motor skills (Rourke & Finlayson, 1978).
2. Reading fluency < 90 and word recognition one-half standard deviation higher will reflect a problem in which accuracy of word reading is less of a problem than automaticity of word reading (Lyon, Shaywitz, & Shaywitz, 2003). The most reliable correlate is rapid automatized naming of letters.
3. Reading comprehension < 90 and 7 points below word recogniton. This pattern often reflects problems with vocabulary and receptive language, working memory, and attention, with strengths in phonological processing (Stothard & Hulme, 1996).
4. Math computations < 90, word recogniton and spelling > 90 and at least 7 points higher. Children with difficulties that involve only math show this pattern, which is associated with problems with executive functions/attention, working memory, motor skills, and spatial skills; phonological processing and vocabulary are often strengths (Rourke & Finlayson, 1978). If spelling is also < 90, this is essentially the same pattern with a more significant motor problem.
5. Spelling < 90. This pattern reflects (1) motor deficits in a young child or (2) residuals of earlier phonological language problems that have been remediated or compensated in older children and adults. The pattern is common in adults with a history of word recognition difficulties. Fluency is often impaired.
6. Word recognition, reading fluency, reading comprehension, spelling, and arithmetic < 90. This pattern represents a problem with word recognition and math characterized by pervasive language and working memory problems more severe than in children with poor decoding and better development of math skills (Rourke & Finlayson, 1978). It is likely a comorbid association of word recognition and math difficulties.

Note. The patterns are based on relations of word recognition, reading fluency, reading comprehension, arithmetic, and spelling. Any score below the 25th percentile (standard score = 90) is assumed to indicate at least mild impairment. A difference of one-half standard deviation is assumed to be important (±7 standard score points). These patterns are unrelated to IQ scores. The patterns are prototypes; the rules should be loosely applied. Adapted from Fletcher, Francis, Morris, and Lyon (2005b).

Assessing Contextual Factors and Related Conditions

Identifying LDs must take into account factors that extend beyond test scores (see Figure 1.1; Waber, 2010). The decision process should focus on what is needed for intervention. This requires an assessment of contextual variables and the presence of comorbid disorders that influence decisions about what sort of plan will be most effective for an individual child. Low achievement is related to many contextual variables, which is why the flexibility in special education guidelines allows interdisciplinary teams to base decisions on factors that go beyond test scores. The purpose of assessment is ultimately to develop an intervention plan.

General Considerations

As we saw in Chapter 3, most definitions of LDs refer to contextual factors and other conditions as "exclusionary." These components of the definition are designed to identify other causes of low achievement that would not represent LDs, the best example being an intellectual disability or a sensory disorder. Contextual factors could also include low motivation, but this is often secondary to poor performance; whether motivation "causes" or represents a true exclusion is unclear. As part of a comprehensive assessment, factors such as motivation, prior efforts at instruction, school attendance and participation in intervention, failure and retention, and related issues that are essential for an effective intervention plan should be documented.

The exclusions also include conditions where the primary cause of low achievement is related to emotional and behavioral difficulties. Here it can be difficult to determine whether such a condition is a primary cause of low achievement, a comorbid condition, or a result of low achievement. As we noted in Chapter 2, ADHD commonly co-occurs with LDs. Grills et al. (2013) found that assessments of anxiety contributed uniquely to the prediction of intervention response in first- and second-grade children. For both examples, the real issue is not about exclusion, but about the design of effective interventions. Children with ADHD who are treated have higher achievement levels; LDs associated with ADHD tend to be more severe. If a child is struggling to read and is anxious, a treatment program addressing anxiety is critical. Sometimes children struggle to achieve when there are family issues, such as divorce. Again, the issue in the assessment is about the intervention plan.

Other exclusionary factors include poverty and economic disadvantage as well as minority language status. It is important to recognize that exclusions based on economic disadvantage were originally placed in special education legislation because Congress was concerned about the blending of special education legislation with educational programs targeting economically disadvantaged students. Exactly how this determination figures into "causes" of LD is unclear. Cognitive testing does not permit separation of biological and environmental causes of low achievement, especially because poverty affects brain development and cognitive skills develop reciprocally with academic skills. Identifying meaningful differences in the cognitive or brain activation profiles, or intervention response of economically disadvantaged low achievers and low achievers with LDs is difficult. There may well be gene by environment interactions (see Chapter 6), and the heritability of LDs may vary depending on poverty, although this is only known at the group level and would be difficult and questionably meaningful to assess in individual people. Again, it is our view that the most meaningful assessment is instructional response; a person who does not respond adequately to instruction (i.e., intractability) may well have an LD.

Minority language status is another conundrum. Clearly children who immigrate or who grow up in households where the language at home is different from the language of instruction are at risk for achievement difficulties. It is not clear how to differentiate a child with achievement difficulties due to LDs from those due to minority language status. At the very least, many children will require assessments of oral language proficiency and achievement in both languages. A person who cannot perform on achievement tests because he or she does not have proficiency in English should not be identified with LDs.

Specific Assessment Guidelines

Parent and teacher rating scales of behavior and academic adjustment, along with parent-completed developmental and medical history forms, should be routinely obtained. These scales may identify behavioral comorbidities and historical factors (e.g., history of brain trauma) that are important to screen. If there is evidence for behavioral comorbidity, the guidelines for identifying these disorders should be followed, which often include a semistructured interview based on DSM-5 criteria in addition to more specific ratings of the behavioral difficulties of concern. For ADHD, it is important to follow guidelines such as those outlined by the American Academy of Pediatrics (2011). These guidelines include rating scales focused specifically on DSM-IV (now DSM-5) behaviors reflecting inattention and hyperactivity/impulsivity in school and home settings and evidence of adaptive impairment in multiple contexts (e.g., grades, peer relationships, and social relationships). If the broad rating scales suggest an internalizing problem, such as anxiety or depression, interviews and rating scales specific to these areas should be completed, and may include self-report measures. Note that the ADHD guidelines do not recommend identification based on cognitive-processing assessments because of lack of evidence that such measures contribute to identification. Simply referring a child for educational interventions without identifying and treating contextual factors will increase the probability of a poor intervention response.

In other domains, assessments are dependent on the question. If an intellectual disability is suspected, IQ, adaptive behavior, and related assessments consistent with this classification can be administered. The definitions and assessment guidelines from the American Academy of Intellectual and Developmental Disabilities should be followed (Schalock et al., 2010; American Association on Intellectual and Developmental Disabilities, n.d.). These guidelines include intelligence scores on a multifactorial test and adaptive behavior measures (in one of three domains: social, conceptual, and daily living skills, or a composite), and age of onset before 18 years. This three-pronged definition specifically applied 95% confidence intervals to the IQ and adaptive behavior measures, given the measurement error of the test, so that eligible IQ scores could range from 65 to 75. This does not mean that IQ testing is routinely needed and, in fact, is generally not needed for identification of LDs. A person with achievement scores in reading comprehension or math within two standard deviations of the mean (i.e., inconsistent with traditional legal definitions of intellectual disability) or development of adaptive behavior obviously inconsistent with an intellectual disability is unlikely to demonstrate levels of performance on IQ tests consistent with intellectual disability. A score at levels consistent with an intellectual disability would not be interpreted as indicating intellectual disability in the absence of adaptive behavior deficits or strengths in reading comprehension or math that extend beyond the development of basic skills.

Autism spectrum disorders are another example of a disorder with specific assessment guidelines that often include IQ tests because of the co-occurrence of intellectual disabilities. However, specific assessment procedures for autism spectrum behavior are also needed.

Some children with low achievement scores may also have oral language disorders requiring speech and language intervention that will require referral and additional evaluation. Such problems are also commonly seen in LDs, and oral language disorders increase the risk for developing academic problems (Bishop & Snowling, 2004; Snowling & Melby-Lervåg, 2016). Simple assessments with vocabulary measures will help identify children for whom overall language development is an issue and allow for screening of which children may benefit from more formal assessments of intelligence and language development. Again, these problems typically extend beyond the academic domain and represent additional areas that require intervention.

Other major considerations are related to English language learners. People who are struggling to read in their nonnative language should not be considered LD unless there is clear evidence that the problems also occur in the native language. It may be necessary to administer formal tests of language proficiency and academic skill development in the native language and in English to evaluate this possibility. For Spanish, the Bateria III Woodcock–Munoz (Woodcock, Munoz-Sandoval, McGrew, & Mather, 2007) is very useful because it is co-normed in Spanish and English and comparable to the WJ (in English). This question does not need to be routinely assessed in children whose language exposure is exclusively English, but can be a major issue in areas where significant segments of the population are not native English speakers.

Cultural and Linguistic Sensitivity

Any assessment should take into account the cultural and linguistic sensitivity of the measure. A significant discussion of these issues is beyond the scope of this book. Many of these issues are explicitly addressed in the Standards for

Psychological and Educational Standards (Joint Committee on Standards . . . , 2014). Considerations for assessments conducted in schools are also explicitly laid out in IDEA 2004, which indicates that selected instruments must be racially and culturally fair, administered in the native language, used for purposes for which they are reliable and valid, administered *as designed* by trained and knowledgeable personnel, and tailored to area of educational need, adapted to physical and sensory disabilities.

The Standards are even more explicit, recognizing that normative standards are developed under controlled circumstances and that deviations from standard assessment instructions make normative comparisons much less straightforward. An examiner must incorporate this knowledge of the ideal circumstances under which the normative standards were based into any test administration that deviates for the methods used to develop and standardize the measure. For instance, performance differs in predictable ways based on racial and SES characteristics that may not be addressed by the normative sample of many standardized tests based on population characteristics. Issues like dialect must be taken into account. There are guidelines for the assessment of minority language people, with recommendations to always try and assess in the minority language if appropriate tools exist. If translators are used, considerable caution must be used when interpreting the results. In general, it is incumbent on an examiner to be aware of what a test is designed for and when extrapolations are made beyond these purposes. These are generally considered issues of fairness that affect the validity of the interpretation of test scores.

GIFTED (TWICE EXCEPTIONAL) LDs

A persistent criticism of the type of identification and assessment approach that we have advocated is that people who have unusually high aptitude or achievement in one area and meet criteria for "giftedness" will not be identified because their level of achievement will be discrepant with their advanced abilities, but not low enough to be detected or to warrant identification as LD (Reynolds & Shaywitz, 2009). These people are referred to as "gifted LD" or "twice-exceptional" students. In terms of RTI, it is important to note that if the hybrid method we have outlined were implemented, significant achievement discrepancies would be revealed when norm-referenced tests or different CBMs are used. No reasonable person advocates for identification based solely on universal screening and progress monitoring. Significant, uneven development across academic domains can be assessed as well. Moreover, RTI identification frameworks do not preclude referrals for evaluation of concerns that may represent a disability, and it is not necessary to wait until completion of multiple tiers of intervention to make a referral. Students with brain injuries, autism spectrum disorder, intellectual disabilities, and oral language disorders may or may not show achievement deficits in universal screening, but may have difficulties in other areas of adaptation that warrant evaluation, identification, and services as outlined above. Also, if the person is characterized by low achievement in one domain, he or she would be flagged as at risk if he or she meets criteria for low achievement. The key would be identifying the area of high achievement and differentiating instruction across these domains. Even for people who clearly have LDs in one domain with evidence of "giftedness" in another domain, the academic weakness may be restricted such that specialized instruction in a domain characterized by robust achievement would not be necessary.

The concept of "gifted" LD becomes controversial when an achievement discrepancy exists relative to IQ or another cognitive assessment, or if it is based on cognitive assessment with no evaluation of achievement levels but is above an identified threshold for low achievement. A person may have a very high IQ score and outstanding math skills, but average-level achievement in reading. Or the person may show extremely strong visual processing skills despite being only average in phonological awareness. Not surprisingly, many advocates of identification of gifted LDs support intraindividual discrepancy (and IQ–achievement discrepancy) methods, anchoring identification in a cognitive discrepancy method. Thus, Gilman et al. (2013) suggest that

> the emphasis on below-grade-level (or lower) performance, without regard to ability or potential weaknesses, misses twice-exceptional students. Those who perform at grade level, by using advanced conceptual abilities and hard work to compensate, may still require interventions and accommodations to manage increasing educational demands. Otherwise, college and even high school graduation may be out of reach. (p. 1)

These authors advocated an approach that examines skill discrepancies, citing others in the field: “First, a comprehensive individualized evaluation that employs an intra-individual, rather than an inter-individual approach toward ability and achievement analysis is critical” (Foley-Nicpon, Allmon, Sieck, & Stinson, 2011, p. 7).

The notion of people who are both gifted and LD has been criticized from its introduction (Cohen & Vaughn, 1994; Vaughn, 1989) due to the absence of specific identification criteria for twice-exceptional students and for the uncritical acceptance of the hypothesis. Lovett and Lewandowski (2006) noted that different identification and evaluation strategies had been proposed, including IQ–achievement discrepancies, ability subtest scatter-based on IQ test profiles, profile analysis based on patterns across cognitive tests, all of which are problematic. They also observed more general issues with the assessment of giftedness, an elusive term that is generally equated to high IQ test performance.

The problems with cognitive discrepancy approaches are well known (see Chapter 3), and there is no reason to think that identification of this infrequently co-occurring discrepancy is improved by focusing on the extremes of the distribution. In fact, because the measures used to assess the discrepancy are correlated, regression to the mean is even more of a concern. To illustrate using IQ–achievement discrepancy, if IQ and achievement have a population correlation of .6, and a regression-based definition of discrepancy is used with a criterion of a 1.5 standard error difference, achievement would have to be 32 points lower than IQ at an IQ score of 130 to meet this criterion. This occurs for highly reliable tests. Even slight reductions in reliability, such as when scatter on IQ subtests is used, magnifies these difficulties, which is why such methods have been widely discredited for any form of identification (Watkins & Cavinez, 2004).

The legal issues revolve around the standard used to make these judgments, which usually include evidence of adaptive impairment. Most courts dealing with disability use an “average person” standard and focus on evidence of low achievement as an indicator of adaptive impairment. However, in some cases where accommodations have been requested for college or professional examinations, the criterion has been revised to refer to skill discrepancies or variations in the process by which the scores are obtained (Mapou, 2013). The key is still the indication of adaptive impairment. Thus, does a person with a high IQ score who has average reading or math skills have adaptive impairment? This would need careful consideration taking into account the psychometric issues involved in establishing any form of severe discrepancy.

We think the most important considerations for any person are a careful evaluation of achievement skills, hopefully based on progress monitoring and norm-referenced assessments. *Automaticity of skills is an especially important consideration*. In addition, instructional history and response should be considered as outlined above. It is very inappropriate to deny eligibility for accommodations for a person who has received extensive intervention because this will affect achievement levels as well as level of achievement. It is just as important to identify areas of strength as well as weakness in developing intervention plans. There may well be a need for special programming for "gifted" people, but this begs the question of how "gifted" is defined and what type of specialized programming is needed. Similarly, for "gifted LD," what identification criteria are most appropriate? In Chapter 2, we outlined approaches to classification research that can be used to evaluate a classification hypothesis. Unfortunately, this type of research has not been completed and the idea of twice-exceptional students clearly needs a strong empirical evaluation.

CONCLUSIONS: ASSESSMENT

Based on our evaluation of classification models and identification methods in Chapter 3, we proposed a hybrid method that incorporated assessment of instructional response, low achievement, and contextual factors and other conditions for the identification of children as LD. We did not advocate for extensive assessments of cognitive, neuropsychological, or intellectual skills in order to identify children with LDs due to the lack of evidence that such assessments contribute to intervention or that discrepancies on those tests provide information not apparent in profiles of achievement tests (see Chapter 3). Our recommendations concerning assessment assume that the person is old enough to expect that reading, math, and written expression skills have begun to develop. It is entirely appropriate to administer cognitive or neuropsychological tests to children below such developmental ages, particularly in an effort to identify risk characteristics. Even these assessments should be relatively brief and targeted to specific academic areas (e.g., phonological awareness and letter–sound knowledge in kindergarten as predictors of reading ability). In general, LDs should not be identified in preschoolers. Even in grade 1, the reliability of identification will be lower because of maturational issues and the limited floors of many achievement tests in this age range (S. E. Shaywitz et al., 1992).

The heart of the identification model and approach to assessment is the focus on the measurement of instructional response. Although some may see our model as appropriate only for schools, there is little evidence that evaluating a person in a single status assessment based on IQ–achievement discrepancy, low achievement, or patterns on cognitive and neuropsychological tests leads to better intervention. Such assessments do not have direct implications for treatment. Further, if the "diagnosis" is based on a single assessment, it may not be adequately reliable. More important, as soon as it is apparent that the person has an achievement problem, progress monitoring and interventions begins. *People should not be identified with LDs until a systematic attempt at instruction has been made.* However, given the need for more research on what constitutes appropriately intensive intervention, optimal methods for estimating slope and intercept effects, as well as cut points to validly differentiate adequate from inadequate responders, intervention response cannot be the sole criterion for identification. Nonetheless, inadequate response as "unexpected underachievement" epitomizes the essential construct of LDs.

CHAPTER 5

Effective Instruction for Students with Learning Disabilities

A Multi-Tiered System of Supports

In this chapter, we address general principles of instructional design and intervention for students with LDs. We begin by discussing issues related to the implementation of MTSS service delivery models, which we believe is the general education context that can provide the strongest instructional outcomes for all students, including those eventually identified with LDs. Intervention for students with or at risk for LDs cannot begin early enough; prevention and remediation are best integrated through MTSS service delivery models. Because of the key role of strong core classroom instruction, the leadership and involvement of general education is essential. Different entitlement programs cannot be treated as silos and should be part of any approach to service delivery that focuses on all students at risk without regard to labels. Entitlement programs, including special education, are potential resources for these efforts, within the minor constraints established in federal law. We then turn to principles of effective instruction for which we believe there is evidence to guide the development of intensive intervention for students with LDs. This includes progress monitoring, which we regard as an essential element of instruction. We also discuss differences between effective and ineffective instruction.

One thread running throughout this chapter is an emphasis on "validated" instructional programs, which may be researcher-developed and/or commercial. We believe that for a program to be deemed "validated," at least one high-quality randomized control trial must indicate statistically significant and practically important effects on important outcomes and other high-quality randomized control trials must not provide evidence to the contrary. But it is important that practitioners appreciate the fact that no instructional program, even those validated using one or more randomized control studies, works for all students. Schools must therefore assume that validated intervention programs will work for most, but not all, individuals—even when teachers implement those programs with greater intensity than in the validation studies (i.e., more sessions per week or longer duration per session or smaller group size), as in intensive intervention. When inadequate response occurs, we assume that the student has individual needs that are unusual or specialized and that the student requires adjustments to the program to meet those needs.

INTERVENTION OUTCOMES FOR STUDENTS WITH LDS ARE POOR

Disappointing outcomes for students with LDs are unfortunate and avoidable given the wealth of information about how to design intervention productively and conduct assessments in ways that support intensive, differentiated intervention. A strong need exists for schools to reorient service delivery for students with LDs in ways that provide opportunities for these students to receive the intensive intervention they require on the skills and strategies that are foundational to success in and out of school. As these opportunities become available, intervention needs to be designed according to validated principles—or using validated programs—and implemented with strong intensity.

Why do we say that intervention outcomes are poor? Here we refer to the disconnection between what we know from research about effective instruction and its implementation in many public and private school settings (see Chapter 11). Consider, for example, pull-out programs commonly used in public schools and often described as "resource" rooms. Foorman et al. (1997) provided interventions in public school special education resource rooms to students with identified reading disabilities in grades 2 and 3. The students were randomized to one of two programs in which phonics was taught explicitly, one of which was an alphabetic (synthetic) phonics program based on an Orton–Gillingham method and the other an analytic phonics method (Recipe for Reading). Students in these two groups were compared with a group that received an intervention that involved teaching sight-word recognition skill. None of the intervention groups in this study showed gains that could not be predicted based on initial status, with little evidence of robust differences among the instructional groups. As Foorman et al. (1997) showed, conditions in the schools mitigated success. Students received only two-thirds of the planned instruction over the school year; they missed many hours of general and special education instruction; and the size of the instructional groups may have been too large to promote adequate implementation of any of the programs. Interventions shown to be efficacious in highly controlled studies may lack effectiveness when implemented in natural classrooms. Careful construction of evidence-based intervention programs is only half the equation; adequate implementation is necessary for success.

In a naturalistic study, Bentum and Aaron (2003) found that 4 years in resource room placement was associated with no growth in reading and a decline in IQ standard scores. Raw scores increased, so this is not a loss in IQ —just reduced growth relative to the normative group. This phenomenon was clearly illustrated by an analysis of state testing data by Hanushek, Kain, and Rivkin (1998), which found that placement in special education was associated with growth in reading of 0.04 standard deviations per year and 0.12 standard deviations in math. What this means in standard scores is illustrated in Figure 5.1, which shows that, for a student starting in special education at the second percentile in reading, 4 years in special education would be associated with improvement from the second to the third percentile. This change is negligible, especially when the resources invested in special education programs are considered.

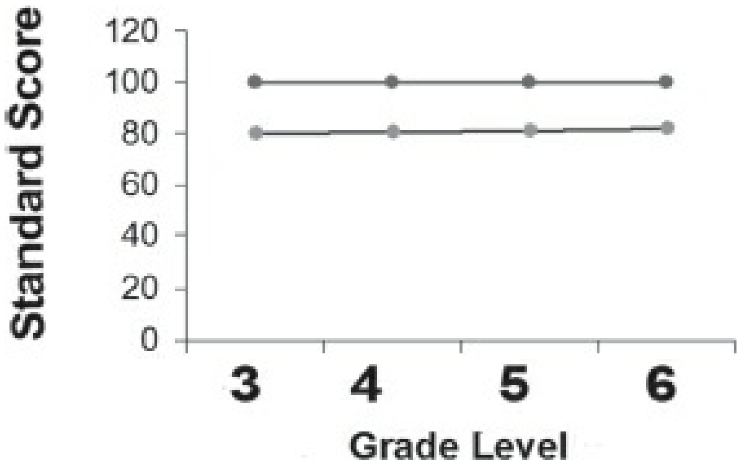

FIGURE 5.1. Changes in reading standard scores associated with growth of 0.04 standard deviations per year when a child is placed in special education from grades 3 to 6. The lines are flat, indicating that the child maintained his or her status on age-adjusted scores, but did not close the gap. Adapted from Hanushek, Kain, and Rivkin (1998). Courtesy Whitney Roper.

Without this type of special education support, however, standard scores may have declined. Vaughn et al. (2012) documented this in comparing growth in reading over 3 years of intervention. Participating students were identified as poor comprehenders in grade 6, but over three-quarters had problems with basic reading skills (Cirino et al., 2013). As illustrated in Figure 5.2, there was slight acceleration in children who received the researchers' intervention, which was based on validated principles of intervention. For students who were struggling readers and served through standard school-based programs, there was a clear decline in standard scores. Note that these programs were not specifically identified as special education programs, but many students were identified for special education.

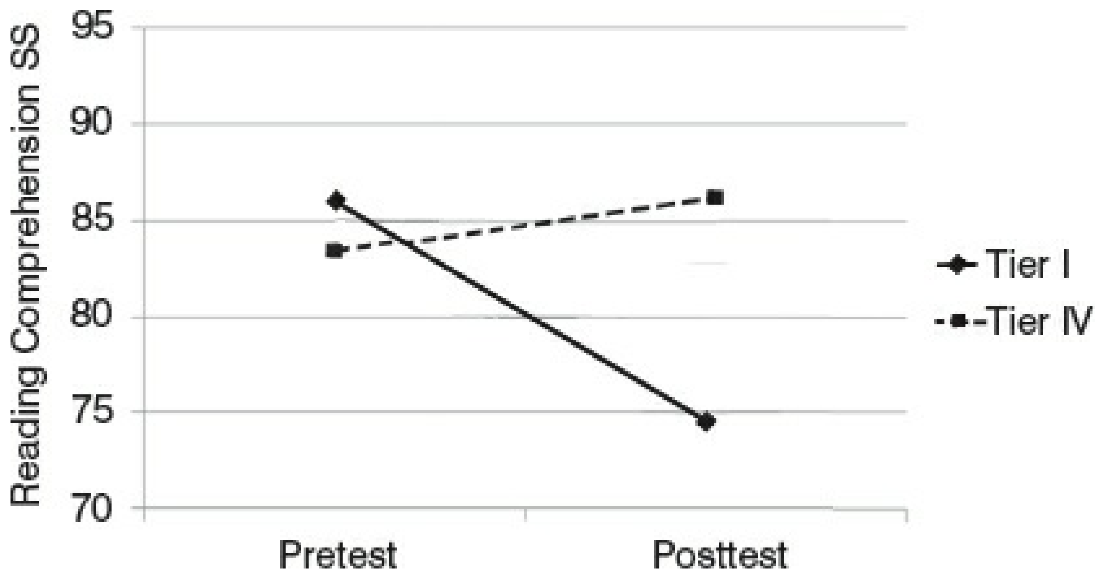

FIGURE 5.2. Changes in reading comprehension standard scores by adolescents in a business-as-usual comparison group (Tier 1) versus adolescents who received 3 years of intervention (Tier IV). Students who received intervention show slightly accelerated gains in reading comprehension, but adolescents who did not receive the research-based intervention show a standard score decline. Data from Vaughn et al. (2012). Courtesy Whitney Roper.

On the other hand, whereas inclusive approaches to special education in which intervention occurs through support in the general education classrooms have been associated with improvements for children with some disabilities, inclusion has not resulted in improved academic outcomes for children identified with LDs. Older research studies (e.g., Vaughn, Moody, & Schumm, 1998; Zigmond & Baker, 1996) suggested that about most students identified with LDs in reading showed little growth in reading despite researcher-supported inclusive practices. In a more recent experimental study, L. S. Fuchs et al. (2015b) randomly assigned fourth-grade students whose prior mathematics achievement was at or below the 10th percentile to receive instruction in inclusive classrooms or to receive intensive intervention delivered in small groups. Instruction focused on students' understanding of and/or procedural skills with fractions and was designed in line with validated principles of explicit instruction as outlined below. The very low-performing students who received inclusive instruction performed significantly and substantially worse than counterparts who received small-group intervention. This was the case even though typically developing classmates in the same inclusive classrooms profited nicely from the same general education fractions instruction that had failed to help very low-performing inclusive students.

Note that we are not advocating for separate placements for students with LDs, which are often unnecessary and may be ineffective from academic and social perspectives for some students. At the same time, we think there needs to be a continuum of services, led by general education, available to *all* students who struggle with learning and behavior in school, as part of a comprehensive effort to include quality intervention in schools. Hence our support for intensive intervention for students with LDs, embedded within a MTSS approach to service delivery that includes special education.

MULTI-TIERED SYSTEM OF SUPPORTS

As we discussed briefly in Chapters 2 and 3, MTSS approaches represent service delivery systems in which schools provide layered interventions as a continuum that begins in general education classes (Tier 1) and increase in intensity in subsequent tiers. In general, greater intensity is achieved through increased time for instruction in smaller groups of students with more differentiation of content. Figure 5.3 provides a schematic of a three-tier service delivery model, commonly represented as a triangle. We have tipped the triangle to emphasize the idea that Tier 1 instruction is for all students, with subsequent tiers increasing intensity for students who do not respond adequately and need additional intervention. Students always begin with differentiated instruction in the general education classroom (Tier 1). Tier 2 is typically small-group instruction, which is most effective if it is aligned with the general education curriculum and delivered using a standardized protocol (Foorman, Herrera, Dombek, Schatschneider, & Petscher, 2017). Tier 3 should be very intense—smaller groups or even individualized instruction, sometimes less aligned with core instruction (e.g., an intense focus on decoding because the child is not learning these skills).

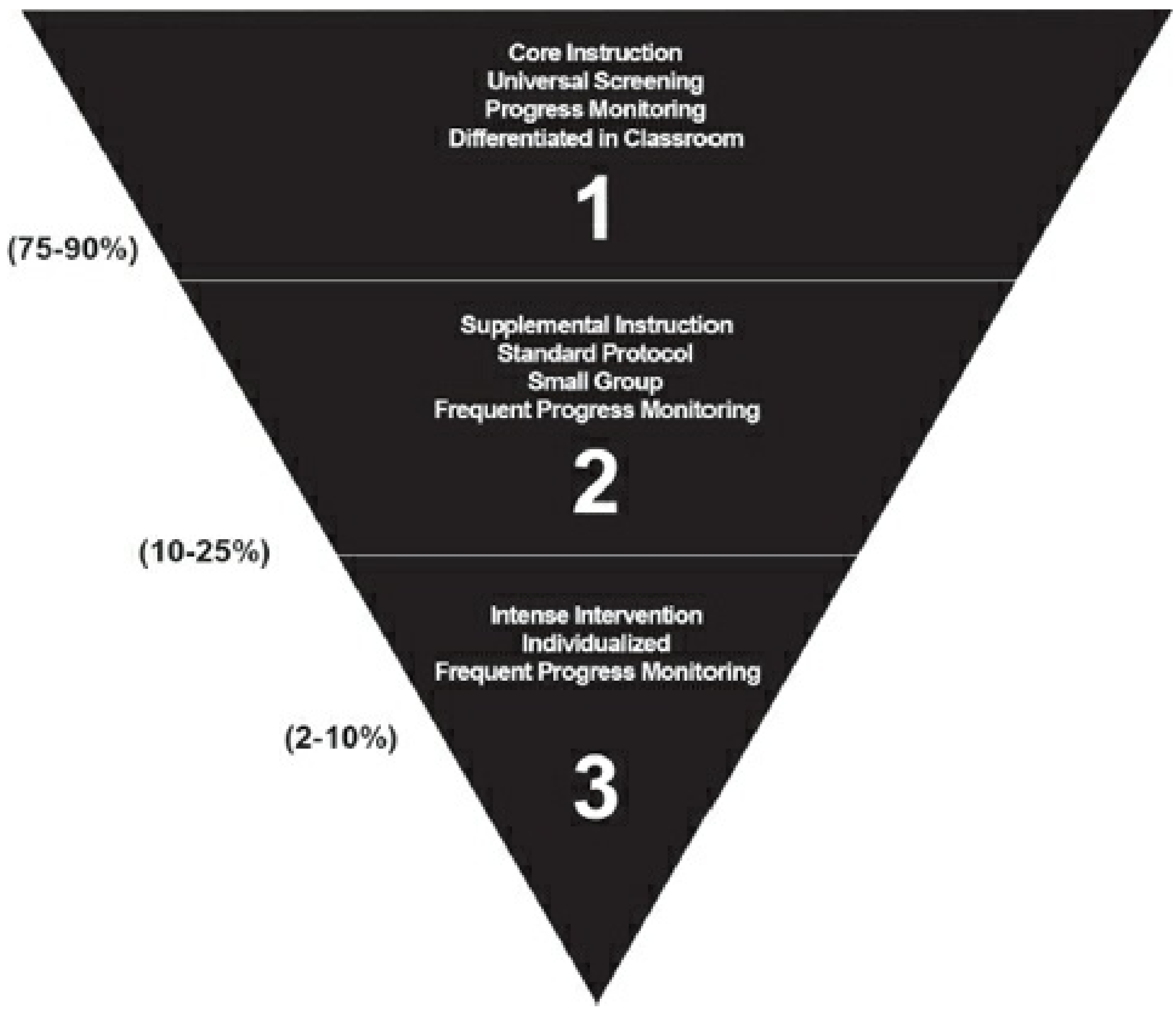

FIGURE 5.3. A three-tier model for service delivery in an MTSS. Courtesy Whitney Roper.

Also common to most MTSS frameworks are (1) universal screening of all children for academic and behavioral difficulties beginning with school entry; and (2) progress monitoring for students identified as at risk (see Figure 3.8; Fletcher & Vaughn, 2009; Vaughn & Fuchs, 2003). The effect is to put all students in a school in a surveillance system where their risk status and learning progress can be quickly identified based on performance data. Children who do not respond adequately may be referred for a comprehensive evaluation for eligibility for special education, which in some models occurs after Tier 2 and in other models follows another round of Tier 2 or more intensive intervention (Tier 3 and beyond). In reality, however, referral for a comprehensive evaluation could occur at any point in the MTSS process, especially if there is a question about oral language difficulties or minority language status, severely low initial academic performance, or the presence of another disability. The timing of the referral to special education is driven by intervention–response concerns.

Origins

Multi-tiered frameworks were significantly influenced by public health models of disease prevention that differentiated primary, secondary, and tertiary levels of intervention, which increase in cost and intensity depending on the person's response to treatment (Vaughn, Wanzek, & Fletcher, 2007). There are two historical origins of school-based implementations of MTSS that began with efforts to prevent behavioral and academic problems. The first origin has its roots in schoolwide efforts to prevent behavior problems (Donovan & Cross, 2002). These models are associated with a *problem-solving* process. This approach involves shared decision-making teams that identify a behavior problem. The team meets and proposes strategies to address the problem. The team has methods for evaluating the results. If the intervention is not successful, the team determines whether the problem is with implementation of the plan, in which case assistance is provided. Alternatively, the student may need a different approach or a more intensive intervention (Kovaleski et al., 2013; National Association of State Directors of Special Education, 2006).

The second origin is rooted in efforts to prevent reading difficulties in children. These implementations typically involve *standardized protocols* to deliver interventions, which increase in intensity and differentiation depending on the student's instructional response (Vaughn & Fuchs, 2003; L. S. Fuchs & Vaughn, 2012). In a common implementation of a standard protocol model (Figure 5.3), all students are screened and those at risk for academic problems are assessed frequently (every 1 to 4 weeks) on short-duration measures designed to assess progress over time (Stecker, Fuchs, & Fuchs, 2005). Classroom teachers receive professional development in effective instruction and ways to enhance differentiation and intensity through flexible grouping strategies and evaluations of progress (Tier 1, primary intervention). Children who do not achieve specified levels of progress based on local or national benchmarks receive additional instruction in small groups of three to five students for 20–40 minutes three to five times per week (Tier 2, secondary intervention). If students do not make adequate progress in secondary intervention, an even more intensive and more individualized intervention (Tier 3, tertiary intervention) is provided, which may involve smaller groups or 1:1 instruction, increased time in intervention (45–60 minutes daily), and/or a more specialized teacher. It may also involve adaptation of the standard protocol, based on weekly or biweekly progress monitoring, to individualize the protocol to better address students' learning needs (Powell & Stecker, 2014). These models for reading and math intervention link with special education because inadequate instructional response allows for determination of adequate and inadequate responders, a key to the assessment approach outlined in Chapter 4. These models thus link intensive intervention with classroom instruction in an integrated approach to service delivery (Fletcher & Vaughn, 2009).

Is MTSS Effective?

There are many reasons to implement MTSS frameworks, including the effort to improve academic and behavioral outcomes in all children. For students with LDs, MTSS frameworks offer several advantages. First, the focus shifts from who is eligible to concerns about providing effective instruction. Identification is not dependent on teacher referral, which is known to be biased toward behavioral difficulties, leading to overidentification of males and minorities as LD (Donovan & Cross, 2002; Shaywitz, Shaywitz, Fletcher, & Escobar, 1990). Second, MTSS frameworks allow placement of students into intervention immediately rather than after time-consuming and often delayed expensive assessments. Third, if a referral is made to special education, the RTI component provides data indicating how the student has responded to various interventions. Fourth, the adequacy of instruction has been measured through systematic collection of data.

An important key to effective implementation of MTSS models is strong core classroom instruction. Although effective Tier 1 instruction reduces the number of students at risk, significant numbers of students (as many as 20–25% in early reading; Vaughn, Wanzek, Woodruff, & Linan-Thompson,

2006) require supplemental interventions by trained personnel (e.g., classroom teachers, paraprofessionals). Over the past 20 years, many school districts have implemented MTSS models from kindergarten to high school (Jimerson, Burns, & VanDerHeyden, 2015), with the implementation of the System to Enhance Educational Performance (STEEP; Witt, Daly, & Noell, 2000) in Vail, Arizona, and different Mississippi public schools representing a strong example (VanDerHeyden et al., 2007). Unfortunately, we are not aware of strong *experimental* trials addressing the effectiveness of approaches based on MTSS. An experimental study would require randomization at the level of the school or district, which would be costly and difficult to implement. However, some demonstrations of layered interventions with multiple tiers beginning in general education and continuing for at least one subsequent tier suggest improved achievement and behavior in elementary school children, reduced special education referrals and placements, and other positive outcomes stem from multi-tiered education (see Jimerson et al., 2015; VanDerHeyden et al., 2007).

Implementation Issues

Implementation of MTSS frameworks has been a struggle for many schools. Strong core instruction and an early intervention program that reduces the number of at-risk students who feed into remedial programs at the middle or high school level are keys to effective implementation. These must be considered prerequisites to any successful implementation. This is in part because implementing MTSS programs at the middle or high school level involves students who have established LDs often related to inadequate instruction in traditional systems and require more intensive and more sustained intervention. Implementing MTSS at the middle and high school level often fails due to a tendency to try to implement MTSS models based on elementary school principles (Vaughn & Fletcher, 2012). One reason is that students are often much further behind their peers at the middle or high school level. Another is that growth is slower at the middle or high school level, so progress monitoring can be conducted less frequently and changes in intervention strategies can be implemented less frequently (Tolar et al., 2014). A third is that universal screening conducted in grade school MTSS models is probably less important for older students because after grade 3, schools often have statewide accountability and other measures in place that can be used as screens. These assessments provide an efficient measure of older students' skills and make the time taken for universal screening unnecessary (Vaughn & Fletcher, 2012). We cannot overemphasize the point that students at the middle or high school level require greater intensity and more sustained intervention to produce meaningful change, so moving through a series of tiers to get to intensive intervention can be inefficient. Even at the elementary level, some students are so far behind that they need to be triaged immediately to more intensive and more sustained intervention.

The difficulties with scaling were clearly demonstrated in a recent study of the implementation of MTSS approaches for elementary school reading (Balu et al., 2015). The study employed a reference sample of representative schools in 13 states to evaluate service intensity and whether more services were provided to poorer readers. A second sample represented 146 schools that self-reported implementing an MTSS model for at least three years. Using a regression discontinuity design, the researchers identified students who scored just below the school-specified cut point for intervention services and compared them to students who scored just above the cut point. It would be expected that the group scoring just below the threshold would demonstrated better reading skills than students scoring above the threshold. The controversial finding was that students who met eligibility criteria did not show improvement in reading skills and in grade 1, showed small, but negative impacts of intervention. However, the study was not an evaluation of the MTSS framework, instead asking a limited question: Does the use of a universal screening system improve student learning? It did not address the more important question: Does the use of MTSS improve student learning? Implementation of MTSS was inadequate. The amount of intervention time provided to students identified as at risk for reading failure was about 6 minutes. In many implementations, the teachers delivering services did not have specialized training in reading intervention, with most schools relying on classroom teachers to deliver Tier 2 intervention (see D. Fuchs & L. S. Fuchs, 2017, for a more complete analysis of the Balu et al., 2015, evaluation).

Differentiating Tier 1, 2, and 3 Instruction

There is nothing magical about three tiers of instruction. Mostly commonly, the first two tiers occur in the context of general education, with the third tier being special education. Sometimes special education is a more intense level of intervention, but it may also represent access to entitlements for accommodations and related services for a person with an identified disability.

The key is to attend to the periodic or ongoing progress-monitoring data that accompany intervention implementation. This provides the basis for schools to distinguish adequate responders from inadequate responders (instructional response is on a continuum and often resource-driven) and for teachers to move to more intense interventions when the rate of improvement falls short of school benchmarks. Some implementations rigidly adhere to a sequence of tiers even when it is clear that the student needs a more intense intervention. As we discuss below, the data should drive decision making and permit direct movement to more intensive intervention when warranted.

Tier 1

Tiers differ in the level of intensity, which is moderated through differences in instructional time, group size (dosage), the learning environment, and instructional individualization. In Tier 1, instruction occurs as part of the core general education language arts and math programs. In elementary school, language arts programs involving reading processes like word study, comprehension, and automaticity, as well as writing, often occur for 90 minutes of the classroom day. Math should commonly receive 45–60 minutes of allocated instruction with the content of reading, writing, and math varying depending on grade and the development of the child. Through progress monitoring, the teacher can determine who is making inadequate progress and increase differentiation and intensity, within the constraints of what’s possible in the classroom, by creating smaller groups that are homogeneous in terms of learning needs and teaching to small groups as well as large groups. There is extensive discussion of effective Tier 1 instruction in publically available materials (e.g., Texas ReadSource [*www.texasreadsource.org/PDSRIWebApp/jsp/index.jsp*]; RTI Action network [*www.rtinetwork.org/essential/tieredinstruction/tier1*]). Core math instruction is described in Chapter 8.

Tier 2

Tier 2 typically changes the learning environment, increases dosage by providing small-group instruction with three to five students and a tutor, and focuses on specific aspects of reading, writing, and math with which the students in the small groups struggle and are placed together because of similar instructional needs for 20–40 minutes per session, usually three to five times per week. This occurs in addition to Tier 1 instruction, so the dosage has been increased. Smaller groups permit increases in the explicitness of instruction as well as the opportunity to incorporate attention to self-regulation and other executive functions the student has not yet achieved. Smaller groups also increase students' opportunities to respond and receive corrective feedback. Progress monitoring and instructional changes are more frequent, and feedback and cumulative review are more targeted.

Tier 3

Tier 3 involves greater instructional dosage by increasing time on task and reducing the size of the instructional group. Often instruction at this level is one on one, although groups of two or three students are also common. Here the specificity of skill focus may also increase to the subset that the specific student has not mastered (e.g., informational text comprehension, understanding fractions). There is a greater emphasis on adapting an intervention program using a validated data-based individualization approach to tailor strategies to the child's specific needs (D. Fuchs, L. S. Fuchs, & Vaughn, 2014b).

In some MTSS approaches, additional tiers beyond these three are used, but we feel that the differentiation provided by a three-tier model captures the essential distinctions in instructional intensity that need to be made. This involves strong core classroom instruction, which should be effective for most students. It relies on a second level of intervention to provide increased dosage in small-group intervention for struggling students by changing the classroom environment and relying on a validated standard protocol. The third tier is a higher dosage, more targeted, often more sustained, and commonly more individualized environment.

Some Examples from Research

To illustrate the value of MTSS, we present two examples from elementary schools. In the first example, three tiers of reading intervention were provided over 2 years to students who were entering first grade (Mathes et al., 2005; Denton, Fletcher, Anthony, & Francis, 2006b). For both of these studies, the entire first grades at six elementary schools were screened over two successive years with the Texas Primary Reading Inventory (TPRI; Foorman et al., 2004) and a list of sight words to reduce false positive errors. Children at risk were then randomized into one of three groups: a group that only received enhanced Tier 1 reading intervention versus two Tier 2 pull-out interventions delivered daily for 40 minutes for about 30 weeks. There were about 30 Tier 1 first-grade teachers and six Tier 2 teachers, who taught six groups daily, three at each of two different schools. Progress was monitored for all students in the study, including the three at-risk samples and a fourth group of not-at-risk students randomly selected from the large sample of children who passed the TPRI screen. There were about 100 students in each of these four groups.

For the Tier 1 group, classroom teachers participated in a district-lead professional development program focused on a comprehensive approach to classroom literacy, with emphasis on explicit instruction. The district purchased new basal reading curricula with additional professional development in their implementation using publisher resources. Teachers, principals, and parents received progress-monitoring graphs each 9-week grading period. The intervention teachers and researchers were available for coaching for all students.

The two Tier 2 interventions were *comprehensive,* meaning that the content included word study, reading for automaticity, and comprehension lessons. Both were *explicitly* taught and delivered in *small groups* of three students during school times when language arts and math were not being taught. Differentiation occurred depending on student needs and was the basis for forming groups. The differences were pedagogical. One intervention (Proactive) was a manualized, scripted intervention with an explicit scope and sequence. For the other intervention (Responsive), students were taught in the context of reading and writing. There was no scope and sequence and teachers were expected to use their own judgment in determining intervention needs. There were sets of activities for teachers to use when these determinations were made.

The results consistently demonstrated better outcomes in word reading, fluency, and comprehension for the two groups receiving Tier 2 intervention. There were no major differences between the two intervention groups except in domains reflecting instructional emphasis (e.g., the Proactive group had slightly stronger word decoding skills). However, as Figure 5.4 demonstrates, the progress-monitoring data revealed clear improvements in the Tier 1 group, which also closed the gap relative to typically achieving students. Indeed, an inspection of norm-referenced assessments at the end of the year showed reading scores well above average in the typical group, and in different parts of the average range for the three at-risk groups. Successful implementation of an MTSS model should lead to improved achievement in all students.

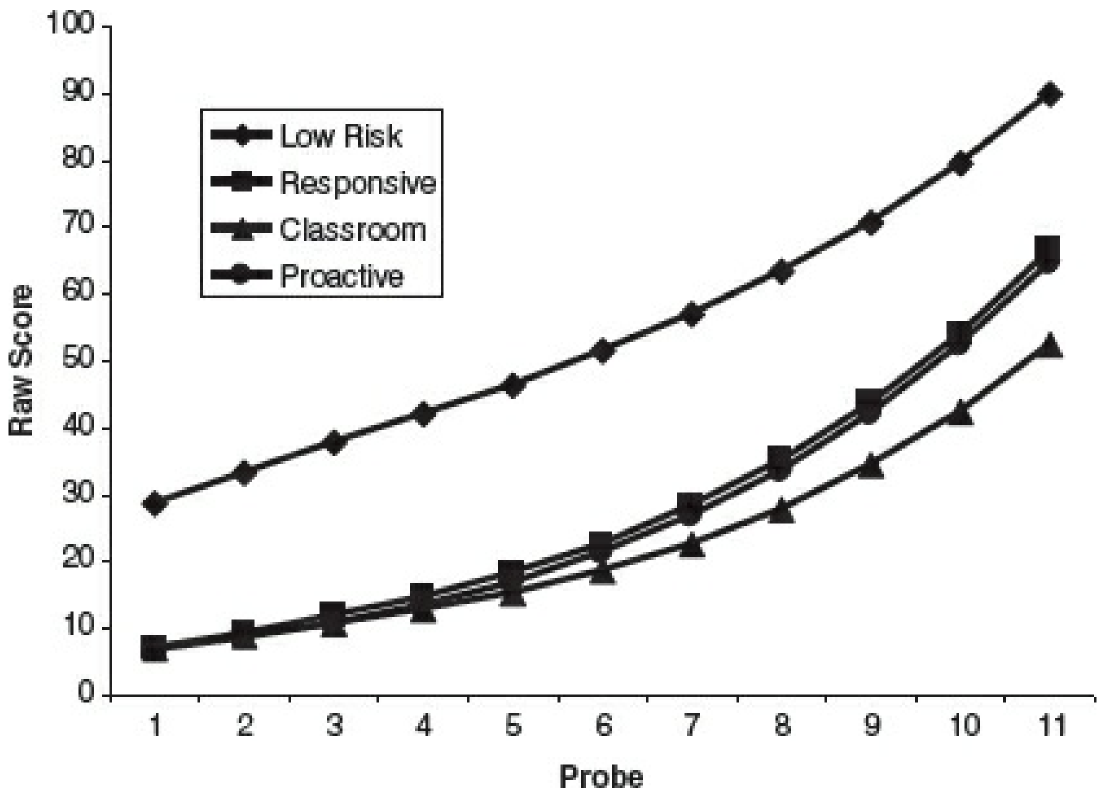

FIGURE 5.4. Growth in reading fluency based on curriculum-based assessments every 3 weeks for students in grade 1 who were (1) identified as being at low risk for reading problems, (2) participated in one of two small-group interventions (responsive, proactive), or (3) received only enhanced general education classroom intervention. The groups that received the small-group interventions showed faster rates of growth and higher end-of-year performance as compared with the at-risk group that received only enhanced classroom instruction. From Mathes et al. (2005, p. 169). Copyright © 2005 International Literacy Association. Reprinted by permission.

At the end of the intervention, there were seven out of 165 (4%) students who received Tier 1 and Tier 2 instruction and 15 out of 92 (16%) who received only Tier 1 instruction that did not meet criteria for adequate response based on word-reading scores below the 30th percentile. An additional five students did not meet benchmark fluency criteria. Since the screener was designed to identify the bottom 20% of students as at risk, this means that Tier 1 instruction reduced the at-risk group from 20% to 3.2% of the school population; for those receiving Tier 1 and Tier 2 instruction, the number was below 1%.

Denton et al. (2006b) took a subset of these students, now in grades 1 to 3 (because of retention), and added some additional very poor readers who had not participated in the intervention. These additional 27 students were placed in an intensive intervention using a program based on intensive word-level skills (Phono-Graphix; McGuiness, McGuiness, & McGuiness, 1996; *www.phono-graphix.com/research.php*). For this intervention, students worked in groups of two students for two 50-minute sessions per day for 8 weeks. The intervention was based on Torgesen et al. (2001; see Chapter 6), who obtained impressive results for word reading and comprehension using this highly intensive approach, but with different programs. Because Torgesen et al. did not get significant gains in fluency, a second 8-week, 1-hour-a-day intervention focusing on a modified version of Read Naturally (Ihnot, Mastoff, Gavin, & Hendrickson, 2001) with additional comprehension components. The intervention took place in the same schools as the Tier 2 study and was delivered by the same intervention teachers trained for 2 weeks in Proactive and Responsive approaches before school began. Students entered the study in a staggered 8-week design so that an initial untreated group could serve as a comparison. Students began the intervention with word-reading scores, on average, around the 15th percentile.

Relative to baseline, there were significant gains in word reading, fluency, and comprehension corresponding to the nature of the two 8-week interventions. In addition, about half the intervention group (including students taught with the Proactive approach and those taught with the Responsive approach) showed reading scores above benchmarks at the end of the intervention. Figure 5.5 shows the individual standard score gains in word reading, with an average of about one-half standard deviation. About half the students made very significant gains in reading, with others also showing satisfactory improvement, leaving a small number of clearly inadequately responsive students. It is noteworthy that the intensive intervention was accomplished in the context of typical schooling, showing that intensive intervention is possible in schools. In addition, it is noteworthy that the Responsive intervention was part of a subsequent scaling study in which the screening, progress monitoring, and Tier 2 curriculum were introduced. The support was reduced over time until this new set of schools was implementing MTSS without researcher support. There was only a small drop-off in gains relative to the researcher-introduced implementations (Denton et al., 2010).

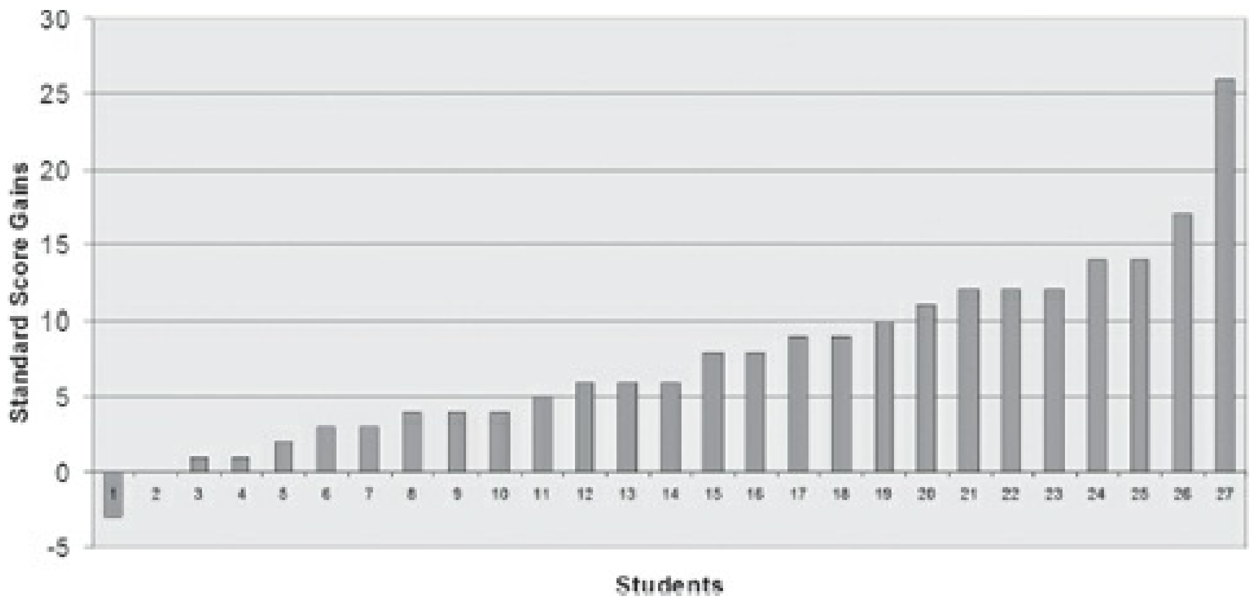

FIGURE 5.5. Standard score changes on measures of word reading for each student who participated in the Tier 3 reading intervention in Denton et al. (2006b). The average change was about 8 points (0.5 standard deviations) and the range was –3 points to +26 points. Courtesy Carolyn Denton.

In math, L. S. Fuchs et al. (2008a) assessed the effects of Tier 2 (small-group tutoring) with and without validated classroom instruction at Tier 1 on at-risk students' math problem solving. Stratifying within schools, 119 third-grade classes were randomly assigned to conventional or researcher-validated problem-solving instruction (referred to as "Hot Math"). The validated classroom program occurred at the whole-class level twice weekly and was implemented as part of the teachers' classroom mathematics program. Across both classrooms, 243 students who were identified as at risk were randomly assigned, within classroom conditions, to receive Hot Math tutoring or not. Students were tested on problem-solving and math applications measures before and after 16 weeks of intervention. Tutored students who received validated classroom instruction achieved higher math scores than tutored students who received conventional classroom instruction, with a large effect size of 1.34 standard deviations. At the same time, the effect size favoring at-risk students who received validated tutoring achieved over at-risk students who did not receive validated tutoring was 1.18, demonstrating added value for at-risk students at both tiers.

These studies demonstrate approaches to multi-tiered instruction in reading and math in which there is a clear differentiation among the tiers. At the level of Tier 2 or Tier 3 intervention, the standard protocol intervention programs incorporate validated principles for students with LDs. In the next section, we highlight this set of principles. We first address intervention design, as reflected in validated programs. Then we discuss assessment that supports effective instructional design to meet individual student needs.

PRINCIPLES FOR DESIGNING INTERVENTIONS FOR STUDENTS WITH AND AT RISK FOR LDs

Table 5.1 summarizes 10 principles that distinguish intervention for students who may be identified with LD or at risk for identification because they are struggling to develop academic skills. The first is *instructional explicitness*. Torgesen (2004) described explicit instruction as "instruction that does not leave anything to chance and does not make assumptions about skills and knowledge that children will acquire on their own" (p. 363). Instruction is explicit when teachers tell students what they need to know with direct explanations, formally sharing new knowledge and modeling the use of the skill or strategy. Explicitness is facilitated by provision of background knowledge and vocabulary, advance organizers, guided and independent practice, corrective feedback, and maintenance checks. Teachers plan lessons with clear goals that progress deliberately from less to more challenging skills and content (Denton, Fletcher, Taylor, Barth, & Vaughn, 2014). They model ("I do"), provide guided practice ("We do"), provide independent practice ("You do"), and check for maintenance (Vaughn, personal communication).

TABLE 5.1. Principles of Effective Instruction for Students with LDs

1. Instructional explicitness
2. Minimization of the learning challenge
3. Proper terminology
4. Speeded practice
5. Cumulative review
6. Simple and direct language
7. Incorporation of self-regulation

strategies

8. Comprehensive instructional approaches
9. Extended duration and time on task
10. Progress monitoring

In thinking about the need for explicit instruction, many typically developing students profit from general education programs that rely, at least in part, on a less implicit instructional style that is oriented toward discovery and inductive learning. Students who develop LDs, however, have failed to profit from such programs and require a different approach. In an analysis of the National Reading Panel (National Institute of Child Health and Human Development, 2000) meta-analysis, Stuebing, Barth, Cirino, Francis, and Fletcher (2008) found a clear advantage for reading programs that included explicit instruction compared to reading programs that were less explicit. They suggested that explicitness occurs on a continuum and can take a variety of forms. Although organizing instructional plans is helpful to teachers, they suggested that there was, in fact, little evidence showing that manualized or "systematic" instruction was superior to less scripted programs emphasizing teacher judgments in the context of reading and writing, so long as the instruction was explicit and comprehensive. For example, Mathes et al. (2005) at Tier 2 and Torgesen et al. (2001) at Tier 3 obtained similar results for explicit approaches that were scripted and systematic versus explicit programs that taught reading in the context of reading and writing, but with no script. None of these programs would be considered examples of discovery learning, constructivist, or inductive.

In math, a meta-analysis of 58 mathematics studies (Kroesbergen & Van Luit, 2003) revealed that, like students with LDs in reading, students with mathematics LDs benefit more from explicit instruction than from discovery-oriented methods. The explicitness principle for LD intervention is incorporated within virtually all mathematics programs with proven efficacy. Instructional steps within an explicit approach include the following:

- Begin by sharing worked examples (e.g., mathematics problems, completely solved and showing all supporting work; completed text summaries, with marked text and notes revealing how the summary was derived).

- Explain to the student how the worked example was completed, step by step, and what the teacher was thinking when completing each step.
- Provide a list of these steps. Help the student apply the steps to the worked example and memorize the list.
- Post the steps and fade the list so the student refers to the poster only when needed.
- Present the same example, with one step of the strategy missing (partially worked example). Require the student to complete that step and to explain how/why he or she does it the prescribed way.
- Give the student opportunities to practice that step of the strategy, so he or she generates many correct responses.
- Gradually fade steps from worked examples, so the student assumes responsibility for more steps.
- Once the student can independently complete entire examples and explain his or her correct responses, build fluency and plan for maintenance.
- Explicitly teach for transfer instruction. With explicit transfer instruction, teachers explain how novel features in text and in mathematics problems can make each seem unfamiliar, even though those novel features are irrelevant; that problem or text is the same problem type or text structure the student has already mastered. Teachers explicitly teach the student to search novel material to understand how it fits with the types of text or problems he or she already knows how to handle. Teachers also present examples emphasizing the same problem type or text structure, as irrelevant features change from one example to the next. In addition, teachers provide practice in sorting novel problems and text in terms of irrelevant changes, and teachers gradually increase the challenges associated with these irrelevant features to increase transfer distance. Explicitly teaching for transfer is critical for students with LDs, because many studies demonstrate the challenge of transfer for these students. (L. S. Fuchs et al., 2005)

Explicitness is not, however, sufficient. A second principle of effective intervention is *instructional design that minimizes the learning challenge*. The goal is to anticipate and plan instruction to avoid confusion and therefore sabotage long-term success. Instructional design that minimizes the learning challenge is accomplished by a task analysis that provides the most efficient method for succeeding with the instructional objective or standard. By *efficient,* we mean the simplest and quickest strategy for producing correct work. Unfortunately, teachers often introduce students to highly inefficient strategies. For example, word problems in the primary grades are often taught with a key word strategy. Students learn that if the word *more* appears in the problem, they should add to find the answer. This is a highly inefficient method for the following reason: the word *more* signals the need to add in only 50% of primary grade word problems. So although this strategy is simple and quick to apply, it fails the efficiency test because it often produces an incorrect answer.

A more efficient instructional design is required. In reading, a scope and sequence often ensures more efficiency by the teacher. Organizing phonics rules, comprehension strategies, and automaticity practice as units may promote efficiency. In the math area, almost all primary grade word problems fall into three categories: *combine* word problems (e.g., "Harry has 5 crayons. Jose has 4. How many crayons do they have?"); *compare* word problems (e.g., "Harry has 5 crayons. Jose has 4 less than Harry. How many more crayons does Harry have?"—notice that subtraction is required even though *more* is in the problem); and *change* word problems (e.g., "Harry had 5 crayons. Tomorrow he'll get 4. How crayons will he have then?"). Helping students understand this scheme provides the means for teaching a strategy for only three problem types, rather than expecting students to view each word problem as a unique challenge. However, identifying this categorical scheme for instruction is just the first step in an efficient instructional design. The teacher then needs to identify efficient strategies, which straddle the three problem types to the maximum degree, while teaching students strategies to reliably differentiate among the problem types and apply strategies that are specific to each problem type. The teacher's responsibility is to design this set of strategies or to obtain a program that organizes instruction in a highly efficient manner.

A third instructional design principle of effective intervention, especially for students with LDs, is teaching and encouraging students to use *proper terminology* for key concepts or procedures (e.g., *irrelevant information, main ideas*). This can ease the learning burden by giving students the vocabulary that most essentially captures important concepts and procedures. Once these terms are taught, the teacher and student should use them consistently. In both reading and math, key vocabulary should be taught explicitly, particularly because students who are behind and in need of remediation lag in their development of vocabulary and background knowledge, which are highly related (Ahmed et al., 2016).

The fourth and fifth instructional design principle of effective intervention for students with LDs concern practice—*speeded practice* and *cumulative review*—which are essential instructional ingredients for most learners, but especially for students with LDs. For these students, practice needs to be designed to develop fluency with foundational skills, with the goal of freeing up attention for higher-order aspects of the learning task. Some research suggests an important role for *speeded practice*. In reading, the efficacy of timed repeated readings of same and different passages is well established as a general practice (National Institute of Child Health and Human Development, 2000; Therrien, 2004) and specifically for students with LDs (Chard, Vaughn, & Tyler, 2002).

In a recent randomized control trial on math instruction, L. S. Fuchs et al. (2013b) contrasted the efficacy of number knowledge tutoring with speeded versus nonspeeded practice on at-risk students' development of arithmetic competence. Tutoring occurred for 16 weeks, three times per week. In each 30-minute session, 25 minutes were identical in the two conditions. The difference between tutoring condition occurred in the last 5 minutes of each session: practice was either speeded or nonspeeded. This seemingly small distinction between conditions resulted in a substantial difference in outcomes. The posttest arithmetical skill of children in the speeded practice condition was one-half standard deviation stronger than for children in the nonspeeded condition, and speeded practice helped students achieve greater reliance on retrieval (the most sophisticated strategy for deriving answers to arithmetic problems and a characteristic weakness in students with mathematics LDs). Moreover, in the nonspeeded condition, learning was weaker for children with limitations in reasoning ability than for children with stronger reasoning ability in the same intervention condition. In contrast, students in the speeded condition responded similarly well, regardless of their reasoning ability. This shows how instructional design can compensate for the types of limitations students with LDs often experience in cognitive and linguistic abilities (see principle 2).

Cumulative review must be integrated in a systematic way to ensure retention of previously mastered content and to help students make effective discriminations among related types of problems or decoding patterns or reading genres, and so on. This type of review is especially important for students with LDs, who often need assistance with retention and retrieval. Cumulative review also helps connect different parts of the overall instructional program, which needs to be explicit for students with LDs.

The sixth instructional design principle for students with LDs addresses the kinds of limitations in oral language these students often experience when identified with LDs in reading, math, and written expression. Oral language difficulties are often comorbid with reading and writing disabilities (Bishop & Snowling, 2004; Snowling & Melby-Lervåg, 2016); when students have word problem difficulties in math, language variables are often associated (L. S. Fuchs et al., 2006a). Even when an oral language problem has not been identified as an area of disability, students with LDs are often inefficient processors of language. Therefore, the *language of instruction must be simple and direct.* Teachers must use short sentences, the active voice, unambiguous pronoun antecedents, and other methods for communicating clearly and simplifying explanations. Teachers should also require students to repeat explanations in their own words, while the students incorporate important terminology. This is one strategy for frequently checking students' understanding of the material at hand.

A seventh principle concerns the difficulties these students have with attentional control, motivation, and self-regulation, which may adversely affect their task-orientation, persistence with challenge, and learning (L. S. Fuchs et al., 2005a, 2006a). By the time students enter intensive intervention, they have experienced repeated failure, causing many to avoid the emotional stress associated with reading or mathematics. They no longer try to learn for fear of failing. For this reason, intensive intervention must incorporate *motivators to help students regulate their attention and behavior and to work hard,* and for many students, tangible incentives are required.

The eighth principle is *comprehensiveness*. All too often, instruction for LDs is skill-based and narrow, focused on learning phonics rules or math facts. In fact, as the reading example above shows (Mathes et al., 2005), instructional programs in reading are more effective if they address all three major components of learning to read, especially in children impaired at the basic level: word recognition, automaticity, and comprehension. In math, L. S. Fuchs et al. (2008a) found that students struggling with math learned math facts as well in the context of word problems as they did with targeted instruction in math facts. Note that learning math facts in the context of word problems may develop more flexible, transferable, and efficient math skills and strategies. Interventions can certainly be targeted, but still should have in mind that the goal is to develop proficiency with reading comprehension and a variety of integrated math domains.

The ninth principle is to *extend duration and time on task*. Many interventions fail to last long enough or to provide enough time in instruction for individuals with LDs, especially if they demonstrate intractability. Note that the Tier 2 intervention in Mathes et al. (2005) was 40 minutes daily for 30 weeks and the Tier 3 interventions were 2 hours per day for 8–16 weeks in Denton et al. (2006b) and Torgesen et al. (2001). In subsequent attempts to reduce time and duration to, for example, 8–16 weeks for 30 minutes daily in grade 1, results like Mathes et al. (2005) have not been obtained (Denton et al., 2011). Unfortunately, as the example of Foorman et al. (1997) shows, supplanting instruction through some types of pull-out programs actually reduces the amount of instructional time. Tiers 2 and 3 must typically increase time on task and the duration of intervention. To accomplish this, the number of students who need intensive intervention for LDs needs to be reduced through strong Tier 1 instruction.

The tenth principle of effective intensive intervention instructional design is *ongoing progress monitoring* to ensure that students with LDs are responding to generally well-designed intervention. This includes the *use of progress-monitoring systems that help teachers know when to make an adjustment to the instructional program and how to generate ideas for productive adjustments,* as described in Chapter 4. The use of assessment systems for monitoring student response and for adjusting those programs is an essential element of effective intensive intervention

In Chapter 4, CBM progress monitoring commonly involves weekly to biweekly assessments, each time on a different material. In this situation, the variation in these individual passage estimates is overcome because no decision is based on a single score. Rather, performance is summarized across time, while relying on multiple readings on multiple passages. In the future, form-equated methods will emerge that will make progress monitoring easier to equate over differences in material and difficulty level. For instruction and in contrast to identification (see Chapter 4), data on both slopes and final status is essential. Such ongoing progress monitoring is required to provide constant feedback to teachers concerning program effectiveness for the individual student. With ongoing monitoring, a student who does not respond adequately can be identified promptly, and the teacher can immediately adjust the intervention to tailor it to the student's needs. After the instructional adjustment has been designed, the teacher implements the program change and continues to conduct ongoing progress monitoring. In this way, the teacher continues to evaluate the success of the intervention for the student with an LD and to make adjustments to the program whenever inadequate response occurs. In MTSS approaches, this distinguishes intensive intervention at Tier 3 from Tier 2, in which the validated program is implemented as designed. In essence, ongoing progress monitoring is used to determine whether a validated treatment program, when implemented intensively, is in fact effective for a given student with an LD.

When progress monitoring reveals that a student is failing to respond as expected, it is then used for a second purpose: to differentiate instruction for that student to represent an individually differentiated instructional program. Multiple randomized control trials show that when teachers use ongoing progress monitoring in this way, they plan instruction more effectively and produce stronger academic outcomes for students with LDs (Stecker et al., 2005). In fact, many consider the use of ongoing progress monitoring to inductively formulate instructional programs over time to be a signature feature of effective special education. For assessing response to ongoing intervention, L. S. Fuchs and Fuchs (1998) and Speece and Case (2001) reported that indices based on both slope and intercept were more predictive of long-term outcomes than slope or intercept alone, the rationale being that a student could begin an intervention well below benchmark standards, but have a very positive response that would be masked by the intercept or an end-of-year benchmark alone.

Although research supports the efficacy of a variety of instructional methods for promoting academic achievement among students with LDs (Swanson, Harris, & Graham, 2013), the heterogeneity of this population, combined with the severe and multifaceted nature of their needs, results in a high rate of inadequate responsiveness to validated interventions that is high, ranging between 10 and 50%, depending on the intervention and the criteria for "inadequate response." For this reason, academic outcomes for students with LDs can be enhanced substantially when student progress is systematically monitored while validated interventions are being implemented. With progress monitoring, teachers and others gauge the extent to which an individual student is responding to an instructional intervention. When response is inadequate, teachers can quickly revise the program and then monitor the impact of those revisions.

Most importantly, when CBM is used to determine the need for revisions to student programs, better end-of-year academic outcomes result than when CBM is not used. CBM enhances instructional planning and student learning by helping teachers set ambitious student goals, by assisting in determining when instructional adaptations are necessary to prompt better student growth, and by providing ideas for potentially effective teaching adjustments (L. S. Fuchs, D. Fuchs, & Hamlett, 1989).

Another key way in which CBM can enhance instructional decision making is in assessing the adequacy of student progress and determining whether instructional adaptation is necessary. When actual growth rate (slope of the observed line) is less than the expected growth rate (slope of the goal line), the teacher modifies the instructional program to promote stronger learning (L. S. Fuchs et al., 1989). Simply collecting CBM data exerts only a small effect on student learning. To significantly enhance student outcomes, teachers need to use the CBM data, almost like an ongoing experiment, to build effective programs.

For helping teachers determine when adjustments are required in students' programs and for identifying when goal increases are warranted, the

CBM total scores are used. In addition, by inspecting the graph of performance indicators over time, teachers may formulate ideas for potentially effective instructional adaptations. For example, a flat or decelerating slope might generate hypotheses about lack of maintenance of previously learned material or about motivational problems. Nevertheless, to obtain rich descriptions of student performance, alternative ways of summarizing and describing student performance are necessary. Because CBM assesses performance on the year's curriculum at each testing, rich descriptions of strengths and weaknesses in the curriculum can be generated.

Figure 5.6 offers an example of a CBM graph showing program development and progress for a child in third-grade mathematics. Each dot represents performance on one occasion on one alternate form of a CBM test that systematically sampled the third-grade curriculum. The vertical dotted line denotes the setting of the goal (also see G at year's end); the dotted vertical line indicates the rate of progress required to achieve the year-end goal; and the solid vertical lines show when the teacher revised the instructional program in an attempt to boost the rate of progress. The last set of data points reveals a stronger rate of growth (the four most recent scores are all above the goal line), so the decision was to increase the goal. The boxes at the bottom represent mastery of the skills taught in the third-grade curriculum. The first stack of boxes shows no mastery (i.e., no dark boxes); in mid-April, Stephen had mastered three skill areas (measurement, money, decimals); had probably mastered two additional skills (counting, applied computation); and had partially mastered four more skills (number concepts, names of numbers, charts/graphs, fractions), leaving only word problems as attempted but not mastered. Stephen's teacher could look across rows of the skills profile to see, for example, that applied computation had gone from (1) not attempted, to (2) attempted but not mastered, to (3) partially mastered, then back to (4) attempted but not mastered after the winter break, to (5) partially mastered again, to (6) probably mastered in March–April.

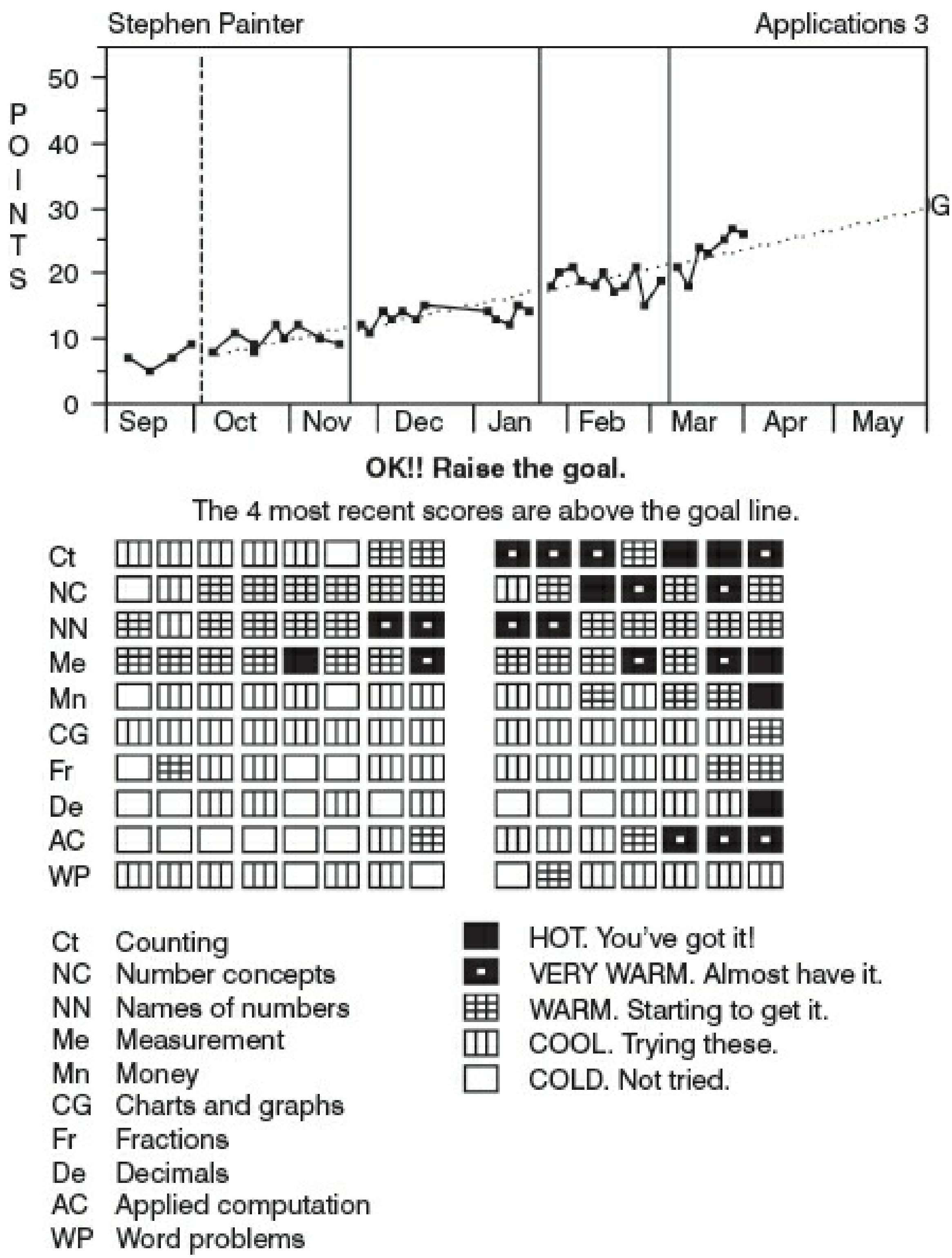

FIGURE 5.6. A CBM graph showing program development and progress for "Steven Painter" in a grade 3 mathematics concepts and applications curriculum. Each dot represents performance on one occasion on one alternate form of a CBM test that systematically sampled the grade 3 curriculum. The vertical dotted line denotes the setting of the goal (also see G at year's end); the dotted vertical line

indicates the rate of progress required to achieve the year-end goal; and the solid vertical lines show when the teacher revised the instructional program in an attempt to boost the rate of progress. The last set of data points reveals a stronger rate of growth (the four most recent scores are all above the goal line), so the decision was to increase the goal. The boxes at the bottom represent mastery of the skills taught in the grade 3 curriculum. The first stack of boxes shows no mastery (i.e., no dark boxes); in mid-April, the student had mastered three skill areas (measurement, money, decimals); had probably mastered two additional skills (counting, applied computation); and had partially mastered four more skills (number concepts, names of numbers, charts/graphs, fractions), leaving only word problems as attempted but not mastered. The student's teacher could look across rows of the skills profile to see, for example, that applied computation had gone from (1) not attempted, to (2) attempted but not mastered, to (3) partially mastered, then back to (4) attempted but not mastered after the winter break, to (5) partially mastered again, to (6) probably mastered in March–April.

CHARACTERISTICS OF INEFFECTIVE INSTRUCTION

There are several approaches to instruction for students with LDs that are demonstrably *in*effective. Some of these were reviewed by Pennington (2009, 2011). In addition, ineffective intervention has general characteristics that are summarized in Table 5.2. *The largest contributor to ineffectiveness is that the intervention does not occur in the context of reading, math, and written language.* Basically, it is very easy to eliminate many commonly proposed instructional approaches for students with LDs because they don't teach reading, math, or written language. Widely publicized approaches to assisting students with LDs in reading by slowing down the speed of temporal processing of words have been shown in multiple studies to lack generalization to improved reading (see review by Olson, 2011, and meta-analysis by Strong, Torgerson, Torgerson, & Hulme, 2011). Teaching working memory skills using computer programs enhances task-specific working memory, but shows little generalization to math or reading (Melby-Lervåg, Redick, & Hulme, 2016). This general principle has been observed for many years in efforts to train cognitive processes (Mann, 1979) and is simply ineffective if it occurs outside the context of reading, math, and written language (Kearns & Fuchs, 2013).

TABLE 5.2. Characteristics of Ineffective Interventions for LDs

1. Doesn't focus on academic skills.

2. Defines academic proficiency narrowly.
3. Doesn't increase instructional time, intensity, or differentiation.
4. Doesn't continually monitor progress and adjust instruction or change program.
5. Teaches for the sake of learning rules, not to master principles.
6. Doesn't engage the child in reading instructional-level material or practice in math and writing.
7. Waits for the child to fail; leaves the child behind.

What would happen if these methods were combined with academic instruction is an interesting question for which research will likely emerge. Importantly, this general principle can be expanded to exercise (Denton, 2011) and perceptual–motor training (Arter & Jenkins, 1979), optometric and related lower level oculomotor and visual efficiency training, and special lenses (American Academy of Pediatrics, 2009; Barrett, 2009; Fletcher & Currie, 2011), and other "shortcut" solutions to the difficulties presented by LDs. *If the intervention does not teach reading, math, and/or written language, do not expect improvement in academic skills.*

CONCLUSIONS: PRINCIPLES OF INTERVENTION

In this chapter, we addressed general features of instruction for students with LDs. We emphasized the need to provide intensive Tier 2 intervention in the context of strong core instruction (Tier 1), which usually reduces the number of students who require more specialized intensive intervention (Tier 3). We also emphasized the importance of early intervention, a topic that we will integrate with neurobiological research in subsequent chapters that highlight the importance of early, intensive intervention for reducing effects of genetic risk and developing the neural systems that mediate reading, math, and writing development. When remediation occurs, it must be much more intensive than in current practice, which we believe could be the mandate of special education intervention programs. To accomplish these goals, we highlighted MTSS frameworks for the service delivery context. Ideally, these frameworks integrate general and special education and provide for the continuum of services needed to improve academic and behavioral outcomes for all students. We highlighted general principles of effective and ineffective intervention for individuals with LDs. Ongoing assessment through progress monitoring is essential for all students who struggle.

CHAPTER 6

Word-Level Reading Disabilities (Dyslexia)

This chapter examines the scientific evidence bearing on the study of people with reading difficulties that involve single-word reading and spelling, or dyslexia. Consistent with the organization of subsequent chapters addressing LDs in other domains, we discuss word-level reading disabilities beginning with definition and classification, including epidemiology, sex ratio, and developmental course. We then examine academic skill deficits, core cognitive processes, and neurobiological factors (brains and genes). We conclude with a comprehensive review of instructional interventions and remediation efforts. This format varies slightly depending on the nature of the literature in each domain.

DEFINITION AND CLASSIFICATION

Word-level reading disability (WLRD) is synonymous with "dyslexia," which has been described throughout the 20th century as "word blindness," "visual agnosia for words," and "specific reading disability" (Doris, 1993). Early definitions of dyslexia were built in part on an older history of efforts to understand children with "unexpected" reading difficulties. Described initially as "word blindness" by ophthalmologists (e.g., Morgan, 1896), initial observations involved case examples of children who were unable to read despite clearly adequate intelligence, strengths in other domains, and an absence of brain injury. In his case description, Morgan (1896, p. 378) described a 14-year-old boy who "has always been a bright and intelligent boy, quick at games, and in no way inferior to others of his age." He then described his conspicuous problem reading and spelling words, concluding that "He seems to have no power of preserving and storing up the visual impression produced by words—hence the words, though seen, have no significance for him. His visual memory for words is defective or absent; which is equivalent to saying he is . . . "word blind."

The term "dyslexia" became prominent because of the work of Samuel Orton and his colleagues, who developed a theory of dyslexia and interventions (Orton, 1928). We use the terms WLRD and dyslexia interchangeably, and generally use the term "dyslexia" when the source we are describing used it. This usage is deliberate and designed to refer to people who display, as Morgan (1896) so aptly described, a primary and often profound problem in reading and spelling single words in isolation. Difficulties in reading single words co-occur frequently with limitations in vocabulary development and reading comprehension across multiple academic domains whenever text is used to convey information. However, single-word reading difficulties play a primary role because text-level cognitive processes demand accurate and fluent reading of the words in text. The inability to read and spell words is a major source of adaptive difficulty for persons with LDs and should never be minimized. People with these difficulties may learn compensation skills, but compensatory strategies are rarely sufficient for full proficiency with reading comprehension.

The evolution of dyslexia from a vague term to a more precise synonym for WLRD provides an example of how definitions of LDs can move from approaches based on exclusionary criteria that mostly indicate what LDs are not (Rutter, 1982) to inclusionary definitions that focus on a key set of marker variables that lead directly to identification. As an example of an exclusionary approach, consider the definition of dyslexia formulated by the World Federation of Neurology in 1968 as summarized in Critchley (1970): "a disorder manifested by difficulties in learning to read despite conventional instruction, adequate intelligence, and socio-economic opportunity. It is dependent upon fundamental cognitive disabilities, which are frequently of constitutional origin" (p. 11). This definition mirrors that used in the ICD-10 (World Health Organization, 2013) and the now discontinued DSM-IV definitions (American Psychiatric Association, 1994), which relied on IQ–achievement discrepancy formulae. As discussed in Chapter 2, DSM-5 (American Psychiatric Association, 2013a) explicitly rejected IQ–achievement discrepancy criteria, but recommended an inclusionary threshold for low IQ within two standard deviations of the mean to differentiate LD from an intellectual disability. DSM-5 identifies different types of LDs in reading (word-reading accuracy, reading fluency, and reading comprehension). This definition did not identify a category of "dyslexia," but noted that problems with the accuracy and fluency of single word-reading skills address dyslexia, much like IDEA 2004 addresses dyslexia by identifying a category of "basic reading skills."

For a more specific definition of dyslexia, consider the formulation developed in 1994 and revised to take advantage of the rapid progress in research that had occurred over the ensuing decade (Lyon, Shaywitz, & Shaywitz, 2003):

> Dyslexia is a specific learning disability that is neurobiological in origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction. Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede growth of vocabulary and background knowledge. (p. 1)

Figure 6.1 provides a schematic of the components of this definition. Building on the research on academic skill deficits and their cognitive correlates reviewed below, this definition indicates that dyslexia is manifested by difficulties with phonological language, often including, in addition to problems with word reading, a conspicuous problem with acquiring proficiency in spelling and writing. Although the definition emphasizes word-reading accuracy, it also explicitly notes that decoding fluency, or automaticity of word reading, is also involved. Given the difficulties in accurately and fluently reading words, reading comprehension can be affected because inaccurate and nonfluent word reading taxes working memory. The definition is inclusionary because it specifies that people can be identified with dyslexia when they show problems with decoding single words accurately and fluently and spell poorly.

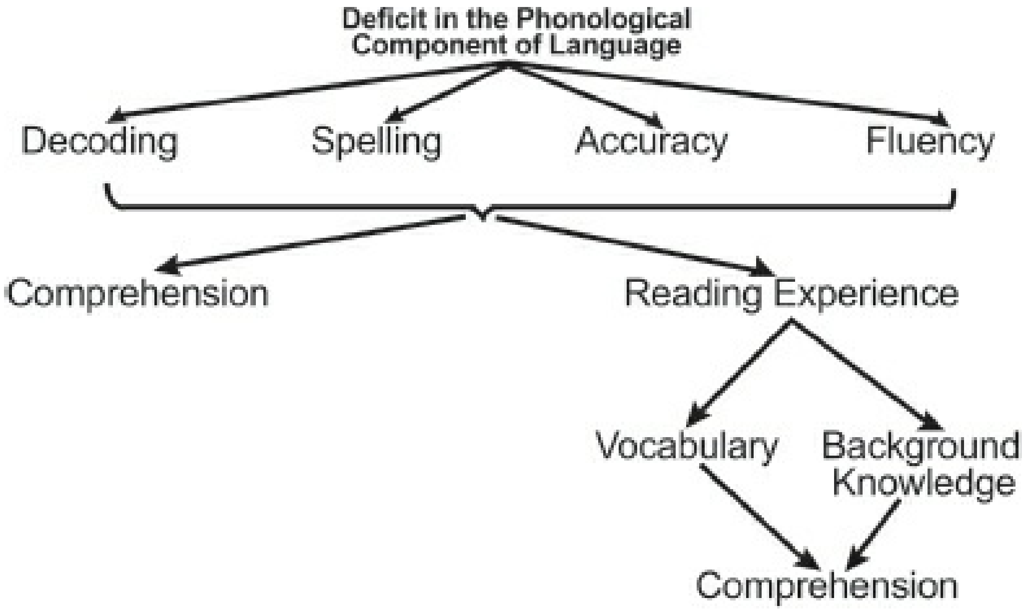

FIGURE 6.1. Components of the definition of dyslexia adopted by the International Dyslexia Association. Courtesy Emerson Dickman.

There remains consensus support for this definition, although we now know that dyslexia is more complicated than a problem with phonological processing, with strong evidence for multiple deficits that influence phonological processing (Pennington, 2006). Nonetheless, Dickman (2017) summarized his survey of over 30 international researchers on dyslexia: Thirty well-known researchers and practitioners took part in the discussion and found little support for changes in the definition.

Prevalence

The prevalence of dyslexia is commonly estimated at 3–7% when applying a cut point of 1.5 standard deviations below the mean on measures of reading achievement (Peterson & Pennington, 2012; Snowling & Melby-Lervåg, 2016). Historically, studies of reading disabilities have generated prevalence estimates of 5–15% in the school-age population (Rutter et al., 2004) but the prevalence of dyslexia has been estimated to be as high as 17.4% in the school-age population (S. E. Shaywitz, 2004). These higher prevalence rates can be misleading given the variations in criteria used to identify reading disabilities and evidence showing that the attributes of LDs, including low achievement, are normally distributed (S. E. Shaywitz et al., 1992). For example, for the 17.4% estimate, prevalence was based on a low achievement threshold at the 25th percentile and/or an IQ–achievement discrepancy regression-based definition of a 1.5 standard error difference between IQ and achievement. By adopting a cut point of the 25th percentile and adding children with reading scores that are above the low achievement threshold, but discrepant with IQ, the prevalence should be over 30%. It is lower because the epidemiological sample from which the estimates were derived has high average IQ and reading scores.

There is little evidence that rates of dyslexia vary significantly across non-English languages (Snowling & Melby-Lervåg, 2016). In one review of prevalence rates from mostly European samples, Moll, Kunze, Neuhoff, Bruder, and Shulte-Körne (2014) examined prevalence estimates of specific reading disability (RD) in isolation and comorbidly with specific arithmetic disability (AD) and spelling disability (SD). Depending on the definition and threshold used, there was a consistent tendency for specific RD (assumed to include SD) to be identified more frequently (range of 2.2–19.9% depending on cut points) than specific AD (range of 1.3–10.3%). However, comorbid associations of RD and AD ranged from 1.0 to 7.6%, and comorbid AD and SD from 2.3 to 8.1%. It is important to again underscore that prevalence rates vary as a function of the definition of RD. As Moll et al. noted: “The total number of children identified with RD, SD or AD simply reflects the cutoff criterion used to classify learning problems” (p. 2).

In an additional epidemiological study of 1,633 third- and fourth-grade students in Germany using DSM-5 inclusionary criteria and a one standard deviation cut point (i.e., < 16th percentile), Moll et al. reported prevalence figures of 6.5% for specific RD, 3.7% for RD and SD, 1.8% for RD and AD, and 3.5% for RD, SD, and AD. This totals to about 16% of the population with some form of RD that frequently co-occurs with SD and AD. Interestingly, isolated SD occurred in 6.7% of the population, while isolated RD occurred in 4.8% of the population; 2.8% met criteria for SD and AD.

In an alternative approach to the question of prevalence, Snowling and

Melby-Lervåg (2016) identified 15 independent longitudinal studies focused on samples of individuals at risk for reading failure because of a family history of dyslexia. Across these studies, they found that if the threshold for dyslexia was placed above the 10th percentile, the prevalence was 53%. If the threshold was placed below the 10th percentile, the prevalence was 34%. In samples with no family risk, prevalence rates were significantly reduced. The averaged prevalence across cut points was 12%. When the cut point was placed above the 10th percentile, 16% were identified as having dyslexia. When the threshold was established at below the 10th percentile, 8% of the samples were identified as having dyslexia. Clearly, prevalence is higher in samples with family risk, but this study also indicates how thresholds affect prevalence in these longitudinal studies.

Regardless of the prevalence, dyslexia is the most commonly *identified* form of LD. Lerner (1989) reported that 80–90% of all children served in special education programs had problems with reading. Kavale and Reese (1992) found that more than 90% of children in Iowa with the LD label were identified for reading difficulties. Both studies indicated that most children who have reading problems experience difficulty with word-level skills. Similarly, Leach, Scarborough, and Rescorla (2003) reported that about 80% of an elementary school sample selected because of reading problems had difficulties involving the accuracy of word reading. The remaining 20% had difficulties primarily in listening and reading comprehension. In middle school students who did not pass the Texas state accountability test of reading comprehension, Cirino et al. (2013) found approximately that over 85% had problems with decoding and/or fluency, while 12% had problems with comprehension based on a threshold of the 20th percentile. Thus, most children who are served in special education programs for LDs likely have WLRD as part of their disability. The rate of children with reading problems in U.S. and international surveys are often over 30%, reflecting the impact of SES and criterion-referenced definitions.

Sex Ratio

Dyslexia has frequently been considered to be more common in males than females. However, several studies have reported that the sex ratio between individuals with dyslexia is not significantly different (Flynn & Rahbar, 1994; S. E. Shaywitz et al., 1990; Wood & Felton, 1994), although these studies tended to report a slight male preponderance of about 1.4:1 (Flynn & Rahbar, 1994; S. E. Shaywitz et al., 1990). The conflict between the reported ratios may be related to the practice of sampling from clinic and school settings that were subject to referral bias. Specifically, boys are more likely to display externalizing behaviors that lead to referral, and the hyperactive–impulsive form of ADHD does appear to be more common in boys than girls (Barkley, 2015; S. E. Shaywitz et al., 1990).

More recent analyses question the lower prevalence rates and the ascertainment bias issue. Rutter et al. (2004) reanalyzed data from four independent epidemiological studies that permitted estimates of the sex ratio for reading disability. The authors reported that, across these studies, the sex ratio ranged from about 1.4–2.7:1, with males more frequently identified. They also included findings from additional studies in the United Kingdom and the United States that reported ratios of about 2:1 boys to girls. At the lower end, these ratios are not really different from those indicated in S. E. Shaywitz et al. (1990) and Flynn and Rahbar (1994), in which a ratio of about 1.4:1 was reported. In a large study of 491,103 beginning second graders, Quinn and Wagner (2015) evaluated sex ratios for measures of read-word and nonsense-word decoding fluency. There was clear evidence of increased prevalence in males from the 3rd to the 30th percentile, with greater severity associated with increased male preponderance. At the 3rd percentile, the ratio was 1.6:1 males for nonsense-word decoding and 2.4:1 males for real-word decoding fluency. By focusing on severity, Quinn and Wagner obtained rates comparable to Rutter and Yule (1975), which used a similar cut point. Most importantly, the differences were not attributable to ascertainment bias. *Only 1 in 4 boys and 1 in 7 girls with reading impairment were identified as learning disabled by the schools,* but the sex ratios were similar.

Altogether, these studies establish male preponderance in dyslexia, but not at the magnitude suggested by clinic samples (Peterson & Pennington, 2012). In some studies, ascertainment bias is clearly demonstrated (see Donovan & Cross, 2002), but more recent studies have shown clear associations with severity that are more consistent with male vulnerability than simply ascertainment bias. In a sense, reports of male preponderance may simply indicate that the distribution of reading skills is different in males and females. This begs the question of whether distributions should be pooled in estimating prevalence. Arnett et al. (2017) evaluated sex differences in a large sample of twins. Like other steadies, this study found lower reading scores in males than females, but differences in the sex ratio were found only in the lower levels of performance: 11.6% males versus 6.1% females. However, there was greater variability in male than female performance, essentially suggesting that the distributions were different. Equating the means and variances for these distributional differences reduced the sex difference to 8.3% males and 7.9% females. Because few studies find evidence of sex-based phenotypic differences in the expression of WLRD (Canning, Orr, & Rourke, 1980; Flynn & Rahbar, 1994; Jiménez et al., 2011), more research is needed on the basis for male preponderance. There is evidence for sex differences in brain structure and function (Lambe, 1999), specifically among individuals with dyslexia (Evans, Flowers, Napoliello, & Eden, 2014).

Developmental Course and Outcomes

Dyslexia in particular and RDs in general reflect persistent deficits rather than a developmental lag in linguistic and reading skills (Francis et al., 1996; S. E. Shaywitz et al., 1999). Longitudinal studies show that, of children identified as reading disabled in grade 3, more than 70% maintain this status through grade 12 (see Figure 3.5; S. E. Shaywitz, 2004). Studies of adults with WLRD find that the word-reading difficulties persist and that the core cognitive correlates in the domain of phonological processing also persist (Bruck, 1987; Cirino, Israelian, Morris, & Morris, 2005). Altogether, the persistence of WLRDs is more the rule than the exception and represents chronic problems for the student. These findings highlight the importance of conceptualizing identification practices within the context of RTI (see Chapter 3).

ACADEMIC SKILL DEFICITS

Word Reading

Using the framework introduced in Figure 1.1, the major academic skill deficits characterizing children with dyslexia is a difficulty with the accuracy and/or fluency (automaticity) of single-word decoding, which affects word reading in isolation and in text (Lyon et al., 2003; Stanovich, 1986). However, the extent to which deficits in word-reading accuracy and/or fluency occur varies across the language being read. This is an important issue, particularly when attempting to understand how linguistic demands in different orthographies contribute to different phenotypes of reading failure. To understand dyslexia, an international, cross-linguistic perspective is essential. Research on dyslexia has been overly influenced by studies of English speakers (Share, 2008), with phonological–orthographic relations in English representing an outlier relative to other languages and leading to excessive focus on reading accuracy (see Chapter 10).

Different orthographies reflect different levels of information depending on sound, pattern, and meaning. A shallow or transparent orthography is one that is highly regular in its sound–symbol correspondences. For example, when reading and spelling words in Finnish, Spanish, or Italian, it is relatively easier to decode written words because there is a more direct, or transparent, correspondence between sounds and letters. German written language reflects a less transparent orthography because it presents the reader with both direct and indirect sound–letter relationships. English and French written languages are characterized by more deep or opaque orthographies because the correspondence between letters and sounds is more irregular, with English having the most opaque orthography of the major languages. Chinese, which also has a deep orthography, includes characters that represent morphemes, but is not an alphabetic language.

An example of how orthographic depth influences different word-reading phenotypes can be observed in comparisons of studies of English-speaking children with WLRD and studies of non-English-speaking children reading a more transparent language. Among English-speaking children there is a relatively higher frequency of accuracy errors when reading and spelling words, in contrast to fluency difficulties observed among Spanish and Italian children with WLRD, who have less difficulty in accurately reading and spelling words. Despite variations in error patterns as a function of the depth of orthography, difficulties in phonologically processing are strongly related to WLRD in both more transparent and more opaque written languages

(Wimmer, 1993; Zeigler & Goswami, 2005).

Spelling

The other academic skill deficit usually characteristic of WLRD (dyslexia) is a spelling deficit. Not only is it difficult for individuals with WLRD to read (decode) words in isolation or in text, it is also difficult for them to spell (encode) words in isolation or in text. We return to the issue of spelling in Chapter 9 as part of our discussion of written expression. Spelling (like word reading) is a multidetermined skill requiring a number of cognitive processes to encode written text. That said, similar to WLRD, spelling difficulties are strongly related to deficits in phonological processing. The distinction between word reading and spelling is important because some individuals experience spelling difficulty with word recognition accuracy. To reiterate, reading fluency and spelling are especially apparent in the identification of LDs in people who use languages that have relatively more transparent orthographies, such as German (semitransparent) or Spanish (transparent) (Wimmer & Mayringer, 2002).

CORE COGNITIVE PROCESSES

While there is converging evidence documenting the fundamental academic skills deficits in WLRD, which represents the primary phenotype, some debate continues about nonreading factors (e.g., linguistic, perceptual) that account for limitations in single-word reading and represent strengths and weaknesses associated with the primary phenotypic expression. Two different perspectives continue to exist. The first and more influential proposes that deficits in word recognition are primarily associated with, or caused by, one primary nonreading factor. For example, both phonological awareness and rapid temporal processing have each been offered as single correlated or causal mechanisms. The second view is that deficits in the ability to read single words rapidly and automatically are attributable to multiple factors (e.g., phonological awareness, rapid naming, verbal short-term memory), appearing alone or in combination, thus giving rise to hypothesized subtypes of reading disabilities (Doehring, 1978; Pennington, 2006).

Phonological Awareness and Learning to Read

The predominant core cognitive correlate of WLRD involves phonological awareness, a *metacognitive* understanding that the words we hear and read share internal structures based on sounds (Liberman & Shankweiler, 1991; Share & Stanovich, 1995). Speech sounds, or phonemes, are the smallest parts of speech that make a difference in the meaning of a word. They are described by their phonetic properties, such as their manner or place of articulation, and their acoustic features or patterns of sound waves. English is an alphabetic language containing 44 phonemes. In any alphabetic language, the unit characters (letters) that children learn in order to read and spell words are keyed to the phonological structure of the words (Liberman & Shankweiler, 1991; Lukatela & Turvey, 1998).

A child's primary task in the early development of reading and spelling in an alphabetic language is to develop an understanding of the alphabetic principle—the realization that speech can be segmented into phonemes and that these phonemes are represented in printed forms (Liberman, 1971, 1996). However, developing this awareness that words can be divided into segments of sound is a very difficult task for many children. The sounds are "coarticulated" (overlapped with one another) to permit rapid communication of speech, rather than sound-by-sound pronunciation. This property of coarticulation—critical for speech, although making the job harder for the beginning reader and speller—is explained by Liberman and Shankweiler (1991):

> The advantageous result of co-articulation of speech sounds is that speech can precede at a satisfactory pace—at a pace indeed at which it can be understood. . . . Can you imagine trying to understand speech if it were spelled out to you letter by painful letter? So co-articulation is certainly advantageous for the perception of speech. But a further result of co-articulation, and a much less advantageous one for the would-be reader, is that there is, inevitably, no neat correspondence between the underlying phonological structure and the sound that comes to the ears. Thus, though the word "bag" . . . has three phonological units, and correspondingly three letters in print, it has only one pulse of sound. . . . Beginning readers can understand, and properly take advantage of, the fact that the printed word "bag" has three letters, only if they are aware that the spoken word "bag," with which they are already quite familiar, is divisible into three segments. They will probably not know that spontaneously, because as we

have said, there is only one segment of sound, not three, and because the processes of speech perception that recover the phonological structure are automatic and quite unconscious. (pp. 5–6)

The metacognitive awareness of the phonological structure of language is not the same as a problem processing the perceptual features of speech, the latter possibly representing one of several factors that can make it difficult for the child to develop this overarching understanding of the internal structure of speech. Regardless of the cause, deficient phonological awareness is a primary basis for the accurate recognition of words necessary for basic reading, reading comprehension, spelling, and written expression (Liberman & Shankweiler, 1991; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2002; Share & Stanovich, 1995). The child (or illiterate adult; Castro-Caldas, Petersson, Reis, Stone-Elander, & Ingvar, 1998) must map the phonemic structure of language onto its written form (orthography), which unlike language, is not a natural, evolutionary process. When the child begins to read, the brain changes in ways that permit mediation of word recognition, illustrating a fundamental form of neural plasticity (Dehaene, 2009; Wolf, 2007). As phonological awareness develops and the child understands the alphabetic principle, the words begin to have significance as visual forms with meaning. The process of word recognition is mastered early in the reading process as the brain develops the capacity to mediate this process, moving from a sublexical (phonological) to a lexical process in which the brain rapidly analyzes the orthographic features of words and immediately goes from the word to its meaning. Opportunities to read allow for self-teaching through practice and recoding of grapheme–phoneme correspondences, so the child begins to become self-taught (Share, 1995). However, the development of reading fluency (automaticity of word reading) and the application of word reading to comprehension processes have longer developmental trajectories.

When the child (or illiterate adult) does not understand the relations of sound and print, word recognition is problematic. The longer the child struggles to learn to read words, the more likely it is that a severe reading disability will emerge because, as we discuss below, the child cannot access print and gain enough exposures to orthographic patterns to develop the brain network needed to support reading. Developing automaticity and comprehension becomes increasingly difficult as the child's exposure to the range and frequency of sight words required to support proficient text reading decreases. Reading becomes a laborious and frustrating experience, thereby reducing the motivation to engage in wide reading for both academic and personal interests. Given the relation between wide reading and exposure to new vocabulary, understanding of word meanings is compromised, further interfering with comprehension (see Figure 6.1).

Rapid Automatized Naming

Rapid automatized naming (RAN) refers to the ability to quickly name letters, digits, and nonalphabetic stimuli. In assessing these skills, the tester asks the individual to name these stimuli as quickly as possible when they are presented in multiple rows. There remains significant debate as to whether RAN contributes unique variance to reading outcomes, and thus reflects a core process in word reading (Wolf & Bowers, 1999) or whether the influence is explained through a common phonological processing factor (Shankweiler & Crain, 1986). In addition, the relation of RAN to reading fluency is perhaps stronger, as discussed in Chapter 10.

Wolf and Bowers (1999) argued that RAN deficits are independent of phonological processing. However, while distinct, phonological processing and rapid naming are correlated, Schatschneider, Fletcher, Francis, Carlson, and Foorman (2004) found that kindergarten assessments of phonological awareness and rapid naming of letters were both predictive of word recognition skills at the end of first grade. In studies that model the growth of reading and reading-related skills over time, phonological awareness and rapid naming abilities uniquely predict English reading skills over time (Wagner, Torgesen, & Rashotte, 1994; Wagner, Torgesen, Rashotte, & Hecht, 1997). Wagner et al. (1997) suggested that, due to the high correlation of phonological awareness and rapid naming assessments at the latent variable level, both were determined by phonological processing. Such an interpretation is consistent with the phonological limitation hypothesis of dyslexia (Shankweiler & Crain, 1986).

In a meta-analysis of 137 studies, Araújo, Reis, Petersson, and Faísca (2015) found a correlation of 0.43 for the relation of RAN and reading performance. Although RAN contributed significantly to word reading, text reading accuracy and fluency, nonword-reading accuracy and fluency, and reading comprehension, the relations were much higher for word and text reading. In addition, the relations for accuracy measures were stronger for studies of lower grades and for RAN tasks involving alphanumeric stimuli. Individual differences found on fluency measures were consistently related to grade level. The meta-analysis showed different contributions of RAN tasks to orthographies that varied in the transparency of phonology and orthography. However, little evidence was obtained for an explanation linking RAN solely to an orthographic or phonological explanation, or for explanations based solely on speed of processing.

Even with meta-analysis, a correlational study cannot unravel different sources of variability, so the relation of rapid naming deficits and reading in individuals with dyslexia remains controversial. One review of the evidence concerning the relation of naming speed and dyslexia (Vukovic & Siegel, 2006) concluded that there was little evidence supporting rapid naming difficulties as an isolated deficit specific to individuals with WLRD, stating that "the existing evidence does not support a persistent core deficit in naming speed for readers with dyslexia" (p. 25). In contrast, a twin study by Petrill, Deater-Deckard, Thompson, DeThorne, and Schatschneider (2006) found that phonological awareness and rapid naming were moderately correlated, but factorially distinct at a latent variable level, and that both contributed uniquely to word recognition outcomes. Whereas phonological awareness had both genetic and shared environmental influences, the contribution of rapid naming was primarily genetic. They concluded that "serial naming speed is phenotypically separable from phonological awareness and could constitute a second, etiologically distinct source of variance in reading skills" (p. 120).

The specificity of RAN deficits to WLRD has also been questioned. Waber, Wolff, Forbes, and Weiler (2000) found that RAN was not specifically associated with reading difficulties because relations were found also with ADHD. Note that in Figure 2.5, a RAN composite did not differentiate those with reading and math LDs from each other. Many studies of general populations have shown that rapid naming of letters contributes independent variance to word reading even when phonological processing is controlled (Schatschneider et al., 2004). However, whether this relation holds *specifically* for people with dyslexia is unclear (Vukovic & Siegel, 2006).

Phonological (Working) Memory

The other cognitive process that is significantly related to word recognition skill and to dyslexia involves working memory for verbal and/or acoustic (sound-based) information. In Wagner et al. (1997) and Schatschneider et al. (2004), different phonological memory tests were not found to contribute unique variance once phonological processing was included in the model. However, there are many comparisons of verbal working memory tasks between people with dyslexia and those in a typically achieving control group, with working memory problems commonly observed among people with dyslexia (Siegel & Ryan, 1989). The question is whether the working memory problems are independent of phonological processing.

Melby-Lervåg, Lyster, and Hulme (2012) conducted a meta-analysis of 235 studies of the relation of phonological awareness and verbal short-term memory to word reading. Comparisons of people not impaired and those impaired in word reading showed larger effect size differences for phonological awareness (1.37) than for verbal short-term memory (0.71). In unselected samples, phonological awareness was the strongest unique predictor of word-reading skills even when controlling for verbal short-term memory. They argued in favor of the phonological representation hypothesis (Snowling & Hulme, 1994), which suggests that the development of word recognition skills depends on the availability of the capacity for phonemically structured representations of speech and that the failure to develop these representations causes inadequate development of word-reading skills (Hulme & Snowling, 2009; Shankweiler & Crain, 1986).

In contrast, in a selective meta-analysis of 48 studies, Kudo, Lussier, and Swanson (2015) found effect size differences for good and poor readers ranging from 1.00 for phonological awareness, 0.89 for rapid naming, 0.79 for verbal working memory, 0.56 for short-term memory, 0.48 for visual–spatial memory, and 0.67 for executive processes. Hierarchical modeling of the data in relation to reading skills found that verbal working memory, visual–spatial memory, executive processing, and short-term memory, along with IQ, were significant moderators of effect size differences, questioning the unique contributions of phonological processing. Unfortunately, this study had small samples of studies, did not separate different kinds of reading tasks, and identified few studies from non-English-speaking countries.

In a study of the predictive validity of letter–sound knowledge, phonological awareness, RAN objects and colors, and verbal working memory in English and three more transparent languages (Spanish, Czech, and Slovak), Caravolos et al. (2012) found that the first three variables were all similarly predictive of reading skills before and after the onset of formal literacy instruction. Verbal working memory and IQ were not predictive. The extent to which working memory accounts for variability independent of phonological awareness remains unclear, and working memory does not account for all the variability in phonological awareness tasks (Oakhill & Kyle, 2000).

Other Unitary Processes

There has been no shortage of other hypotheses about causal mechanisms in WLRD. We review these hypotheses below, with a focus on evidentiary support, neurobiological correlates, and the effectiveness of related intervention programs. In contrast to the substantial data on phonological processing, these hypotheses are generally not viewed as strong explanations of WLRD, much less the development of reading proficiency. In many instances task deficits may be a consequence of reading difficulties related to stress or lack of reading experience.

Visual Modality

The history of behavioral research on children with reading disabilities is characterized by various attempts to compete and compare single causal factors such as visual perceptual skills among good and poor readers. These studies invariably beg the question of how the presence of a single factor in children with reading difficulties explains the word-reading problem. Visual modality hypotheses are classic examples of generalizing from a group difference or correlation to a cause.

Visual–Perceptual Deficit Hypotheses

Much of the literature concerned with dyslexia in the 1960s and 1970s evaluated the hypothesis that *visual–perceptual or spatial cognition difficulties* were linked to reading disabilities (Vellutino, 1979). Although it is common to observe the presence of difficulties with copying or matching geometric designs in comparisons of children who are disabled and nondisabled in reading, there is little evidence that the spatial processing problems per se are linked to reading disorders (Vellutino, Fletcher, Scanlon, & Snowling, 2004). There is no doubt that individuals with reading disabilities *do* have problems that extend beyond the reading process. For example, difficulties involving math or attention, or other cognitive and motor difficulties do co-occur among children with WLRD. However, such comorbidity does not, or should not, imply that they play a causal role in the reading deficit. Many studies making such causal claims were, in fact, correlational in nature, rather than experimental.

Magnocellular Hypotheses

This same inferential problem is apparent in more contemporary studies that have attempted to establish *low-level sensory and attention deficits* in the *visual* modality as a cause of dyslexia. In the visual area, there are studies using psychophysical methods involving visual persistence, contrast and flicker sensitivity, and the detection of motion thresholds. Data from these studies are often interpreted to suggest a deficiency in the temporal processing of visual information, which interferes with the acquisition of reading skills (Stein, 2001, 2014). More recent studies evoke the construct of *visual attention* as a more general property of the visual system mediating early visual processing of different kinds of stimuli (Facoetti, Corradi, Ruffino, Gori, & Zorzi, 2010; Schneps, Brockmole, Sonnert, & Pomplun, 2012; Vidyasgar & Pammer, 2010). These studies generally ascribe these problems to specific difficulties in the magnocellular visual pathway, which is viewed as disrupting the processing of letter positions in the neural system (see below) that the brain uses to stabilize the visual appearance of external stimuli. The hypothesized result of these difficulties is that the text has a crowded or jittery appearance, making it difficult for the reader to apprehend the print in an accurate and fluent manner.

The magnocellular system is a transient visual channel that provides short, previsual responses to fast-moving stimuli that are low in spatial frequency. It supports peripheral vision by detecting motion to the side of the visual field. To identify *what* is moving in peripheral vision, central vision is needed, which is supported by the parvocellular visual pathway that operates in ventral brain networks. Conversely, the parvcelluar pathway is a sustained visual channel that provides longer duration responses to slow-moving stimuli of high spatial frequency. Lovegrove, Martin, and Slaghuis (1986) and Stein (2001, 2014) proposed that in children with WLRD, these two systems can inhibit one another, giving rise to ineffective transient system inhibition which interferes with the saccadic suppression of visual information, so that the word cannot be adequately fixated.

Similar effects have been hypothesized to emerge from a visual attention deficit that interferes with the processing of letter strings or the order of letters necessary for word recognition (Facocetti et al., 2010; Vidyasgar & Pammer, 2010). Gori and Facocetti (2015) argued that their data indicated that limitations in visual attention were the primary core deficit in reading failure with less evidence supporting the core importance of phonological awareness. Given their findings, Gori and Facocetti argued that traditional approaches to reading instruction should be replaced by interventions based on remediation of visual attention deficits.

Although individuals with reading disabilities often differ from typically achieving individuals on measures involving the visual system, including different measures of visual attention, it is not clear how the magnocellular system can be involved in word recognition, nor is the evidence consistent vis-à-vis its role. In a review, Boden and Giaschi (2007) identified different ways magnocellular dysfunctions might disrupt reading: (1) contrast sensitivity, (2) position encoding of letters in a word, (3) oculomotor deficits that affect eye movement control of saccades or binocular control, and (4) foveal/parafoveal interactions that lead to sluggishness in temporal processing. However, the authors found mixed evidence for each of these manifestations and concluded that the research base did not establish a role for the magnocellular system in reading, much less reading disability. Skottun and Skoyles (2008) reviewed studies of visual persistence, coherent motion, temporal order judgments, contrast sensitivity, and temporal acuity, concluding that “as far as vision is concerned there is little evidence for a specifically temporal deficit” (p. 666). Like Skottun and Skoyles (2008), Ramus (2003) indicated that lower-level visual-processing deficits were associated with dyslexia, but occurred at low rates and had limited capacity as an explanation of the reading problem, echoing other reviews that also identify problems with selection criteria in many studies (Goswami, 2015; Hulme, 1988).

The role for a visual attention factor is more compelling (Besner et al., 2016), especially when evidence from neuroimaging studies is reviewed, but clearly does not replace the primary role of phonological processing. The visual attention hypothesis needs to be uncoupled from the magnocellular hypothesis and the focus should be on whether visual attention is an additive factor in both good and poor reading. It is not yet known whether the visual attention factors emanating from attention circuits involving involuntary posterior brain systems are also related to more regulatory anterior systems, or if both systems are involved.

Peripheral Vision Hypotheses

Similar interpretative problems have been observed in a variety of efforts to link problems with peripheral vision to dyslexia. These areas are considered difficulties with “visual efficiency” and are usually treated by behavioral optometrists per the American Optometric Association practice guide (Garzia et al., 2008). Visual efficiency includes acuity and refraction, which ensures that objects in the visual fields are sharp and clear, essential for reading. However, ocular motor (eye movements) and accommodative-vergence functions are also evaluated and treated through optometric training interventions focused on control of eye movements. Rayner, Pollatsek, and Bilsky (1995), among others (Kirkby, Webster, Blythe, & Liversedge, 2008), have found that the eye movement difficulties sometimes seen in children with dyslexia are the product of their proficiency in reading as opposed to a cause of their reading problems. In evaluating the binocular control literature, Kirkby et al. concluded that “from the studies reviewed here, it should be clear that results in this area are highly contradictory” (p. 757).

Lower-level vision systems interact with higher cortical systems that process stimulus location or guide motor movements, such as eye movements in reading. Where we look next in a line of text is determined in large part by how well we understand what we have just read. If the word we just read doesn't make sense, our eyes move back to read it again. This has been clearly demonstrated in semitransparent languages like German (Hawelka, Gagl, & Wimmer, 2010).

Scotopic Sensitivity Syndrome

One other visual hypothesis based on the magnocellular system has been labeled as "scotopic sensitivity syndrome" (Irlen, 1994). The proposed intervention for this syndrome requires colored lenses and overlays to improve the efficiency of visual processing of text. Like other visual efficiency hypotheses, scotopic sensitivity (or Irlen) syndrome affects children with and without reading disabilities. As in other hypotheses loosely based on the magnocellular system, this syndrome is posited to make reading difficult because it degrades or jumbles the perception of print. Colored filters, either lenses or overlays, are believed to improve the functioning of the magnocellular system and prevent letters and words from seeming to "jump around the page." It is noteworthy, and somewhat concerning, that over half the population of the United Kingdom population has been reported to be affected by scotopic sensitivity syndrome (Evans & Joseph, 2002). The efficacy of colored overlays or lenses lacks evidentiary support (Kriss & Evans, 2005; Solan & Richman, 1990). Several studies found that colored lenses and filters improved reading *speed* slightly in all people regardless of reading status, but with no effect on accuracy, the primary difficulty in WLRD (Iovino, Fletcher, Breitmeyer, & Foorman, 1999; Kriss & Evans, 2005). The visual deficits are likely products of the reading problem itself. In an fMRI study controlling for reading experience, Olulade, Napoliello, and Eden (2013) concluded that "visual magnocellular dysfunction is not causal to dyslexia but instead may be consequential to impoverished reading" (p. 180). Thus, it is not surprising that groups of professional associations including the American Academy of Ophthalmology, the American Association for Pediatric Ophthalmology and Strabismus, the Council on Children with Disabilities, the American Association of Certified Orthoptists, and the United Kingdom College of Optometrists (Barrett, 2009) do not recommend these interventions

(American Academy of Pediatrics, 2009).

Orthographic Processing

Other attempts to explain the visual processing difficulties observed in dyslexia relate these difficulties to the processing of the orthographic components of written language and argue that such deficits are not related to phonological decoding. As noted earlier, the relation of phonology and orthography in English is sometimes inconsistent and English spellings are commonly irregular (Joshi, Treiman, Carreker, & Moats, 2008–2009; Rayner et al., 2002; Ziegler & Goswami, 2005). Thus, it is hypothesized that visual system deficits are related to the ability to immediately process words that cannot be sounded out automatically. This would imply differences in the ability of people with dyslexia to read pseudowords, exception words, and regular words, which has been infrequently reported. People with dyslexia simply can't read well!

Talcott et al. (2000) found correlations between visual motion sensitivity and orthographic processing even when variance due to phonological processing and IQ was covaried from the relationship. However, this relation was true for all children, regardless of the presence of a disability. In addition, there was no evidence that the relation of orthographic processing to word recognition was stronger than the relation of phonological processing. In German, which has a more regular relation of orthography and phonology than English, Wimmer and Schurz (2010) found little support for hypotheses involving visual attention, magnocellular processing, orthographic automatization, or visual–sequential memory. Only phonological processing in connection with orthographic–phonological connectivity was supported.

Auditory Modality

Rapid Temporal Processing

Hypotheses arguing a specific causal role for auditory processing have also been put forward. The most prominent example is studies carried out by Tallal and colleagues in a series of studies involving children with specific language impairment (Tallal, 1980, 2004). Differences between language-impaired and nonimpaired children were found in the ability to access acoustic stimuli with spectral parameters that changed rapidly in intensity. Problems in processing rapidly changing stimuli were observed for speech and nonspeech stimuli, leading Tallal and associates to hypothesize that both language and reading disabilities are caused by lower-level auditory processing problems involving the perception of rapidly changing stimuli.

Despite some apparent replications (Reed, 1989), Mody, Studdert-Kennedy, and Brady (1997) raised questions about the criteria for identifying poor readers and the comorbidity of ADHD within the sample that were not taken into account in the analysis of performance differences. Subsequent studies by other investigators have controlled for ADHD and used well-established definitions of dyslexia. Waber et al. (2001) identified children with dyslexia but not ADHD from a larger group of children originally referred for evaluation of learning impairments in a clinic setting. Waber et al. (2001) found a significant difference between good and poor readers in their ability to discriminate speech and nonspeech stimuli but not stimuli that showed rapid changes in their acoustic parameters. Breier, Fletcher, Foorman, and Gray (2002) used temporal-order judgment and discrimination tasks in large samples of children with dyslexia without ADHD, dyslexia with ADHD, ADHD without dyslexia, and typically achieving children. Children with dyslexia did not show a specific sensitivity to variations in interstimulus intervals. The results were independent of the presence of ADHD. The lack of consistent evidence demonstrating a link between the rapid processing of speech and nonspeech stimuli can be one reason that intervention programs that present auditory stimuli in a slower manner (e.g., Fast ForWord) (Scientific Learning Corporation, 1998) have not been found to be effective (Gillam et al., 2008; Olson, 2011).

Speech Perception

Alternatives to the temporal processing hypothesis focus more narrowly on speech perception. In contrast to other hypotheses reviewed in this section, studies of speech perception have been undertaken to explain why children present with phonological processing difficulties. Goswami (2011; Goswami et al., 2002) proposed that difficulties in the early specification and neural representation of speech could be a candidate core deficit in children with dyslexia. She and her colleagues suggested that deficits in speech perception could be related to early perceptual difficulties with rhythmic timing, which is important in the development of the ability to discriminate syllables and acoustic structures within speech. Using stimuli to examine the modulation of acoustic structures in segments of speech, Goswami found that children with dyslexia, in contrast to good readers, scored significantly lower on these tasks, with the level of difficulty associated with specific levels of reading and spelling proficiency. Goswami hypothesized that difficulties modulating acoustic structures were most likely referable to difficulties acquiring phonological awareness, thus suggesting an overarching role of the metacognitive understanding that speech has a phonological structure rather than a perceptual deficit per se. In subsequent studies, Goswami and colleagues have expanded this hypothesis to younger children and shown that early speech perception was related to preliteracy skills, such as rhyme awareness (Corriveau, Goswami, & Thompson, 2010).

Goswami (2014) interpreted these findings as indicating that risk for dyslexia may originate in the primary auditory cortex. Kraus and White-Schwoch (2015) also reported relations of early auditory processing and emerging phonological skills that they suggested may represent biomarkers for the emergence of reading impairment (White-Schwoch et al., 2015). Breier et al. (2002) found that children with WLRD exhibited significant problems with speech perception in a sample that excluded children with indications of an oral language disorder and ADHD. Breier, Fletcher, Denton, and Gray (2004) demonstrated that problems with perception of speech sounds characterized kindergarten students at risk for reading difficulties. In an imaging study involving the discrimination of speech sounds, Breier et al. (2003) found that children with dyslexia showed weak activation of temporo–parietal areas of the brain in the left hemisphere that corresponded to areas involving phonological processing. However, not all research finds that these types of speech perception difficulties are related to poor reading. Joanisse, Manis, Keating, and Seidenberg (2000) reported that speech perception deficits characterized only children who were identified with WLRD in the context of an oral language disorder. Likewise, children identified with speech–sound (articulation) disorders are at risk for reading difficulties (see Pennington, 2009), but largely when the articulation problems occurred in conjunction with an oral language disorder (Hayiou-Thomas, Carroll, Leavett, Hulme, & Snowling, 2017). While potentially a cause of phonological awareness and word-reading difficulties, more research needs to be completed.

Anchoring Hypothesis

An interesting body of research has been conducted primarily with Hebrew-speaking adults, but with some studies of children. Studies involve using an auditory task in which people with dyslexia and controls are asked to decide which of two tones has a higher pitch. In one condition the same reference tone was always presented first or second and the other tone was higher. In the second condition, no reference tone was presented. People with dyslexia were less able to learn that the reference tone was the same and use this information to perform the judgment. Thus, they benefited less from repetition, leading the researchers to conclude that they were less able to form or "anchor" perceptual representations with repetition (Ahissar, 2007). In subsequent studies, the researchers have been able to relate this difficulty benefiting from repeated exposure to working memory, phonological processing, and visual tasks involving words and visual features. Thus, it is proposed that anchoring difficulties represent a domain-general difficulty in forming a reliable perceptual representation of a repeated stimulus.

Although a domain-general anchoring deficit may not replace phonological processing as a pivotal component of learning to read, it could help explain difficulties developing perceptual representations of the orthographic features of print, thus interfering with the development of automaticity. Perrachione et al. (2016) suggested that anchoring deficits may be related to more general problems with neural adaptation, a form of plasticity that may be especially reduced in the neural network associated with immediate linkages of orthographic representations of whole words to meaning. We will return to this concept in our discussion of neural adaptation in Chapter 10 and its links to automaticity.

Exactly how anchoring deficits relate to reading requires additional study. Jaffe-Dax, Lieder, Biron, and Ahissar (2016) found that utilization of information about the statistical probabilities of occurrence in a serial frequency discrimination task was reduced in dyslexia. Direct analogues with orthographic representations have not been adequately studied. Using reading tasks usually demonstrates little more than the fact that people with dyslexia do not read as well as controls, and correlational findings in small samples are not persuasive. Di Filippo, Zoccolotti, and Ziegler (2008) reported that children with dyslexia had poorer performance on RAN tasks on repeated and unrepeated stimuli. However, Ahissar and Oganian (2008) suggested that this study was confounded because training trials were repeated prior to beginning the task, which would eliminate observations of anchoring. Beattie, Lu, and Manis (2011) found that adults with dyslexia were similar to controls on letter detection tasks, but had higher thresholds in the presence of external noise than in the absence of external noise. The authors concluded that noise, not an anchoring deficit, is the basis for the perceptual processing problems. In a study of good and poor German readers, Willburger and Landerl (2010) found that only poor readers with attention problems exhibited difficulties with the tone-length anchoring task, suggesting that comorbidity may be a problem for specific ties to dyslexia. More research on children and in samples in other languages is needed.

Cerebellar Hypothesis

Nicolson, Fawcett, and Dean (2001) proposed a cerebellar deficit hypothesis, suggesting that children with dyslexia represented a group that has failed to adequately automatize various skills, a function they argue is mediated by the cerebellum. Their research focused on phonological awareness as an example of procedural, or implicit, learning mediated by the cerebellum. This cerebellar hypothesis has given rise to interventions that specifically attempt to remediate reading deficits by focusing on the motor system.

There is limited support for this controversial theory (Nicolson & Fawcett, 2011), but many null results. Wimmer, Mayringer, and Raberger (1999) did not find that German children with dyslexia differed from controls on a balancing task, provided that ADHD was controlled for in the analysis. In fact, ADHD was a better predictor of performance on this cerebellar task than reading status. In a subsequent study, Raberger and Wimmer (2003) replicated these findings and were also unable to identify a link between balancing and rapid naming ability. Ramus, Pidgeon, and Frith (2003) found no evidence for time estimation deficits in individuals with dyslexia, and no evidence for causal relations of motor function and different phonological and reading skills. In a comparison of three hypotheses about dyslexia involving (1) phonological processing, (2) low-level auditory and visual deficits, and (3) cerebellar functions, Ramus (2003) found the strongest support for phonological deficits, which often occurred in the absence of any sensory or motor disorder. He noted that sensory and motor disorders occur in certain individuals with dyslexia, but was not able to link these with a reading problem. Similarly, Savage et al. (2005) found that measures of motor balance (and speech perception) did not contribute unique variance to reading and spelling outcomes if phonological processing was in the regression model.

Efforts supposedly demonstrating that training the cerebellum through physical exercise improves motor functions and reading (Reynolds, Nicolson, & Hambly, 2003) have proven highly controversial because, among other issues, the children with dyslexia were defined in part by performance on these tasks (Bishop, 2007). The same cerebellar tasks have not been strongly related to reading or to reading difficulties (Barth et al., 2010) and may reflect nonspecific associations with other disorders (Rochelle & Talcott, 2006; Loras, Sigmundsson, Stensdotter, & Talcott, 2014).

At the level of the brain, Kibby, Francher, Markanen, Lewandowski, and Hynd (2003) administered tests of reading and spelling, along with assessments of language functions. They also obtained MRI scans and measured the volume of the cerebellum. Although there were small but significant differences in cerebellum volumes between dyslexic and typically achieving children, which has been reported for different cerebellar structures in several studies (see Eckert et al., 2003), there was no evidence that cerebellum volumes correlated with academic or language skills in either group. However, in a study with careful specification of reading impairment in the participants, Fernandez, Stuebing, Juranek, and Fletcher (2013) found no differences in gray and white matter volumes of the cerebellum in children with dyslexia, but did observe regional reductions in the volume in the anterior lobe of the cerebellum relative to typically developing children. In a follow-up, Fernandez et al. (2016) traced the connectivity of this cerebellar region to cortical areas involved in reading. They found *greater* fractional anisotropy (FA) for children with dyslexia in tracts connecting the cerebellum with temporal–parietal and inferior frontal regions related to phonological processing relative to typical readers, implying reduced axonal integrity in these tracts.

A recent meta-analysis (Eckert, Berninger, Vaden, Gebregziabher, and Tsu, 2016) found weak support for increased cerebellar gray matter in people with dyslexia despite fairly consistent reports from different studies. This pattern was attributed to greater gray matter variability in the cerebellum in people with dyslexia. Danelli et al. (2013) found no overlap in neural activation during word and pseudoword reading for tasks related to cerebellum or magnocellular systems. However, Norton et al. (2014; see Chapter 10) compared brain activation during a printed-word rhyme-judgment task in children with phonological awareness, rapid naming, and both phonological awareness and rapid naming deficits. The researchers found that the naming deficit group showed reduced activation in right cerebellar lobule VI, which was even more reduced in the double deficit group. It would make more sense for cerebellar hypotheses to focus on indices of automaticity, such as rapid naming and reading fluency, since the brain regions mediating single-word reading and phonological awareness are well understood (see below).

Subtypes of Dyslexia

Empirical Subtyping

In an effort to explain the variability in LDs, it has been commonly hypothesized that a number of *subtypes* exist that can be identified on the basis of how people perform on measures of cognitive–linguistic, perceptual span, and other skills (see reviews by Hooper & Willis, 1989; Rourke, 1985). The argument for the existence of subtypes in the population with LDs was based on the practical observation that even though children with LDs may appear similar with respect to their reading deficits (i.e., word recognition deficits), they frequently differ significantly in the development of other skills that may be correlated with basic reading development. Thus, even within well-defined samples of children with dyslexia, there is large within-sample variance on some skills. This observation may explain, in part, why such children have been reported to differ from controls on so many variables unrelated to reading (Doehring, 1978). In the main, however, empirical subtyping studies, popular in the 1970s and 1980s, were confounded by measurement and methodological factors, with limited exceptions (see Morris et al., 1998).

Surface versus Phonological Dyslexia

Another prominent subtyping approach is derived from the dual-route framework of reading (Coltheart, 2005) and is based on a distinction between surface and phonological dyslexia in the acquired alexia literature (Castles & Coltheart, 1993). The dual-route theory stipulates that the reading system comprises a dorsal sublexical system in which phonological rules relate graphemes to phonemes and a ventral visual–orthographic system in which meaning is directly accessed (see below for a more detailed explanation). If the impairment is primarily in the sublexical system, the problem is considered *phonological dyslexia* and represents the common view of WLRD as a disorder caused by impairments in phonological processing. If the lexical system is the primary locus of impairment, the disorder is termed *surface dyslexia* and represents a problem that will be manifested with the orthographic component level of reading. The model thus predicts that people with phonological dyslexia are expected to exhibit poorer reading of pseudowords than exception words. In contrast, people with surface dyslexia are expected to exhibit better reading of pseudowords than exception words.

Findings related to this subtyping hypothesis question whether children with reading problems can be reliably characterized with surface dyslexia. Although a study by Murphy and Pollatsek (1994) reported no evidence for a subtype of surface dyslexia, Manis, Seidenberg, Doi, McBride-Chang, and Peterson (1996) and Stanovich, Siegel, and Gottardo (1997) did find some support for surface dyslexia in younger children. In Manis et al. (1996) the group of children with dyslexia had difficulties in reading both exception words and pseudowords, and the group identified with surface dyslexia performed similarly to controls matched on reading level. This observation was supported by Griffiths and Snowling (2002), who found that measures of phonological processing contributed unique variance to pseudoword reading, including phonological awareness and verbal short-term memory skills. The only unique predictor of exception-word reading was an assessment of reading experience, consistent with the view that orthographic processing had a significant experiential component. In another study, Stanovich et al. (1997) found that most children with WLRD experienced problems with both phonological and orthographic components of word recognition, suggesting that surface dyslexia represented an unstable subtype with a transient delay in the development of word recognition skills. In contrast, phonological dyslexia represented a long-term deficit in the acquisition of word-reading skills. Ziegler et al. (2008) also found multiple phonological, phonemic, and letter processing deficits in comparisons of children with and without dyslexia, suggesting that no single deficit could account for all people with dyslexia. Bergmann and Wimmer (2008) were not able to distinguish surface and phonological dyslexia in German-speaking children with dyslexia even when using speeded tasks because accuracy is not usually a problem in a midtransparent language like German. Sprenger-Charolles, Siegel, Jimenez, and Ziegler (2011) looked for evidence of surface and phonological dyslexia in three languages varying in transparency (English, French, and Spanish). Like Stanovich et al. (1997), they concluded that surface dyslexia had a “delayed developmental trajectory,” while phonological dyslexia was a “deviant developmental trajectory” (i.e., only phonological dyslexia was persistent). Surface dyslexia may appear transiently in younger children. The value of this subtyping hypothesis is its reliance on a theory of word recognition; the evidence for surface, or orthographic subtypes, is weak.

Accuracy versus Fluency

There are subtyping studies that (1) separate poor readers who are inaccurate word readers from those whose problem is with the automaticity of word reading or the fluency with text reading, or (2) separate poor readers according to patterns of impairment on assessments of phonological awareness and rapid naming, which are essentially proxies for word-reading accuracy and text-reading fluency. In these hypothesized subtypes, there is clear acceptance of a subtype represented as a WLRD with phonological processing problems. The critical question is whether a rate subtype can be identified as a subtype of WLRD or as a separate subgroup of reading LDs.

Rate versus Accuracy

Lovett (1987) proposed two subtypes of reading disability, based on the hypothesis that word recognition develops in three successive phases. The three phases are related to response accuracy in identifying printed words, automatic recognition without the need to "sound out" words, followed by developmentally appropriate maximum speed as components of the reading process become consolidated in memory. Children who fail at the first phase are "accuracy-disabled"; those who achieve age-appropriate word recognition but are markedly deficient in the second or third phase are "rate-disabled." In her subtype hypothesis, the rate-deficit group is an outgrowth of earlier problems with accuracy.

The strength of the Lovett subtype research program is its extensive external validation. In a study of the two subtypes (rate-disabled vs. accuracy-disabled) and a normal sample matched on word recognition ability to the rate-disabled group, children in the accuracy-disabled group were deficient in a wide array of oral and written language areas external to the specific reading behaviors used to identify subtype members; the rate-disabled group's deficiencies were more restricted to deficient connected-text reading and spelling (Lovett, 1987; Lovett, Ransby, Hardwick, & Johns, 1989). Reading comprehension was impaired on all measures for the accuracy-disabled group and was highly correlated with word recognition skill, but the rate-disabled group was impaired on only some comprehension measures. Additional subtype–treatment intervention studies find differences between the accuracy- and rate-disabled groups on contextual reading, whereas word recognition improved for both groups.

Double-Deficit Hypothesis

Other research emphasizes the importance of the basic distinction between accuracy and rate, but uses cognitive proxies for this relation. However, the rate subtype is not really considered as an outgrowth of the accuracy subtype, but as an independent group. Wolf and associates (Wolf & Bowers, 1999; Wolf et al., 2003) proposed that although phonological processing contributes considerably to word recognition deficits, accurate and fluent reading of text is also a critical academic skill. Children may demonstrate fluency deficits that are somewhat independent of problems with phonological processing. When isolated deficits in fluency occur, the most reliable correlate occurs on tasks that require rapid naming of letters and digits. Thus Wolf and associates have postulated a "double-deficit model" involving three subtypes: one characterized by deficits in both phonological processing and rapid naming; another with impairments only in phonological processing; and a third with impairments only in rapid naming.

Three lines of evidence support the validity of the double-deficit hypothesis. First, subtyping studies have compared children who have deficits in both phonological awareness and rapid naming with children who have only a single deficit (Wolf & Bowers, 1999). These studies show that children with double deficits have more severe reading difficulties than children who have only single deficits. The naming-speed group, unlike the double-deficit or phonological deficit group, does not appear to be significantly impaired in phonological processing or decoding (Wimmer & Mayringer, 2002). However, some investigators do not identify all the subtypes predicted by the double-deficit hypothesis (Waber, Forbes, Wolff, & Weiler, 2004). Second, as we discussed above, naming-speed tasks, especially the ability to name letters rapidly, consistently contribute independently to variance in reading achievement beyond what can be attributed to phonological awareness ability. Finally, the cluster analysis study of Morris et al. (1998) found evidence for a reliable subtype with impairment in both phonological awareness and naming speed, as well as subtypes with impairment in only phonological awareness or speed of processing.

As a subtyping hypothesis, there are several issues posed by the double-deficit framework (Vellutino et al., 2004). The most significant is whether the phonological awareness and rapid-naming deficits are really independent within the double-deficit group. It may be that within this group both deficits are driven by the severe problem with phonological processing that characterizes this group (Compton, DeFries, & Olson, 2001; Schatschneider, Carlson, Francis, Foorman, & Fletcher, 2002). There are also inherent methodological problems identified by Schatschneider et al. (2002) and Compton et al. (2001) that involve difficulties in defining single- versus double-deficit typologies. When both phonological processing and rapid naming are impaired, a child is more severely impaired in both dimensions, which makes it difficult to match single- and double-deficit-impaired children. Finally, it is unclear what is captured by the double-deficit hypothesis that is not captured by rate-accuracy hypotheses given that phonological awareness is more correlated with word recognition and rapid naming with fluency.

Summary: Cognitive Processes

There is substantial evidence for a sequential relation in which phonological awareness must develop in order to learn to read words; the relations then become reciprocal. This relation is not unchallenged (Castles & Coltheart, 2004; but see Hulme, Snowling, Caravolas, & Carroll, 2005). Hulme, Bowyer-Crane, Carrol, Duff, and Snowling (2012) completed a mediation analysis of data from a randomized trial of kindergarten children at risk for reading problems. Children who received preliteracy instruction had higher reading skills, while those who received oral language training had better language abilities. For the mediation analysis, these groups were collapsed and path analyses were conducted to test the hypothesis that the letter–sound/phonological awareness skills at the end of the intervention mediated the growth of reading skills 5 months after the intervention. These results were supported, suggesting a causal relation.

The results from this randomized trial were consistent with longitudinal studies showing temporal continuity of phonological awareness skills as precursors of reading ability (Wagner et al., 1994). Taken together, the data supporting phonological awareness as a pivotal factor in reading development and disabilities is well established (Hulme & Snowling, 2009; Shankweiler & Crain, 1986). The relation of RAN and phonological memory is less clear, especially because they are correlated with phonological awareness. Some argue that deficits in RAN are manifestations of phonological processing difficulties, while others argue for independence. As we discuss in Chapter 10, RAN seems more related to fluency than to accuracy, but both fluency and accuracy defeats are related to WLRD.

Other unitary explanations of the cognitive processes associated with WLRD are less compelling, although any single-deficit theory is not adequate for explaining all cases of WLRD. Any theory and related hypotheses vis-à-vis WLRD must demonstrate that the hypothetical constructs (visual, auditory, cerebellar) are essential for proficient word reading and deficient in children with WLRD. Studies of visual processes should attempt to link more formally with theories of word recognition in an effort to more fully develop these hypotheses. Speech perception problems may make it more difficult to develop phonological awareness and grasp the alphabetic principle, but the specificity of such deficits to WLRD is not well established. However, it is not likely that speech perception problems explain the range of reading difficulties because there are multiple barriers to acquiring language and literacy across different populations (Pennington, 2009) and among children who struggle to acquire oral language. Studies of other lower-level auditory deficit explanations, especially those based on rapid temporal processing, do not provide compelling explanations of the core reading problem observed in children with dyslexia. The evidence for behavioral (reading) deficits being mediated by the cerebellum in children with dyslexia is inconsistent, and little support emerges from intervention studies targeting motor functions hypothesized to be subserved by the cerebellum. A major goal should be to identify how the cerebellum impacts word-reading fluency, especially given the evidence for differences in cerebellar structure in people with and without dyslexia. This is a plausible approach given the potential role of the cerebellum in automaticity of cognitive functions.

NEUROBIOLOGICAL FACTORS

The hypothesis that LDs are "unexpected" stems in part from the belief that if children who experience low achievement due to factors such as economic disadvantage and inadequate instruction are excluded from the LD category, the cause in those who have low achievement not due to the exclusions must be intrinsic to the child, that is, neurobiological in origin (see Chapter 2). The intrinsic nature of LDs was inferred initially from the linguistic and behavioral characteristics of adults with documented brain injury. As the field progressed, definitions of LDs continued to attribute them to intrinsic (brain) rather than extrinsic (e.g., environmental, instructional) causes, even though there was no objective way to adequately assess the presence of putative brain damage or dysfunction. This problem was constantly dismissed as a matter that technology would eventually resolve! This conviction was reinforced by the common nonspecific association of indirect indices of neurological dysfunction with LDs, including perceptual–motor problems (i.e., difficulty in copying geometric figures), paraclassical or “soft” neurological signs (e.g., gross motor clumsiness, fine motor incoordination), and anomalies on electrophysiological measures (Dykman, Ackerman, Clements, & Peters, 1971). Even at the time, the lack of specificity of these observations to either LDs or neurological integrity was widely acknowledged (Rutter, 1982). Over the past three decades, the quality of the evidence has improved. It is now possible to clearly support the hypothesis that LDs in general, and dyslexia in particular, have a locus in neurobiological factors. But the evidence also suggests that causal models in which neurobiological deficits produce a child with dyslexia are simplistic and do not take into account the complex interplay of the brain and the environment in development. In this section, we review studies of (1) brain function, (2) brain structure, and (3) genetics. Most of these studies explicitly identified children as reading-disabled on the basis of word recognition and phonological processing abilities, so they tend to be specific to dyslexia.

Brain Function

Multiple functional neuroimaging modalities have been used in the study of dyslexia, representing different methods for assessing variations in brain activation patterns in relation to cognition. The modalities (see Papanicolaou, 2017) include: positron emission tomography (PET); functional magnetic resonance imaging (fMRI); magnetic source imaging (MSI); and magnetic resonance spectroscopy (MRS). In this section, we also mention measures involving electrophysiological methods in context, but do not describe these studies in detail, as their potential for brain mapping is less well developed than that of these other methods.

Functional imaging modalities attempt to measure changes in the brain that occur during cognitive processing, and then to construct maps that demonstrate where (and sometimes when) in the brain these changes occurred. For example, metabolic changes reflected by glucose utilization or shifts in blood flow from one part of the brain to another can be measured for different tasks, depending on the mental operations and regions of the brain involved in the operation. These changes can be recorded by PET or fMRI.

Similarly, neurons make connections to support a particular activity. When neurons make connections, the properties of these neurons change, which in turn alter brain electrical activity. This activity can be recorded by an electroencephalogram (EEG). Changes can also occur in the magnetic fields surrounding these electrical sources when a person performs an activity. MSI measures these changes, providing information about what brain areas produce the magnetic signals. MRS measures changes in brain chemistry, such as lactate or glutamine, in response to some type of challenge (Hunter & Wang, 2001).

These modalities also vary in their spatial and temporal sensitivity. Metabolic techniques like PET and fMRI assess brain activity that occurs after the cognitive activity has occurred. They do not occur in real time. In fMRI, serial magnetic resonance images are acquired so rapidly that they can be used to capture the changes in blood flow associated with cognitive activity. Thus, spatial resolution with fMRI is excellent.

Reading and the Brain

Previous research has used all four functional imaging modalities, with converging findings suggesting that tasks requiring word reading in a variety of paradigms are associated with increased activation in three primary left hemisphere areas: the basal surface of the temporal lobe extending into the occipital region (occipitotemporal); the posterior portion of the superior and middle temporal gyri, extending into middle, supramarginal and angular gyri (temporoparietal); and the inferior frontal lobe regions (Dehaene, 2009; Price, 2012). This network is often referred to as a “triangle model,” and is depicted in Figure 6.2, which also shows the semantic regions in the left inferior temporal lobe. The area involving language regions in the middle temporal role is often referred to as a dorsal route (Panel B) and is more involved in processing segments of words (i.e., sublexical). The ventral route mediates direct access to meaning through the occipitotemporal regions (i.e., lexical), including an area that is termed the “word form” area that is hypothesized to deal with larger chunks and whole words.

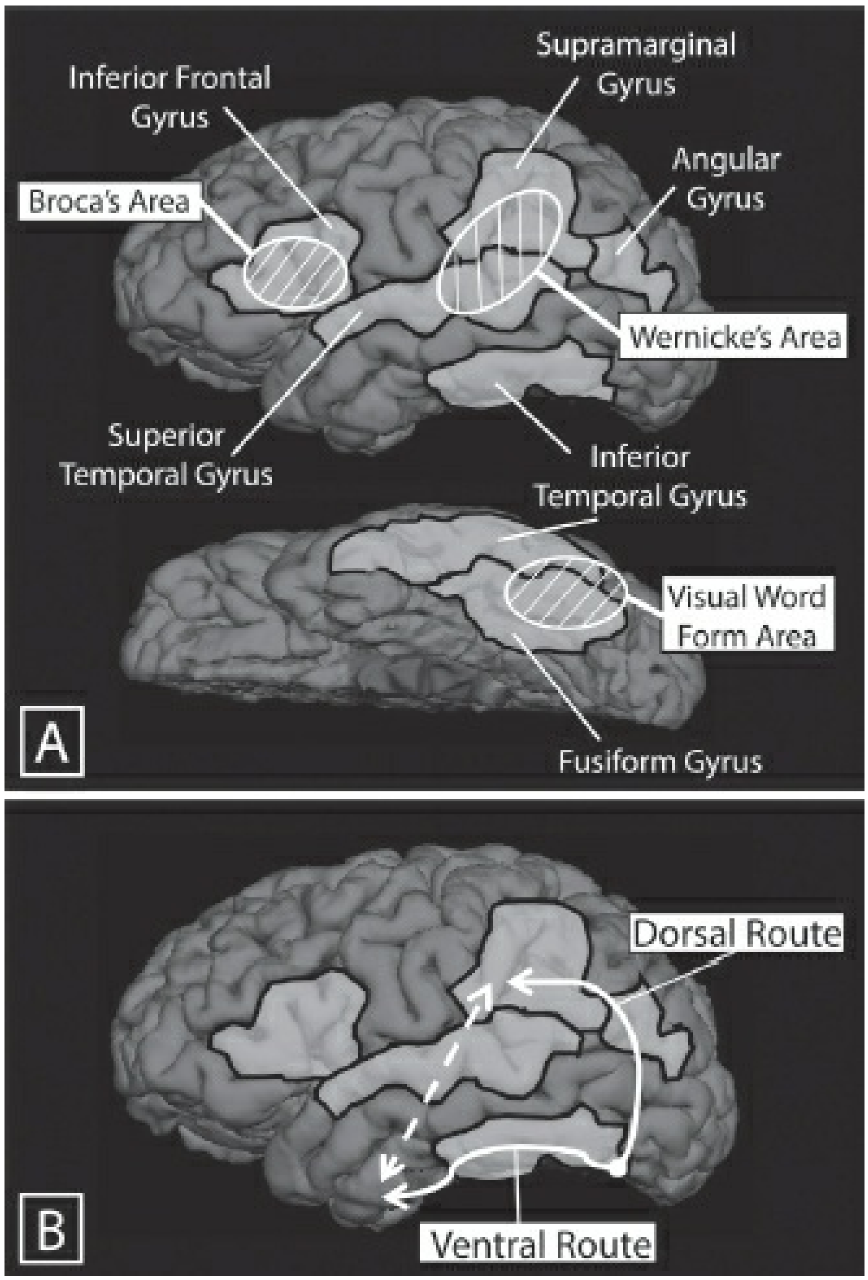

FIGURE 6.2. Simple model of the neural network for reading, showing major participating areas of the brain. In panel A, regions of interest making up the dorsal and ventral streams are shown. In panel B, the dorsal and ventral streams are illustrated, culminating in reciprocal links with semantic systems

in the broadly defined inferior temporal region. Courtesy Victoria Williams.

This triangle network does not develop independently of exposure to print and is dependent on evolutionarily based networks for language and visual processing that allow the extraction of language from print (Dehaene, 2009; Wolf, 2007). The evolutionary basis of language and its relation to reading is well known and stems from the discovery of the alphabetic principle (Liberman, 1996). Less recognized is the programming of occipitotemporal systems with an evolutionary basis for visual processing and attention that subserve, for example, face and object processing, and which through exposure and training become highly specialized for processing print (Vogel, Petersen, & Schlagger, 2014). The evidence for disruption of these systems in people with WLRD has been documented through meta-analyses of the PET and fMRI literature as well as MSI and functional connectivity studies that are generally not included in the meta-analyses. It is important to recognize that the brain operates as a network, or circuit, and that reading and dyslexia do not boil down to a single brain region and certainly not to a lesion or pattern of lesions, as in brain injury.

From MSI studies, we have an understanding of the time course of activation of this network (Simos et al., 2009). After initial activation of the primary visual cortex (not depicted in Figure 6.2), within 70 milliseconds of exposure to a word, the visual association cortex in the occipitotemporal region is activated. This area is responsible for graphemic analysis, rapid processing of orthographic relations, and multimodal integration of different features of the word. It is organized in a posterior to anterior direction, with more posterior regions extending into the occipital regions engaged in sensory processing, middle areas specialized for object and pattern processing, and more anterior areas, including the fusiform gyrus, providing immediate recognition of patterns (see Figure 6.2, Panel A). Although not precisely localized to the fusiform gyrus, this area is considered a “word form” area that facilitates immediate word recognition based on orthographic patterns with direct access to brain systems mediating the sound and meaning of the word (Dehaene, 2009; Price, 2012). As identified in Figure 6.2 (Panel B), this component of the network is often referred to as a *ventral* pathway and is contrasted with a *dorsal* pathway that extends from the occipitotemporal regions to an area roughly corresponding to Wernicke’s area, which includes the superior temporal and supramarginal gyri. This part of the network is activated simultaneously about 750 milliseconds after exposure to a word. It is responsible for phonological processing and letter–sound correspondence. This pathway may extend to the angular gyrus, which historically has been seen as operating as a relay station that links information across modalities, integrates phonological and semantic information, and may be involved in regulating visual attention to words. An area roughly corresponding to Broca's area and the inferior frontal region is responsible for phonological processing that involves articulatory mapping as in the pronunciation of words and access to lexical representations of words. It is also activated simultaneously with other regions of the dorsal pathway.

Why Is a Network of Brain Regions Needed?

To understand why a network of brain regions is necessary, consider the multiple processing demands of a single word like *cat*." This word has a visual representation, which is the word as encountered on a page. Identification of *cat* relies initially on recognition of the patterns and the statistical properties of the order of the three letters (e.g., *ca* is more likely than *ct*), or *orthography*. The word is also composed of three phonemes. In a brand new reader never exposed to print, the letters would be recognizable and could be matched, but not understood as letters, much less a word. In order to understand the letters as a word, the *phonological structure* must be accessed. There would need to be some capacity for phonologically representing speech (phonological awareness) and for mapping the phonological representations onto speech (the alphabetic principle). The novice reader, who may be a very young child or someone who is illiterate and never exposed to written language, would have no capacity for processing the orthographic components as letters or to represent the letters as a set of phonemes. As Morgan (1896) described, the words have no significance to the person even though the letters can be seen, matched, and drawn. This is why reading cannot be considered a product of evolution or as "natural" as learning to talk (Liberman, 1996). It is also why reading words is a complex cognitive skill that cannot be localized to a single area of the brain. In fact, the other component that must be processed, which is the meaning of *cat,* relies upon a *semantic* system that is part of oral language and is already substantially developed by the time the child is faced with the task of learning to read. However, if the only word encountered is *cat,* accessing the specific meaning in the text will be difficult. There are many types of cats, with adjectives like *big, house,* or *jungle* modifying the exact meaning of *cat*. For meaning to occur at a fundamental level, the word *cat* must be read as part of a sentence or within the context of a story.

What is specific to reading is the need to map phonemes onto the orthography, and eventually to process the orthographic patterns at the "speed of sight" (Seidenberg, 2017). The skilled reader does this automatically and devotes little attention to the phonological representation of the words. But the unskilled reader (a beginning child or an illiterate adult) or the person with dyslexia cannot easily map the phonological representations onto orthography, so automaticity does not develop. Most importantly, to understand dyslexia, it is essential that this fundamental problem—the inability to recognize or spell an isolated word—be addressed. Because text reading among individuals with WLRD is frequently impaired, there are few opportunities to take advantage of context to assist in decoding single words. This finding explains why instructional approaches that teach the use of context or multiple cues to read unfamiliar words are ineffective for people with dyslexia (and many beginning readers; Adams, 1990). The alphabetic principal must be taught, and taught quite explicitly to many people with and without dyslexia. In addition, component visual processes must become highly tuned to attend to and process letter patterns. Obviously, letter-by-letter or sound-by-sound reading would be so slow that the person would not be able to create a coherent mental representation of the text because of the limits of working memory. In turn, slow and nonautomatic reading of words negates access to longer term memory for conceptual understanding. In contrast, the skilled reader goes directly to the meaning of the word and text from the patterns made by the visual representation of the word through the occipitotemporal region of the neural network involved in deciphering written language (Dehaene, 2009).

Dual-Route Theory and Connectionism

Implicit in the triangle model is the idea that there are multiple pathways for learning to read words. The brain, through experience (instruction and reading exposure), develops the capacity for phonological and orthographic processing of written language by utilizing areas of the brain intended for phonological processing of language and for visual processing. The easiest way to understand the operation of this network is through the cascading dual-route model, which we have referred to throughout this chapter (Castles & Coltheart, 1993). There are several variations in these models (Dehaene, 2009) as well as alternatives involving connectionist models (Plaut, McClelland, Seidenberg, & Patterson, 1996).

Dual-Route Models

In the dual-route model of word reading, the network of regions involves two distinct neural processing pathways. Shifts occur depending on symbolic rules that differentially activate these pathways. The *dorsal phonological system* is responsible for sublexical decoding and is largely phonological and essential for dealing with novel words. As depicted in Figure 6.2, it involves the left temporoparietal cortex, including the angular gyrus, posterior superior temporal gyrus, and the supramarginal gyrus. Reading through this route requires phonological decoding to mediate between orthographic representations and meaning. The *ventral lexical system* allows for immediate recognition of words based on the orthographic pattern and direct access to the semantic system. As readers become more experienced, the ventral system becomes highly tuned to orthographic patterns, so that the dorsal system is bypassed unless a novel word is encountered. By retrieving whole-word forms, there is immediate access to the meaning of the word and reading becomes increasingly automatic, the key to proficient reading. To reiterate, in beginning, dyslexic, or illiterate readers, the ventral system is not highly specialized for print and sublexical representations are essential for learning how to map words to meanings.

Connectionist Models

Computational modeling of the processes in word and text reading have generated somewhat different explanations of the neural and cognitive processes involved in reading. Connectionist models are based on weighted contributions of distributed visual and language processes. Although the two models converge on the idea of separate neural pathways for lexical (semantic) and sublexical (phonological) processing, connectionist models suggest that these operations occur simultaneously and in parallel based on the weighted connections of different components of the word, so that the underlying mechanism is singular, not dual. Reflecting the absence of assumptions about how reading is represented in the brain, which in a dual-route model stem from studies of acquired reading disorders and the distinction of surface and phonological (deep) alexia (reviewed above for dyslexia) a connectionist model proposes that the reading network emerges based on exposure to words and text. Written language is coded as distributed representations that contain multiple units that allow the brain to become increasingly proficient in automatic word recognition of familiar and novel words. Training and feedback activate different orthographic patterns and lead to the development of computational rules that become increasingly automatic. Simultaneous paths address the relation of orthography to phonology through exposure to the meaning of the word. In essence, the relation of orthography and phonology is learned by exposing the brain to a visual representation, attempting to encode the pronunciation of the word, and receiving feedback on the correct pronunciation, thus increasing the strength of connections among the different units of the word: orthographic, phonological, and semantic (Taylor, Rastle, & Davis, 2013).

The two models converge around the idea of dorsal and ventral pathways and are becoming increasingly integrated (Foorman, 1994; Seidenberg, 2017; Taylor et al., 2013). Much of the empirical evidence that supports this model of the brain circuits that mediate word reading is also consistent with studies of acquired reading difficulties secondary to brain damage (Dehaene, Cohen, Sigman, & Vinckier, 2005) and the effects of transient interference with normal function in specific brain areas because of neurosurgical operations (Simos et al., 2000). Most critically, there is a substantial research base on how disruptions of this network are related to dyslexia.

Functional Neural Correlates of WLRD

Here we review functional neuroimaging studies of children and adults with dyslexia. In this review, most studies are cross-sectional. There are interesting questions about developmental trajectories that are not well addressed by these meta-analyses, which require longitudinal research and were recently reviewed by Black, Xia, and Hoeft (2017). Whenever possible, we rely upon meta-analyses, which synthesize different studies—typically with small samples—into a larger empirical data set. Meta-analysis is important because individual studies often give somewhat different results because of variations in participants, languages, tasks, scanners, and imaging sequences. Some of this study-specific variation is minimized by a meta-analytic approach. The general results of these meta-analyses are identified in Plate 1, which shows a composite across multiple studies of differences in patterns of brain activation in children and adults with dyslexia versus controls using fMRI (Richlan, Kronbichler, & Wimmer, 2011). In the top part of Plate 1, children show reduced activation of the left inferior parietal region (dorsal), extending to the supramarginal gyrus. There is also a significant pattern of underactivation in the ventral occipitotemporal region. Adults also show underactivation in the dorsal regions, but more in the superior temporal gyrus. There is much greater underactivation of the ventral occipitotemporal region. There is also underactivation of the inferior frontal gyrus. The main area of overlap is in the anterior portion of the left occipitotemporal region. In the bottom part of Plate 1, latency data from the MSI study by Rezaie et al. (2011b) are shown for prominent portions of the dorsal (superior temporal) and ventral (fusiform gyrus) circuits. The delay in latency to reading pseudowords is bilateral and striking in the poor readers. In the next section, we build on these data through meta-analysis.

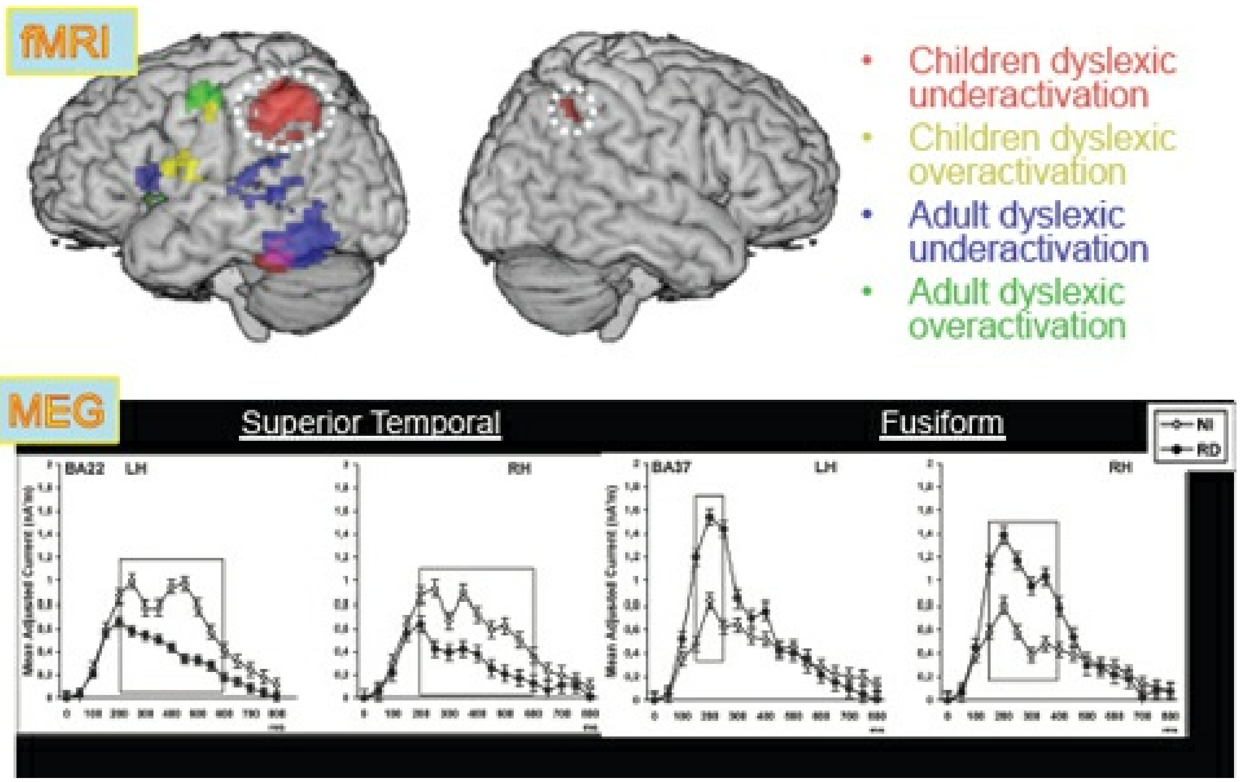

PLATE 1. Brain areas in the dorsal and ventral streams that show underactivation and overactivation in children and adults with dyslexia in meta-analytic synthesis. The top figure is from a meta-analysis of fMRI data from Richlan, Kronbichler, and Wimmer (2011, p. 1738). Copyright © 2011 Elsevier Publications. Reprinted by permission. The bottom figure represents MSI latency data to a pseudoword reading task from Rezaie et al. (2011b, p. 879). Copyright © 2011 Cambridge University Press. Reprinted by permission.

Meta-Analyses of Functional Imaging Studies

In early meta-analyses addressing differences in brain activation in children and adults with dyslexia and typical readers, many of the synthesized studies focused on adults, but tended to converge on the pivotal roles of the occipitotemporal region and the temporoparietal region, with less convergence on the role of the inferior frontal gyrus and the angular gyrus. The issue with the inferior frontal gyrus is whether this region is overactivated or underactivated in poor readers. In meta-analyses of typical readers (e.g., Cattinelli, Borghese, Gallucci, & Paulesu, 2013), the inferior frontal gyrus has been related to requirements for effortful processing, while the angular gyrus was clearly involved in recognition of real words, and even more with pseudowords. Studies of people with reading difficulties tend to use easier tasks that minimize level of effort to accommodate the reading problems. But the basis for the inconsistent results for the angular gyrus is not clear and more attention has been devoted to the occipitotemporal region as an organizer and relay station for any form of print, including pseudowords.

In an early meta-analysis, Maisog, Einbinder, Flowers, Turkeltaub, and Eden (2008) reported synthesized PET and fMRI studies of adults identified as "dyslexic" and controls. The results indicated reduced activation in left hemisphere regions involving occipitotemporal, temporopartial (inferior parietal cortex, superior temporal gyrus), thalamus, and inferior frontal gyrus (see Figure 6.2 and Plate 1). In the right hemisphere, reduced activation was apparent in the fusiform, post central, and superior temporal gyri. There was no evidence for differences in cerebellar activation or for *increased* activation of the inferior frontal region, but there was increased activation in poor readers in the right thalamus and anterior insula. The most robust finding was reduced activation in the occipitotemporal regions of the left hemisphere, particularly the fusiform gyrus representing the visual word form area.

A second meta-analysis (Richlan, Kronbichler, & Wimmer, 2009) involved 17 fMRI or PET studies of reading and/or phonological processing in samples identified with "dyslexia" versus "controls" unrestricted for age. Across studies, groups with dyslexia showed reductions in activation of left hemisphere regions including the temporoparietal (inferior parietal, superior temporal, middle and inferior temporal regions), and occipitotemporal regions, including the fusiform gyrus. Underactivation of the inferior frontal gyrus and hyperactivity of the primary motor cortex and anterior insula were also reported. The angular gyrus was again reported as showing no differences in activation despite individual studies highlighting its role as part of the dorsal system. Richlan et al. suggested that involvement of the inferior frontal region may be age-dependent, with children with WLRD more likely to show overactivation.

In a third meta-analysis, Richlan et al. (2011) synthesized findings from nine studies of children and nine studies of adults identified with "dyslexia" and controls. They found the strongest evidence for reduced activation of the left occipitotemporal region in both children and adults, with a tendency for children to show more underactivation in the left inferior parietal regions (Plate 1) regions.

In more recent meta-analyses, Paulesu, Danelli, and Berlingeri (2014) synthesized findings from 53 studies that included children and adults with dyslexia. As is the case in most meta-analyses, the studies varied in how participants were identified, primary languages, reading and nonreading tasks, and scanner/imaging sequences. The researchers used a hierarchical clustering approach to determine whether there were different activation sites for good and poor readers. In the controls, there was clear activation of multiple areas within the ventral occipitotemporal region, which was not present in people with dyslexia. They also found activation of the dorsal middle temporal and supramarginal gyri, which was not present in people with dyslexia. These patterns were most apparent on reading tasks and on phonological tasks in the auditory modality. The dorsal activations also included inferior parietal and supplementary motor areas involved in motor control and the superior parietal area, which the investigators argued was related to task demands for eye movements and spatial attention. In the right hemisphere, there was consistent evidence for activation of the superior parietal lobe hypothesized to undergird spatial attention in controls, but not in people with dyslexia. Among individuals with dyslexia, activation was present in areas adjacent to the superior parietal lobe, suggesting a more disorganized response to visual stimuli. There was little evidence of activations involving the cerebellum or inferior frontal regions, or for neural systems associated with the magnocellular hypothesis, presumably reflecting responses to the need for eye movements. As shown in Plate 1, *underactivation of the ventral left occipitotemporal and dorsal left temporoparietal regions were consistent findings among individuals with dyslexia*. There was overlap of phonological and reading tasks and the auditory and visual tasks, with less overlap for nonreading tasks.

In their meta-analysis, Pollack, Luk, and Christodoulou (2015) focused only on fMRI and PET studies that used reading-related tasks in adults and children. They identified 13 studies across six languages and conducted separate meta-analyses for the groups with dyslexia and age-comparable controls. The controls showed activation of a left hemisphere network that included left frontal and temporal lobe components extending back to the fusiform gyrus. Poor readers did not show this left-lateralized pattern, instead showing a more bilateral pattern. There were no differences in activation of the inferior frontal or left insula regions, although poor readers tended to show bilateral activation.

Most recently, Martin, Kronbichler, and Richlan (2016) examined 28 functional neuroimaging studies of dyslexia equally divided between languages like English with opaque orthographies and transparent orthographies. Separate meta-analyses of these two sets of 14 studies revealed underactivation of the left occipitotemporal region in the group with dyslexia, which included the word-form area, and underactivation of the left middle and temporal regions. In comparing across orthographies, there was more underactivation in the opaque compared to the transparent orthographies in the left inferior frontal gyrus, left precuneus, and right superior temporal gyrus, with overactivation in the left anterior insula. Transparent orthographies compared to opaque orthographies were associated with underactivation of the left fusiform gyrus, left temporoparietal, left pars orbitalis, and left frontal operculum, with overactivation in the left precentral gyrus. These results show that underactivation of the ventral and dorsal pathways are universal, with variations depending on characteristics of the orthography, especially in frontal and potential compensatory regions.

Altogether, these six meta-analyses show multiple dysfunctional neural signatures in dyslexia within the reading network apparent in typical readers. As Plate 1 shows, most striking is the underactivation of the ventral occipitotemporal region, which is heavily involved in orthographic processing. Also replicated was involvement of the middle and superior temporal regions that are part of the dorsal stream and important for phonological processing. Hyperactivation of the inferior frontal regions in people with dyslexia was not consistently apparent, nor was there clear evidence from the meta-analyses for involvement of the angular gyrus or cerebellum. Equally striking is increased activation of a variety of right hemisphere regions, leading to a diffuse, somewhat bilateral, pattern of activation in people with dyslexia. *In essence, this research suggests a failure to establish the left hemisphere network needed to mediate reading.* There is also evidence that the patterns are age-related and change over time in good and poor readers (Black et al., 2017). To fully appreciate this functional network, it is important to consider other imaging modalities that were generally not included in these meta-analyses.

Magnetic Source Imaging

Studies using MSI, also known as magnetoencephalography (MEG), have observed reliable differences in activation patterns of children with well-defined WLRD and typically achieving children. Unlike PET and fMRI, MSI directly assesses neuronal signaling in real time (Cheyne & Papanicolaou, 2017). Neurons are constantly active, firing bursts of electrical signals in response to cognitive processes. Although the sources of these currents cannot be directly measured, they also produce small amounts of electromagnetic energy that emanate around the source and travel outside the head. These minute magnetic signals are not distorted by the passive electrical properties of brain tissue and can be captured by MSI in the form of a magnetic field (or magnetic flux) distributed along the head surface. Local, transient changes in magnetic flux are recorded by superconducting loops of wire (magnetometers) that are contained in a helmet-like device covering the head. Based on the recorded changes in the surface distribution of electromagnetic energy, researchers obtain precise location estimates for active neurons using simple statistical modeling techniques.

In a series of studies of adults and children who vary in reading proficiency, good and poor readers defined on the basis of word-reading criteria did not differ in activation patterns when they *listened* to words, showing patterns predominantly in the left hemisphere consistent with expectations for an auditory listening task (see Simos, Rezaie, Fletcher, Juranek, & Papanicolaou, 2011a). However, on *printed* word recognition tasks, striking differences in the activation patterns of good and poor readers occur. In proficient reading children, the primary pattern shows initial activation of the occipital areas of the brain that support primary visual processing. Then the ventral occipitotemporal regions in both hemispheres were activated, followed by simultaneous activation of three areas in the *left* temporoparietal region (essentially the angular, supramarginal, and superior temporal gyri, encompassing Wernicke's area). In the children with reading problems, the same pattern and time course was apparent, but the temporoparietal areas of the *right* hemisphere were activated.

Simos et al. (2011a) obtained MSI data from continuous word recognition tasks (auditory and visual) in children who experienced reading difficulties ($N = 44$) and typical readers ($N = 40$). Minimum norm estimates of regional neurophysiological activity were obtained from magnetoencephalographic recordings at 3 millisecond intervals. There were no differences on the auditory task. On the reading task, poor readers showed reduced activity in the superior and middle temporal gyri bilaterally during late phases of word reading. Increased activity in prefrontal, mesial temporal, and ventral occipitotemporal cortices, bilaterally, was apparent. The temporal profile of activity in the group with reading difficulties was markedly different from the typical reader group, showing simultaneous activity peaks in temporal, inferior parietal, and prefrontal regions. In other studies, the patterns observed in children with dyslexia have not been apparent in children with ADHD, but do occur in children with comorbid WLRD and ADHD (Simos et al., 2011b) and are not related to IQ or IQ discrepancy (Simos et al., 2014).

On the whole, the findings are similar to those from the PET and fMRI studies, but the differences between good and poor readers are more strikingly lateralized, especially in areas associated with temporoparietal regions. The posterior to anterior time course of word reading is clearly apparent.

Functional Connectivity

In older PET and MRI studies, the correlations of different brain regions were assessed to examine connectivity during task-related activation paradigms. For example, Horwitz, Rumsey, and Donahue (1998) used PET and computed within-task, cross-participant correlations that showed the left angular gyrus was disconnected from the occipitotemporal and temporal lobe regions during a word-reading task. Pugh et al. (2000) used fMRI across a series of print-related tasks to examine the covariance of the angular gyrus with other regions of the reading network. They found much lower correlations among left hemisphere brain regions associated with the phonological processing demands of different tasks in children with dyslexia compared with controls. Stanberry et al. (2006) used functional connectivity MRI to examine the synchronicity of the hemodynamic response to a visual phoneme mapping task with resting and activated states. This method requires selection of brain regions for “seeding” to determine functional connectivity. They found reduced connectivity of the left inferior frontal gyrus with other frontal, occipital, and cerebellar regions in adults with dyslexia. There was little connectivity with the angular gyrus, but connectivity of the angular gyrus was apparent with other brain regions involved in reading in controls. Using similar methods, but focusing on seeding the occipitotemporal region, van der Mark et al. (2011) found reduced connectivity of the visual word-form area with the inferior frontal and inferior parietal language areas during reading in the left hemisphere of children with dyslexia.

Newer developments in fMRI involve the assessment of brain activation in resting states, that is, without specific activation tasks. These methods involve low-frequency scanning of fluctuations in blood flow during periods when the participant is not formally engaged in a specific task. Because the functional networks should show connectivity even at rest, seeding relevant areas of the reading network should elucidate patterns of connectivity. Thus, Hampson et al. (2006) found positive correlations in resting state of the inferior frontal region and angular gyrus, as well as the occipitotemporal region (including the fusiform gyrus). In a study of 25 adults varying in reading proficiency, Koyama et al. (2010) seeded six left hemisphere regions identified as involved in word reading: the left fusiform gyrus, the left superior temporal gyrus, the left temporoparietal junction, the left precentral gyrus, the left inferior frontal gyrus, and the posterior part of the left inferior occipital gyrus. There were positive relations among each of these six regions of interest, especially the left temporoparietal junction (including the angular gyrus) and left frontal and temporal regions, a finding consistent with the correlational studies of connectivity reviewed above (Pugh et al., 2000). There was also strong connectivity of the left fusiform gyrus and the left inferior frontal gyrus. Interestingly, patterns of connectivity tended to be more bilateral than in task-activated fMRI. In children with dyslexia, a small pilot study of five children found reduced connectivity between the left and right inferior frontal gyri (Farris et al., 2011). Koyama et al. (2013) seeded 12 regions, some of which were regions of interest in the reading network and others related to motor and attention skills implicated in other studies. The groups of interest for this review involved 11 children with a history of dyslexia and low word-reading scores. For this group, there was reduced connectivity in between the left intraparietal sulcus and the left middle frontal gyrus, with generally reduced connectivity within frontal systems related to attentional control. These regions correlated with reading scores and behavior ratings of inattention, although only a few participants were identified with ADHD. They did not observe differences in the connectivity of reading-related areas.

Finn et al. (2014) completed a whole-brain analysis of resting state functional connectivity in a large sample of children and adults with dyslexia. They found reduced connectivity among people with dyslexia in areas associated with different components of the primary visual pathway and adjacent association areas involved in reading. In addition, similar to the Koyama et al. (2013) findings, the connectivity of these regions with prefrontal regions associated with attention was weaker in groups with dyslexia, suggesting that top-down processes associated with regulation of attention were not developed as well in people with dyslexia. Language was less lateralized in the left hemisphere of children with dyslexia, with stronger connectivity to homologous right hemisphere areas; this pattern was less apparent in poor-reading adults. In addition, there was not differential connectivity of the visual word-form area with other regions of the reading network; in poor-reading adults, the connectivity was bilateral. The researchers argued that this pattern suggested a stronger role of visual attention in dyslexia than reported in previous fMRI studies, consistent with other studies showing connectivity of the occipitotemporal region and frontal regions associated with attentional control in proficient readers (Vogel, Miezin, Petersen, & Schlagger, 2012).

In a study of German adolescents with dyslexia that used reading task-based fMRI and resting state fMRI, Schurz et al. (2015) examined activation and connectivity in nodes defined according to the three components of the left hemisphere reading network: occipitotemporal (fusiform, inferior temporal), temporoparietal superior temporal (left inferior parietal), and the middle temporal and inferior frontal gyri. They reported reduced connectivity in dyslexia of multiple left posterior regions in the occipitotemporal and temporoparietal components and the inferior frontal gyrus. A similar pattern was apparent on task activation. They also found evidence of reduced connectivity of the inferior parietal lobe and the middle frontal gyrus, which they interpreted as evidence of executive control of attention.

Functional Neuroimaging Studies and Neural Plasticity: Intervention

Another approach to understanding how the neural network functions in people with dyslexia is to evaluate potential changes in functional and structural organization of the brain before and after a reading intervention. These studies began to emerge in 2000 and used different functional imaging modalities, including fMRI (Aylward et al., 2003; Bach, Richardson, Brandeis, Martin, & Brem, 2013; Davis et al., 2011; Eden et al., 2004; Farris et al., 2011; Farris, Ring, Black, Lyon, & Odegard, 2016; Gaab, Gabrieli, Deutsch, Tallal, & Temple, 2007; Gebauer et al., 2012c; Odegard, Ring, Smith, Biggan, & Black, 2008; Hoeft et al., 2007; Meyler, Keller, Cherkassky, Gabrieli, & Just, 2008; Richards et al., 2006; Richards & Berninger, 2008; B. A. Shaywitz et al., 2004; Temple et al., 2003; Yamada et al., 2011), MSI (Simos et al., 2002a, 2002b, 2005, 2007a, 2007b; Rezaie et al., 2011a, 2011b), MRS (Richards et al., 2000), and resting state fMRI (Koyama et al., 2013).

Many of the fMRI and MSI studies were reviewed by Barquero, Davis, and Cutting (2014), who also conducted a meta-analysis of eight fMRI studies that provided post-intervention imaging data. In their narrative review, the authors observed a predominance of changes in the reading network that were normalizing as well as changes that were compensatory and often outside the reading network in Figure 6.2. The most consistent evidence involved the middle and superior temporal gyri, which are part of the dorsal (sublexical) pathway. In many studies there was significant normalization (increased activation) of activity in the left hemisphere regions, with reduced right temporoparietal patterns of activation that were much more apparent before intervention. This pattern may well reflect overengagement of right hemisphere dorsal systems prior to intervention, with subsequent increases in left hemisphere ventral and dorsal regions activation in interventions that had a significant decoding component. For the inferior frontal gyrus, most studies reported underactivation or no differences in activation prior to invention; there was an increase in activation that was often bilateral after intervention. In the occipitotemporal region, there was either no difference in activation or underactivation before treatment; after intervention, activity increased in many studies, sometimes bilaterally. MSI studies also report reduced latencies (i.e., increased efficiency) to word stimuli after intervention, often bilaterally. Across studies, there was considerable variability in the degree adjacent areas to these regions were activated, which may reflect differences in the participants and interventions.

The meta-analysis did not strongly replicate the narrative review, reflecting the small sample of studies (and participants) and the exclusion of the MSI studies. There was increased activity in the left thalamus, right insula/inferior frontal gyrus, left inferior frontal gyrus, right posterior cingulate, and left middle occipitotemporal regions. Activations of the left inferior frontal gyrus (articulatory coding) and occipitotemporal regions (orthographic processing) are not unexpected. The right inferior frontal and posterior cingulate are often related to attention. Normalization of the middle and superior temporal regions was not apparent, but this may be due to exclusion of the MSI studies, which have reliably shown changes in activation of these areas to intensive interventions in well-defined children with or at risk for dyslexia. MSI is very sensitive to the temporal lobe regions involved in language, reading, and memory, which is why it is used for presurgical mapping in epilepsy. In reading studies, MSI is less sensitive to differences in the occipitotemporal regions often revealed in fMRI studies, except that latency differences are commonly reported in MSI studies, which are not measurable with fMRI. In the section below on intervention, we present additional figures showing neural changes associated with intervention response.

Functional imaging studies have also shown that intervention response can be predicted from degree of activation. Rezaie et al. (2011a, 2011b) reported that the degree of dorsal system activation predicted intervention response using MSI in a subset of adolescents participating in Vaughn et al. (2010a; see Plate 6). In fMRI studies, Hoeft et al. (2011) found that baseline activity in the right inferior frontal gyrus at baseline predicted word-reading scores 1 year later in a longitudinal study with no specified intervention. In a subset of the fifth-grade children imaged in Meyler et al. (2008), Hoeft et al. (2007) reported that increased activation in the right fusiform, middle occipital, and left middle temporal gyri, accompanied by reduced activation in the right middle frontal gyrus, predicted intervention response. Davis et al. (2011) found increased activation in the left middle and superior temporal regions, aligning with the MSI studies. Odegard et al. (2008) conducted a postintervention comparison of adequate and inadequate responders and found differences reflecting increased activation in the left inferior parietal region in responders and greater activation of right middle temporal regions in inadequate responders. Farris et al. (2016) reported greater activation in the left inferior frontal region was associated with larger improvements in sight-word efficiency and phonological decoding for the group of children with dyslexia, with increased activation in the right inferior frontal gyrus also associated with larger improvements in reading, replicating Hoeft et al. (2007).

The Lesson of Illiteracy

There is additional neuroimaging research on adult illiteracy that provides insight into the development of the neural network that mediates good and poor reading. These studies often compare illiterates who have never been taught to read with adults who have had reading instruction later in life, so-called late learners. There appear to be functional and structural brain changes that occur as a person with illiteracy and no formal reading instruction learns to read. Dehaene, Cohen, Morais, and Kalinsky (2015) made two major conclusions from these studies related to functional and anatomical modifications after literacy was introduced.

The first conclusion is that the neural network that emerges with instruction for children and adolescents can be observed in adults acquiring reading skills for the first time, but usually not to the same degree of activation. This is most likely due to the adults having a generally lower level of reading following intervention than people who learn to read as children. Automaticity is a major problem and it may be that print exposure is just insufficient to program the ventral systems. The second conclusion is that not only does the acquisition of literacy rewrite the organization of the brain, acquiring literacy also enhances other skills, presumably due to the enhancements of the neural circuitry mediating reading.

As Dehaene et al. (2015) noted, learning to read is a process of extracting language from vision regardless of the age of the learner. The need to learn to read by scaffolding literacy onto brain areas specialized for language and visual processing is apparent in studies of illiterates and in young children, with rapid specialization of the ventral occipitotemporal region supporting visual processing for processing print (Vogel et al., 2014). The processing of the features of letter and word forms becomes a very specialized function of this brain region even as other aspects of visual processing continue to be supported, but are functionally reorganized. The specialization begins to occur in a matter of weeks in beginning readers (Brem et al., 2010) or in a few days in illiterates (Perrone-Bertolotti et al., 2014). Once the person becomes proficient, the posterior occipitotemporal (ventral) region processes print more rapidly than visual stimuli that resembles print, and is universally specialized across languages and types of writing, including alphabetic and logographic (e.g., Chinese) forms of writing. As this area becomes specialized for print, it does not do so at the expense of other forms of visual processing. In fact, some studies of adult illiterates have shown enhanced visual processing of shapes. Other studies suggest that learning to read in illiterate individuals results in face recognition becoming more strongly lateralized to the right ventral visual regions as the left fusiform gyrus becomes more specialized for reading. This displacement is interesting because face recognition is represented adjacent to the area that becomes specialized for word reading (Dehaene et al., 2010).

The dorsal system, which is already specialized for speech processing through the auditory cortex, also shows changes indicative of functional and structural plasticity in adults. The acquisition of literacy enhances speech processing in the left perisylvan region, which includes the planum temporale. This region becomes much more sensitive to explicit manipulations of the phonemic structure of speech. Preston et al. (2016) found that the degree to which speech and print coactivated different regions in the left hemisphere in kindergarten predicted reading proficiency 2 years later. These relations were especially strong for the inferior frontal gyrus bilaterally and for the left inferior parietal cortex and fusiform gyrus.

Learning to read in illiterates and in children also enhances other skills, such as vocabulary and verbal fluency as more proficient reading increases the amount of print processed by the reader. While there does not appear to be a direct effect of reading skills on executive skills, reading in school enhances organization and self-regulatory skills, including attention (Roberts et al., 2015; Dehaene et al., 2015). These changes in neural representation are clearly experientially induced. Without explicit exposure to print, the brain would not be able to process written language. However, developing literacy is not simply a process of training the brain to recognize orthographic patterns; these patterns must be linked to the linguistic representations in the sublexical and lexical systems (Dehaene, 2009; Preston et al., 2016).

Summary: Functional Neuroimaging

Functional imaging studies of WLRD have proliferated over the past two decades. Although there is variability across different studies, this variability may reflect differences in samples, imaging modalities, and research designs. There are inconsistencies among studies with respect to the engagement of a particular area that may reflect task differences, especially with regard to the involvement of the angular gyrus and the specific role of different regions of the occipitotemporal region (Price, 2012). But what stands out is strong research support for the presence of a "triangle" network signature in proficient readers. In contrast, this neural signature is either dysfunctional or absent in less proficient readers. As Price (2012) stated, "A striking feature is that the same conclusions have been produced over and over again. Although this results in repetitive reading, it is important for validating the findings and demonstrating the remarkable consistency of the functional anatomy across individuals and studies. Yes, there are interesting and relevant sources of inter-subject variability but these are small relative to the consistent effects" (p. 838).

Brain Structure

Before the development of functional neuroimaging methods that could be used with children, noninvasive methods based on MRI were used to try and identify neuroanatomical correlates of dyslexia. These studies became less prominent once the functional studies became possible, but have continued, especially in relation to possible neuroanatomical correlates of the reading network identified in the functional neuroimaging studies. These findings are less consistent than the functional neuroimaging studies, especially when initially completed. Underlying these structural brain studies is a "chicken-and-egg" question that is also apparent in the functional studies: To what extent are differences in brain structure precursors to the reading difficulties or consequences of a lack of reading experience because of an inability to read words, which, in turn, limits access to print and thus stimulation of neural circuits undergirding reading development. It is clear from the intervention-imaging studies that the neural system is malleable and from studies of illiteracy that experience is a key aspect of development of the neural system mediating word reading. But the precise roles of preexisting brain dysfunction and experience are not clear, except that if reading is not taught, the person does not learn to read proficiently and does not develop the neural network necessary to mediate proficiency in reading.

Research on brain structure involves either postmortem studies or the use of imaging techniques such as cerebral computed tomography (CT) and anatomical MRI (aMRI) addressing gray matter (cortical) and white matter (subcortical) volumes, surface area and cortical thickness, and cortical gyrification. There are also studies of white matter connectivity that evaluate axonal integrity using diffusion tensor imaging (DTI).

Postmortem Studies

There are a few postmortem evaluations of the brain anatomy of adults with a history of dyslexia. These cases are rare, as dyslexia is not lethal. These studies, largely by a group led by Galaburda (1993), have involved a total of 10 brains accumulated over several years. The findings indicated that individuals with dyslexia are characterized by differences in the size of specific brain structures (e.g., planum temporale) and the presence of specific neuroanatomical anomalies. The planum temporale, a structure on the planum of the temporal lobe, appears symmetrical in size in the left versus the right hemisphere (Galaburda, Sherman, Rosen, Aboitiz, & Geschwind, 1985). In postmortem studies of adults who presumably did not have reading problems, this structure is often larger in the left hemisphere than in the right hemisphere (Geschwind & Levitsky, 1968). Because this area of the left hemisphere supports language function, the absence of this anatomical difference has been viewed as a partial explanation for language deficiencies that are hypothesized to result in reading problems. In addition, microscopic examination of cortical architecture showed minor focal distortions called "ectopias." Although also common in individuals with no history of dyslexia, these ectopias were more common than would be expected in individuals with a history of dyslexia. They were also more common in the left hemisphere, especially in the perisylvan region associated with speech processing and production. Microscopic examinations of subcortical structures have shown differences relative to normative expectations, particularly in the thalamus (Livingstone, Rosen, Drislane, & Galaburda, 1991). These structures of the thalamus are widely believed to be involved in visual processing. Finally, examinations of the cerebellum in a subset of these brains (Finch, Nicolson, & Fawcett, 2002) revealed larger mean cell sizes in the medial posterior cerebellum relative to normal expectations, as well as unexpected distributions of cells in several parts of the cerebellum.

Altogether, postmortem studies have found evidence of anomalies at both subcortical and cortical levels in many parts of the brain. These anomalies are largely interpreted as problems with neural migration and tend to cluster in the perisylvan regions responsible for language processing, but are spread across the brain. However, these studies are limited because the reading characteristics, educational histories, and important factors that influence brain organization, such as handedness, are difficult to ascertain in a postmortem study.

Anatomical MRI (aMRI)

Given the difficulties involved in obtaining brains for postmortem evaluation, as well as the aforementioned limitations of any postmortem study, investigators have turned to aMRI for the evaluation of potential differences in brain structure. The use of aMRI is desirable because it is noninvasive and is safe for children. The aMRI data can also be segmented and quantified, so that precise measurements of brain structure can be made. The findings can then be correlated with reading performance.

Gray Matter

Volume. The predominant method for assessing gray matter involves measurement of volumes of different brain regions of interest or a structure using voxel-based morphometry. These methods convert the number of voxels in a predetermined MRI cluster to a volumetric measurement across contiguous aMRI slices. Early studies of gray matter volumes examined a variety of structures using manual tracing methods. Given the interest generated by postmortem studies, these included the planum temporale and the temporal lobes. Studies comparing the planum temporale in individuals with and without dyslexia reported both symmetry (Hynd, Semrud-Clikeman, Lorys, Novey, & Eliopulos, 1990) and even reversals in the expected patterns of asymmetry (Hynd et al., 1990) in the groups with dyslexia. Altarelli et al. (2014) used metrics to measure hemispheric brain differences similar to those employed in the postmortem studies and found a greater proportion of boys (but not girls) with dyslexia showing larger right than left planum temporale. However, other studies have not found an association between symmetry of the planum temporale in dyslexia (Eckert et al., 2003; Rumsey et al., 1997; Schultz et al., 1994). The planum temporale is an anatomically difficult area to define. The aMRI assessments of gray matter volume are more consistent in showing reduced volumes in the left temporoparietal regions. Individual studies yield inconsistent findings, reflecting small samples and variations in imaging acquisition and analysis methods (Schultz et al., 1994).

Richlan, Kronbichler, and Wimmer (2013) synthesized nine studies with 134 dyslexic and 132 control participants. Reliable reductions in gray matter were seen in the right superior temporal gyrus and the left superior temporal sulcus. In addition, four of the nine studies reported reductions in gray matter volumes in the left occipitotemporal region, but these were too scattered for reliable convergence. The basis for the reduction in the right and not the left superior temporal sulcus was not clear, but bilateral reductions in the volume of these regions have been reported for young children at risk for reading difficulties. The left temporal sulcus reduction involves the perisylvan language regions and is consistent with the post mortem studies and is perhaps the most reliable finding from aMRI studies of gray matter.

Linkersdoefer, Lonnemann, Lindberg, Hasselhorn, and Fiebach (2012) synthesized findings from 24 studies largely drawn from the functional meta-analyses in Richlan et al. (2009, 2011). In nine studies, gray matter volumes were assessed. Meta-analysis of these studies showed reduced volumes of gray matter in bilateral temporoparietal and left occipitotemporal regions as well as bilateral regions of the cerebellum. These differences coincided with regions of underactivation in the left fusiform and supramarginal gyri. There was also overlap of cerebellar overactivity.

In contrast, Jednorog et al. (2015) synthesized aMRI data from six countries with 236 participants. The only significant difference was reduced left thalamic volume in poor readers. The lack of replication was striking. In a recent meta-analysis that also included direct and uniform assessments of 293 children and adults with reading disability, Eckert et al. (2016) found reductions in gray matter volumes of the left posterior superior temporal sulcus/left middle temporal region and the left orbitofrontal/pars orbitalis region. Volumes in both these relatively large regions (reflecting an effort at synthesis across studies that vary in regions of interest) predicted reading comprehension skills. These findings were robust in the synthesis of studies and in the direct analysis even when age and sex were controlled. However, when total gray matter volume was controlled, the effects were not statistically different. Eckert et al. (2016) suggested that the findings are reliable, but dependent on participants with low total gray matter volume.

Krafnick, Flowers, Luetje, Napoliello, and Eden (2014) argued that gray matter differences between children and adults who are poor readers were not reliable and largely disappear if reading experience is controlled. When comparing age-matched samples of good and poor readers, researchers found that gray matter was reduced in multiple regions of both hemispheres, most reliably in the left if superior sulcus. However, if the groups were matched for reading level, often presumed to control for reading experience, gray matter differences diminished and only the right precentral gyrus showed reduced gray matter in the group with dyslexia. Krafnick et al. concluded that gray matter differences are largely a product of reading experience and not associated with dyslexia. In contrast, Xia, Hoeft, Zhang, and Shu (2016) used both age- and reading-level matched samples of 48 Chinese children with and without dyslexia. The group with dyslexia (age and reading level matched) showed reductions in gray matter volumes in the left temporoparietal region of the reading network, and in the middle frontal gyrus superior occipital, and reductions in white matter in bilateral occipitoparietal regions often associated with attention. It is clear that volumetric differences in gray matter are reduced as samples increase in size and more sources of variability are measured and controlled.

Cortical Thickness and Gyrification. Another approach used in a few recent studies involves estimation of cortical mantle thickness and gyrification (folding across sulci). The volume of gray matter is a product of cortical surface area and its thickness. Cortical thickness is assessed by measuring the proximal distance between the boundary of gray and white matter and the outer surface of the brain. Surface area is an assessment of depth of cortical sulci and their gyral width, reflecting the amount of cortical folding in a region of interest. While surface area provides an indicator of cortical folding, it does not capture the extent of gyrification, which emerges prenatally and resembles an adult brain by the end of gestation.

Some studies of children and adults with dyslexia reveal reductions in cortical thickness in language and reading areas (Altarelli et al., 2014; Clark et al., 2014; V. J. Williams, Juranek, Cirino, & Fletcher, 2018). Frye et al. (2010) found that differences in surface area and not in thickness in a sample of adult poor readers with a history of dyslexia, but there are questions about how these measures were calculated (V. J. Williams et al., 2018). Welcome, Chiarello, Thompson, and Sowell (2011) found no group differences in cortical thickness among adult university students classified as poor, resilient, and typical readers, but the criteria for poor reading were quite lenient. Ma et al. (2015) found a paradoxical result of increased cortical thickness in left fusiform and supramarginal gyri in a sample of children. However, diagnostic criteria were derived from historical criteria with a focus on who in the sample were remediated. In general, the studies using reliable diagnostic criteria yielded more consistent findings.

For gyrification, even fewer studies have been reported. Im, Raschle, Smith, Grant, and Gaab (2016) found increased gyrification reflecting reduced cortical folding in the left occipitotemporal and left temporoparietal regions in a sample of children with dyslexia as well as a preschool group at genetic risk for dyslexia. V. J. Williams et al. (2018) compared gyrification and thickness in a sample of 31 children with dyslexia defined by poor word reading and 45 age-matched controls. This study was characterized by careful application of semiautomated methods that included manual editing of each slice to ensure accuracy of the gray matter/white matter interface. There was increased gyrification in children with dyslexia in left inferior occipitotemporal and left anterior and superior frontal cortices. In addition, there was reduced thickness in many of the same areas, including bilateral inferior temporal, inferior frontal, and occipitoparietal regions. These differences in thickness are displayed in Plate 2 showing considerable convergence with the functionally derived reading network. Convergence of thinner and more gyrified cortex in the left occipitotemporal region is consistent with the role of the ventral pathways for proficient word reading. That the differences are present in gyrification is consistent with studies of children at risk for dyslexia (see below) due to family history because gyrification is established prenatally. Outside the traditional reading network, thinner cortex was found in right orbitofrontal, left anterior cingulate, left superior parietal, and right medial parietal regions, with some overlap with regions of increased gyrification frontally in regions associated with different attention control.

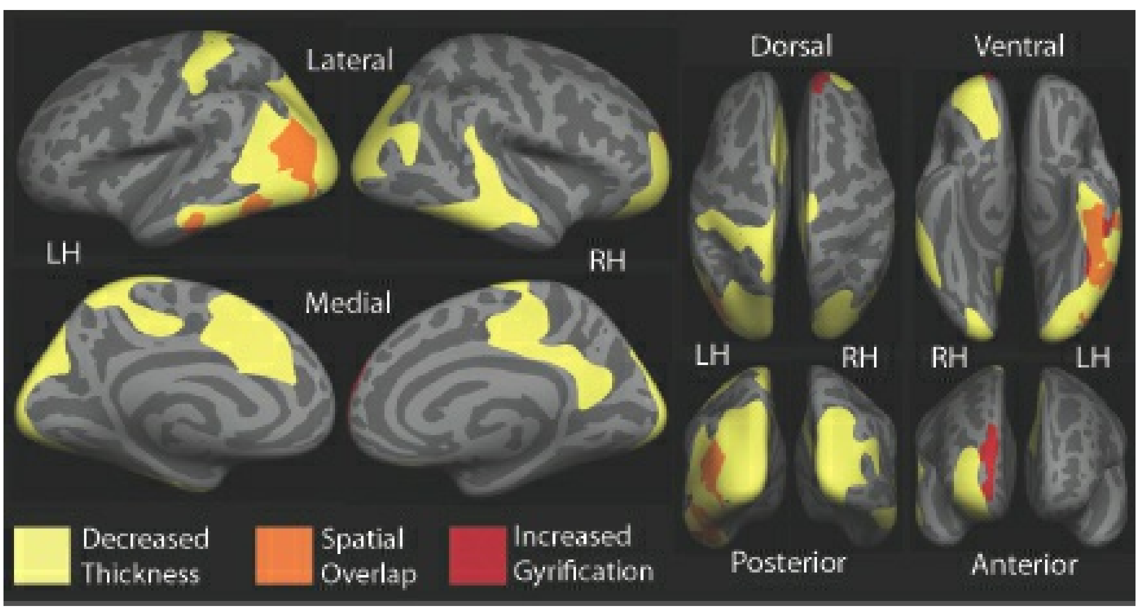

PLATE 2. Areas in which cortical thickness is greater in typical readers and gyrification is reduced in poor readers. Better reading skills predicted thicker cortex in the left hemisphere. Data from Williams et al. (2017). Courtesy Victoria Williams.

White Matter

Fewer studies have examined white matter volumes, focusing instead on studies of connectivity and white matter integrity using DTI. Eckert et al.

(2005) found reduced white matter volumes in the left temporoparietal region. Silani et al. (2005) assessed volumes of gray and white matter in the same participants. In addition to finding reduced gray matter volumes in the left middle temporal gyrus, they found reductions in white matter volumes of the regions involving the arcuate fasciculus. The arcuate fasciculus, also referred to as the superior longitudinal fasciculus, is a long axonal fiber bundle that connects language areas in the posterior temporal and inferior frontal regions (loosely representing Wernicke's area) to the frontal lobes, including Broca's area. As such, it is a pivotal for the connectivity of language and reading networks.

Other studies have examined the corpus callosum, a large fiber bundle connecting the two hemispheres that is essential for interhemispheric communication. These studies did not assess volume, but typically measured the area of the corpus callosum on the midsagittal slice. These studies have produced mixed findings, with some studies reporting differences in the size or shape of different regions (*isthmus and posterior segments*; Duara et al., 1991; Rumsey et al., 1997; Hynd et al., 1995). Conversely, other studies have not found differences in corpus callosum area measures (Schultz et al., 1994).

Interest in area and volumes of white matter pathways has receded because of the advent of DTI, which is especially useful for assessing the integrity of cerebral white matter and brain connectivity. The DTI studies usually involve small samples. A meta-analysis of DTI studies (Vandermosten, Boets, Wouters, & Ghesquière, 2012) identified 10 DTI studies with clearly identified groups of poor readers with dyslexia and controls. The primary finding was reduced integrity in a left temporoparietal region. Fiber tracking showed that this region included connections from the left arcuate fasciculus (AF) and the left superior corona radiata (SCR), which (among other regions) connect the posterior temporal regions and inferior frontal gyrus. This region correlated with reading performance. The meta-analysis also found reduced integrity as indicated by both higher and lower indices in the posterior segment of the corpus callosum. It is interesting that the meta-analysis (as well as many individual studies) did not show strong evidence for reduced integrity in pathways involved in connectivity to the left occipitotemporal region, such as the inferior longitudinal fasciculus (ILF) and the inferior frontooccipital fasciculus (IFOF). As the authors noted, "It can be hypothesized that the reading network is presumably much more complex than two fibers (left AF and SCR) and rather has multiple and bidirectional connections with other white matter tracts, such as the corpus callosum and ventral tracts (i.e., IFOF and ILF)" (p. 1549). At the same time, Christodoulou et al. (2016) found strong associations of fractional anisotropy of the left arcuate fasciculus and reading ability in beginning readers. Studies of white matter integrity remain inconsistent, perhaps epitomized by Christodoulou et al. (2016), who found decreased integrity was differentially associated with some DTI metrics, but not others; paradoxically, decreased FA was associated with stronger phonological awareness in typical readers. Much more research with larger, consistently defined samples is needed.

Neuroimaging Studies of Young At-Risk Children

Implicit throughout this review of functional and structural neuroimaging studies is the question of the extent to which the differences observed in children and adults with dyslexia is a product of reduced experience with print and a consequence of reading failure or actually precede reading difficulties. There are a few longitudinal studies of children imaged as preschoolers with and without family histories of dyslexia. In a series of studies from Gaab's laboratory, Raschle, Chang, and Gaab (2011) conducted voxel-based morphometry on a sample of 20 5-year-olds equally divided into groups with and without a family history of dyslexia. A positive family history was associated with reductions in gray matter volumes in the left occipitotemporal, bilateral temporoparietal, left fusiform gyrus, and right lingual gyrus. There were no differences in frontal or cerebellar volumes. In a functional neuroimaging study of participants from the same cohort, Raschle, Zuk, and Gaab (2012) found bilateral reductions in the occipitotemporal regions and in the left temporoparietal regions during a phonological processing task. In a more recent study, Raschle et al. (2015) evaluated 114 prereaders who had elevated risks of dyslexia because of oral language problems (*n* = 34), often with positive family histories of dyslexia, and low-risk comparison children (*n* = 80). A subset underwent aMRI for voxel-based morphometry. In this subset, those at risk for dyslexia showed reduced gray matter volumes in the left middle temporal, occipitotemporal, and frontal regions. In children with both oral language difficulties and a positive family history, the reductions were greater.

Wang et al. (2016) examined white matter integrity in children from the same cohort followed longitudinally and also in comparison with a cross-sectional group of proficient readers. There was atypical lateralization of the arcuate in prereaders with a positive family history of dyslexia. The longitudinal data showed faster maturation of white matter in the negative family history groups.

In other studies, Vandermosten et al. (2015) performed DTI in a sample of 4- and 5-year-old children, 36 with a positive history of dyslexia and 35 with a negative history. They found reduced integrity of the arcuate fasciculus bilaterally, which was associated with performance on a phonological awareness task. In addition, there was anomalous development of pathways associated with the ventral processing stream, specifically the inferior frontal–occipital tract.

One of the few longitudinal studies (Clark et al., 2014) followed a cohort of Norwegian children at risk and not at risk for dyslexia from 6 years of age until 11 years of age. Identifying small samples with dyslexia and controls at the 11-year-old follow-up, Clark et al. found differences primarily in the primary auditory and visual cortices, and in regions mediating executive functions at the first assessment. Over time, children identified with dyslexia showed overall stable, but reduced, cortical thickness, while typical readers had a thicker cortex that gradually thinned to levels comparable with poor readers, predominantly in areas associated with the reading network. This study was sharply criticized by Kraft et al. (2015) due to the limitations of small sample sizes, failure to control for parent educational level, and failure to relate the reading and spelling data to thickness measures. Kraft et al. reported data from a comparison of 53 prereading children, 25 of whom had family risk for dyslexia and 28 who did not. Cortical thickness was reduced in the left supramarginal gyrus and the left occipitotemporal region in those with positive family histories, consistent with V. J. Williams et al.'s (2018) findings.

Hosseini et al. (2013) evaluated 20 5-year-olds at family risk for dyslexia and 22 controls not at risk. They found a diffuse pattern of cortical thickness anomalies involving left language areas and homologous right hemisphere regions in prereaders with a family history of dyslexia. In studies of children defined as at-risk because of behavioral indices of performance, such as phonological awareness and rapid naming, Yamada et al. (2011) evaluated a small group of 14 5-year-old children at risk and not at risk for reading difficulties. Functional neuroimaging during simple letter recognition tasks revealed less engagement of the occipitotemporal region in at-risk children. After a short period of reading intervention for the at-risk children (and regular kindergarten instruction for the not at-risk group), children at risk engaged the frontal lobes bilaterally, while those not at risk showed activation of the left temporoparietal region. These findings are reminiscent of a MSI study by Simos et al. (2005), which reported bilateral activation of temporoparietal regions in at-risk children at the end of kindergarten (Plate 3) and significantly increased activation of the left hemisphere temporoparietal area after a year of reading instruction when at-risk children responded to the intervention. Those not at risk showed left lateralization of the temporoparietal region at the end of kindergarten that increased with a year of classroom reading instruction.

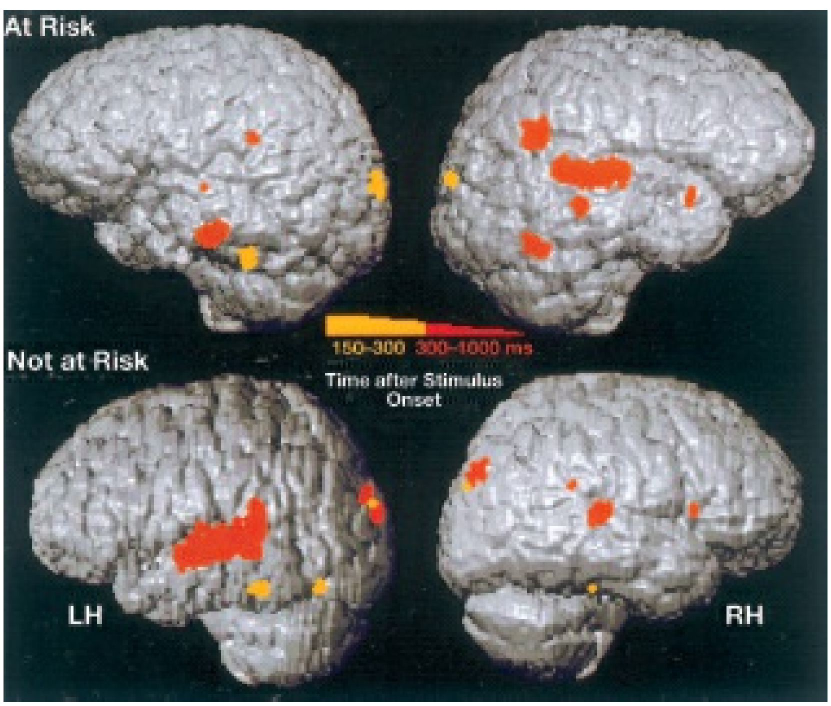

PLATE 3. MSI activation maps from students who at the end of kindergarten were at low risk and high risk for reading problems based on performance of a letter–sound task. Note the absence of activation in the left temporoparietal area in the at-risk child and the clearly lateralized pattern showing

left temporoparietal activation in the not-at-risk child. From Simos et al. (2002, p. 161). Copyright © 2002 Sage Publications. Reprinted by permission.

Saygin et al. (2013) evaluated prereading skills in 40 kindergarten children and performed DTI, focusing on the arcuate fasciculus, inferior longitudinal fasciculus, and the parietal portion of the superior longitudinal fasciculus. Phonological awareness skills were correlated only with the integrity assessment of the left arcuate fasciculus. It is noteworthy that electrophysiological studies with infants show that early anomalies associated with speech processing in the first year of life predicts poor reading at 8 years of age (Molfese, 2000).

The number of studies has increased to a point where a recent meta-analysis addressed specifically structural and functional studies of prereaders (Vandermosten, Hoeft, & Norton, 2016). This study, focused on functional and structural regions identified across studies, identified children at risk because of familial or behavioral risk. The results revealed convergence in relating early deficiencies in phonological processing and development of the left temporoparietal region in those at risk for dyslexia. However, when compared to proficient readers, prereaders used a more widely distributed network for phonological processing that in some studies included the left occipitotemporal region and the cerebellum, with less consistent involvement of right hemisphere regions, especially in those at risk for dyslexia.

Summary: Brain Structure

Samples in individual studies using aMRI are small and heterogeneous. Comparisons across laboratories are also hampered by the use of different neuroimaging methods and data-analytic techniques, leading to difficulties in replicating these findings. Controlling for variation in demographics is also very important (Schultz et al., 1994). These issues, difficult to address in small samples, highlight the value of meta-analysis. Here there are fairly consistent findings indicating reduced volumes of gray matter in the temporoparietal regions. White matter studies consistently implicate the arcuate fasciculus, which connects regions of the dorsal pathway, but have not consistently identified other areas of aberrant connectivity, especially affecting connectivity with the ventral occipitotemporal region. Although it has been suggested that volumetric differences may be a product of experience (Krafnick et al., 2014), other studies using reading-level match designs have shown gray matter differences, especially in the temporal regions (Xia et al., 2016). When prereaders are examined, there is strong evidence that these differences in brain structure, especially in the temporal lobes, exist before formal reading instruction. However, care must be taken in interpreting the results of studies of 5-year-olds. Just because schooling has just begun does not mean that the child has not been exposed to print, taught about letters, or had other exposures to reading that will vary with factors such as parental education and SES, and environmental factors present at conception. As Kraft et al. (2015) suggested, attention to issues like parental education is important because of the influence on the home literacy environment. Yet this is also a statistical conundrum because lower literacy is associated with lower levels of education and families with a history of dyslexia are, on average, less well educated (see Chapter 2). Trying to eliminate this variability leads to nonrepresentative samples of prereaders at risk for dyslexia, so that matching and covarying for characteristics inherent to a group are inappropriate methods of control (Dennis et al., 2009). Moreover, in studies of young children and adult nonreaders, it seems to require relatively little training for the occipitotemporal regions to begin to specialize for print (Brem et al., 2010; Dehaene et al., 2015). Studies of kindergarten children will not fully resolve the question of the role of experience versus preexisting patterns of risk in the development of the neural network that mediates proficient reading.

Genetic Factors

Observations of the cross-generational nature of reading problems stem from the earliest studies of dyslexia (e.g., Hinshelwood, 1917). The risk in the offspring of a parent with a reading disability is approximately four to eight times higher than the risk in the general population depending on whether both parents have a reading and/or spelling problem and the thresholds used to define WLRD (Peterson & Pennington, 2012). Studies of the heritability of dyslexia and other reading disabilities show that the familiality is almost entirely genetic after exposure to schooling (Olson, Keenan, Byrne, & Samuelsson, 2014) but there is also variability due to environmental influences (Petrill et al., 2006). These studies, subject to multiple reviews (Benítez-Burraco, 2010; Elliott & Grigorenko, 2014; Fisher & DeFries, 2002; Grigorenko, 2005; Pennington, 2009; Peterson & Pennington, 2012; Plomin & Kovas, 2005; Scerri & Schulte-Körne, 2010), show a long history of investigations at multiple levels. As Elliott and Grigorenko (2014) summarized, three areas of research converge in demonstrating that dyslexia has a substantial heritable component. These areas involve both family and twin studies of individuals, studies searching for genes involved in dyslexia, and studies examining the role of specific genes (e.g., candidate genes) that congregate within families and may contribute to the manifestation of dyslexia in the general population.

Familial Patterns

Reading problems run in the family. Focused primarily on WLRD, 25–60% of the parents of children who have reading problems also display reading difficulties. The rate is higher in fathers (46%) than in mothers (33%). Children who have parents with reading difficulties are at much higher risk relative to the general population (Snowling & Melby-Lervåg, 2016). In a summary, Scerri and Schulte-Körne (2010) cited research showing that 20–33% of siblings in families with a child with dyslexia developed a reading problem; if one parent had a reading problem, the sibling risk increased to 54–63%; with both parents with a reading problem, the sibling risk was about 75%. Spelling is even more likely to be affected, with 52–62% of siblings experiencing a spelling problem. If ascertainment depends on the parent or school identifying a child as having dyslexia, the rate is closer to 30%. If the child and parent are actually evaluated by research instruments, the rate is significantly higher.

Twin Studies

The limitation of family studies is that environments are also shared between family members. Studies of biologically related family members living together confound genetic and environmental contextual influences, the latter referred to as *shared* influences. Family and twin studies can be used to estimate the amount of variability due to genetic and shared and nonshared environmental contextual factors. Twin studies are especially helpful because these studies can examine the concordance of dyslexia and use deviations from expected concordance rates in identical (monozygotic) and fraternal (dizygotic) twins to assess contextual influences on reading achievement. Monozygotic twins have the same genotype, so the presence of genetic influences leads to the expectation that concordance rates would be much higher in monozygotic than in dizygotic twins, who share, on average, only 50% of the same genotype. If shared family contextual influences are implicated, the concordance of monozygotic and dizygotic twins should be equal (Fisher & DeFries, 2002). Shared environmental influences account for differences between families and could include SES, parental reading practices, and schooling (e.g., same teacher for both twins). Environmental contextual influences could also be nonshared, representing factors that are not genetic and account for differences within families, such as differences in teachers and instructional practices.

Concordance rates are quite high for monozygotic twins (almost always above 80%) relative to dizygotic twins (rarely above 50%). Therefore, these differences in concordance rates presumably reflect genetic effects. The effects observed in twin studies are consistent if groups are formed of good and poor readers or if the entire continuum of reading proficiency is studied. More than any other area, genetic studies embrace a dimensional view of WLRD, with little evidence of genetic effects that are specific to dyslexia.

Statistical methods help separate the variance in reading skills according to genetic and both shared and nonshared environmental contextual influences based on variations in concordance rates (DeFries & Fulker, 1985). In a meta-analysis, 41–74% of the variance in reading achievement and up to 90% of reading-related processes, such as phonological awareness and rapid naming, can be attributed to genetic factors (Grigorenko, 2005). Sex does not appear to influence the association of genetic factors and reading achievement in large-sample twin studies (Elliott & Grigorenko, 2014). Summarizing different large twin studies from Colorado, Florida, Ohio, the

United Kingdom, Scandinavia, and Australia, Olson et al. (2014) reported consistent evidence for heritability estimates for word recognition after grade 1 of 55–83% and shared environmental contextual influences of 2–34%. There was a pronounced trend for lower estimates of heredity prior to the onset of formal reading instruction and for the genetic correlation of reading and heredity estimates to increase with age. Much of the increase is apparent immediately after onset of formal schooling, with a more gradual increase in subsequent years (i.e., the genetic correlation with age is nonlinear). Moreover, heritability estimates are higher at lower ends of reading ability, i.e. where the performance deficit is more apparent (Hawke, Wadsworth, Olson, & DeFries, 2007).

The variation in heritability estimates is likely due to environmental influences and to measurement effects and error (e.g., measuring word-reading accuracy vs. fluency with different tests that have small amounts of unreliability), along with minor amounts of nonshared environmental influences. Age-based variation might reflect different compositions of genes at different age bands. Genetic and environmental influences are not independent. Family history of poor reading may give rise to limited environment–instructional interactions in the home (Olson Forsberg, Gayan, & DeFries, 1999), which is often associated with lower levels of parental education and SES. For example, Friend, DeFries, and Olson (2008) reported higher genetic influences for affected children if their parents were more educated; if less educated, genetic and shared environmental influences were similar. Olson et al. (2014) interpreted these findings as indicating "that children who fail in reading in spite of having highly educated parents (and likely a better environment for learning to read) are more likely to have genetic than environmental constraints on their reading development" (p. 41).

Friend et al. (2009) also examined high and low reading groups and found a similar gene × environment interaction. For stronger readers, genetic influences were higher in association with fewer years of parental education; for poorer readers, genetic influences were lower for children with more years of parental education. Thus, children who learned to read in more disadvantaged backgrounds were more likely to learn to read when genetic influences are stronger. Parental education, home literacy environment, parental reading level, and IQ are all correlated and jointly influenced by genetic and environmental factors that are shared. These effects can be moderated by the quality of the classroom (Taylor, Roehrig, Soden Hensler, Connor, & Schatschneider, 2010). Altogether, the lesson of behavioral genetics is clear: Heritability reflects the influence of the genome when it is unleashed in an environment that is optimized for those who develop reading skills (Elliott & Grigorenko, 2014).

Finally, the genetic impacts are not specific to reading, so that the genetic mechanisms that are associated with one particular LD may also affect other LDs. In an integrative review of genetic research on LDs, Plomin and Kovas (2005) characterized quantitative genetic research on children with LDs as indicating that the effects of some relevant genes are general and not specific to different kinds of LDs. They noted that the genes that had been associated with problems in language, reading, and mathematics are essentially the same genetic constellations that account for normal variation in these domains. In addition, genetic constellations that affect one language or academic domain may affect other components of the disability, which is important for understanding comorbidity (see Chapter 2). Plomin and Kovas relied primarily on large studies such as those from the Colorado group (Olson et al., 1999) and the Twins Early Development Study (TEDS), a study of about 7,500 pairs of twins from the United Kingdom. Summarizing across studies, Plomin and Kovas observed heritability estimates of about 0.6 for reading disability and for reading ability and noted similar findings for analyses based on discrete groups (e.g., dyslexic vs. nondyslexic) as well as studies that analyze reading as a continuous distribution. They noted that "when a gene is found that is associated with a learning disability, the same gene can be expected to be associated with variation in the normal range of ability" (p. 600). Plomin and Kovas also noted the absence of evidence for single-gene defects, stating that "it is generally accepted that genetic influence on common disorders is caused by multiple genes of small effect size rather than a single gene of major effect size" (p. 600). The researchers observed high heritability across a variety of different domains of reading abilities. When examining correlations across different domains of language, reading, and math, Plomin and Kovas found that the domains were often highly correlated, but also observed that the correlations were not perfect, which indicates that there are specific as well as general genes involved in the heritability across these domains. Thus, there are also genetic contributions that are specific to reading, math, and language disorders and development. They concluded by noting that "definitive proof of importance of general genes will come from molecular genetic research that identifies DNA associated with learning disabilities and abilities" (p. 613).

Genomewide Linkage and Association Studies

Family and twin studies help give estimates of the genetic and environmental contributions to individual differences in reading achievement, but do not help identify specific genetic mechanisms underlying heritability. These mechanisms come from molecular genetic studies that attempt to identify specific genetic loci and their contributions to the variation in reading performance. Kornilov and Grigorenko (2018) distinguished two different models of genetic contributions to reading. The most common, which fits the behavioral genetics research, is the *common disorder–common variant* (CDCV) model. Little research supports the idea of a major or single gene defect in relation to WLRD. The identified genetic markers identified so far exercise uniformly small effects on the variability in reading achievement. In the CDCV model, polygenetic inheritance is assumed and multiple genes exert small influences on the expression of a phenotypic trait, like reading. Some variants are expressed in multiple disorders (i.e., generalist genes) and others are specific to the trait (i.e., reading or a related process). Because such a model assumes that liabilities are correlated, they can explain variability in the relation of reading and other domains as well as relations with other cognitive skills. These models also assume dimensionality because the shared risk variants occur in the population at large and are expressed to different degrees depending on reading proficiency and environmental influences.

Kornilov and Grigorenko (2018) contrasted this model with the *common disorder–rare variant* (CDRV) model, where each case emerges because of a single rare variant with a large effect size. These models are consistent with findings from single families with high co-occurrence of reading, spelling, and language problems across multiple family members and across generations. The model also explains why population-based studies find low hereditability of individual genes because such genes do not occur frequently in the population.

There is support for both of these models, which have helped identify potential candidate genes involved in the transmission of dyslexia. With the mapping of the human genome, genomewide association studies are possible, which are not only quite expensive, but also difficult to interpret because of issues about sufficient sample sizes to detect small effects and issues about false positive findings. Nonetheless, at least 15 genomewide association studies have been conducted based on the CDCV, but often guided by findings from the CDRV model (Kornilov & Grigorenko, 2018; Scerri & Schulte-Körne, 2010). In addition, there are studies of specific families from the CDRV model.

Recent reviews converge in identifying nine potential genetic regions related to dyslexia (Plate 4), with strong, replicated evidence for seven of the sites (Benítez-Burraco, 2010; Kornilov & Grigorenko, 2018; Peterson & Pennington, 2012; Scerri & Schulte-Körne, 2010). They have been given numbers from DYX 1 to DYX9. Together, these sites span a large number of genes located on the nine chromosomes: 1p34–p36 (DYX8), 2p15–p16 (DYX3), 3p14.1–q13 (DYX5), 6p22.2 (DYX2), 6q12–q14.1 (DYX4), 11p15.5 (DYX7), 15q21 (DYX1), 18p11 (DYX6), and Xq27 (DYX9). To date, seven candidate genes located within these loci have been supported by at least two molecular–genetic studies carried out in independent samples: *DYX1C1* at DYX1, *DCDC2* and *KIAA0319* at DYX2, *MRPLI9* and *C2ORF3* at DYX3, *ROBO1* at DYX5, and *KIAA0319L* at DYX8. Support is strongest for six genes on sites on chromosomes 15, 6, 2, and 3, with less well-replicated regions on chromosomes 11 and 1. Several of these genes are involved in brain migration and axonal guidance (Peterson & Pennington, 2012; Scerri & Shulte-Körne, 2010), a finding also noted in studies of brain development using animal models (Benítez-Burraco, 2010). For the originally discovered locus on chromosome 6, Grigorenko (2005) reported in her meta-analysis that there were some negative findings, but strong evidence that this locus was involved in dyslexia and specifically for phenotypes identified with assessments of phonological decoding, orthographic coding, single-word reading, and phonemic awareness. There was no evidence that phenotypes defined by rapid naming and spelling were related to this site, but other sites have been related to these skills (Scerri & Schulte-Körne, 2010).

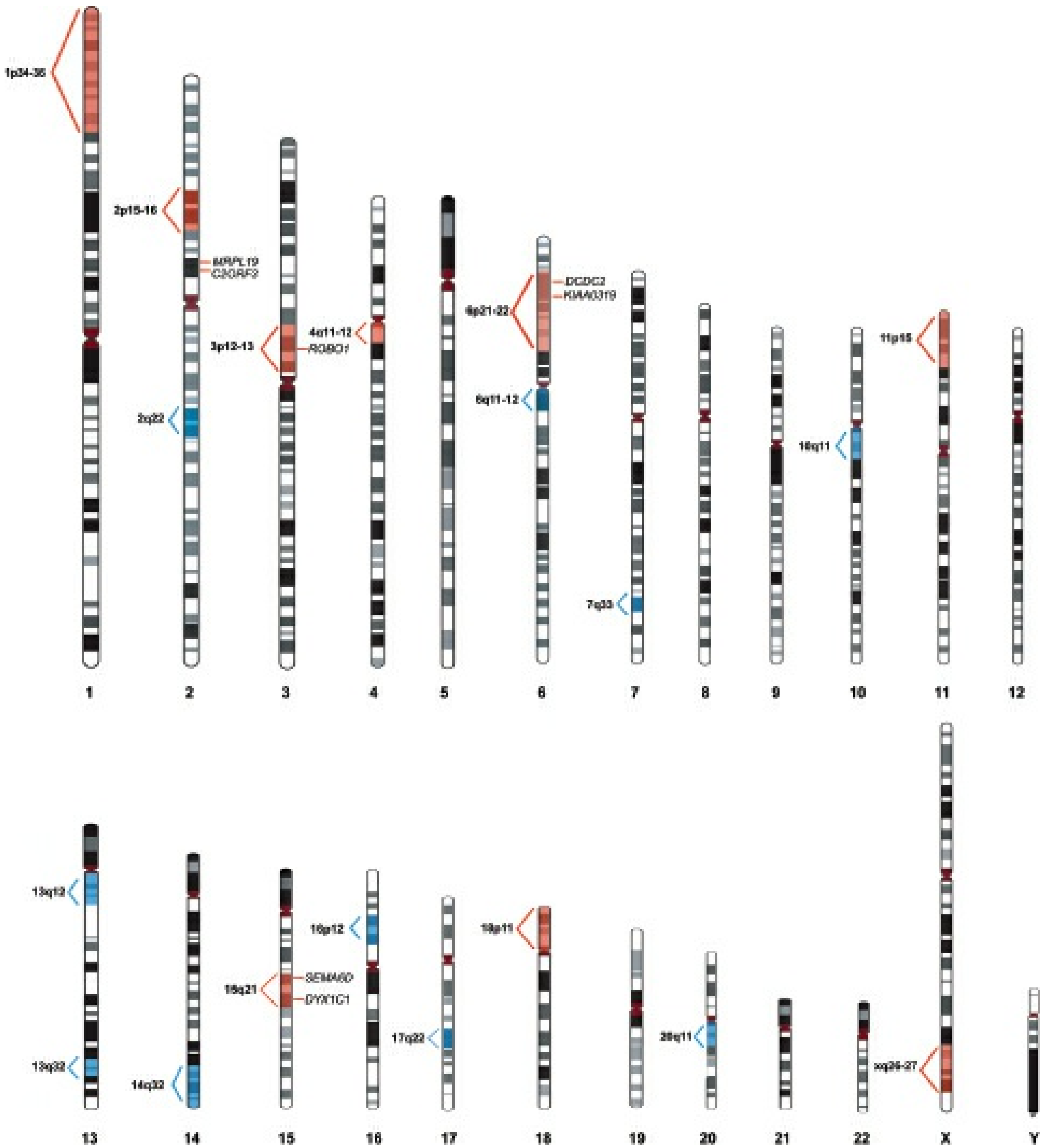

PLATE 4. A summary of identified genetic susceptibility loci in the genome associated with reading disability/dyslexia. From Elliott and Grigorenko (2014, p. 113). Copyright © 2014 Cambridge University Press. Reprinted by permission.

A major problem for genetic studies is the need to compare across different methods for defining the phenotype, an area of controversy for genetic studies of reading disabilities: Does phenotypic variance reflect genetic variance and how much of this variance is due to the phenotype or to measurement error in the tests used to assess the phenotype (Skiba, Landi, Wagner, & Grigorenko, 2011)? The tendency to equate the phenotype to specific tests makes this distinction especially difficult. Nonetheless, the candidate gene findings have been replicated to different degrees across studies for different chromosomes despite this variation in samples, methods of analysis, countries of origin, and definitions of the phenotype.

The search for genetic signature of WLRD continues, with a number of recent whole-genome association studies focusing not only on behavioral (Eicher et al., 2013; Gialluisi et al., 2014; Luciano et al., 2013), but also on brain-based reading-related phenotypes (Eicher et al., 2013; Roeske et al., 2011). The resulting picture, at least at this point, appears to be complex, with the list of candidate genes demonstrably associated with dyslexia continuing to grow, although far from the goal of each novel candidate being replicated in an independent sample (Kornilov & Grigorenko, 2018). The effects of potential candidate genes are small unless pooled into genomewide polygenic scores (Dudbridge, 2013). At this point both the CDCV and CDRV approaches are likely equally valid when applied to complex disorders like WLRD. At the current stage of the field of human genetics/genomics of complex traits, little is known about their joint or specific roles in the etiology of common disorders in general and WLRD in particular. Corresponding theoretical (as well as analytical) frameworks that integrate common and rare variant approaches are scarce.

Summary: Neurobiological Factors

Genetic studies show moderate to high heritability of word-reading skills. Genes that influence reading may have an impact on reading and instructional practices implemented early in development, reflecting genetically driven interactions with the environment that increase estimates of heritability. These estimates do not mean, however, that reading achievement in poverty is due to genetic factors or that genetic factors constrain the effects of intervention, particularly in younger children (Olson et al., 2014). Despite the evidence from familial segregation and twin studies, the effects of individual genes are small, although progress has been made in identifying candidate genes related to dyslexia.

Biology is not destiny. There is clearly a neurobiological substrate underlying WLRD, which is heritable and expressed as individual differences in the readiness of brains to develop the neural network that mediates reading. But this network is malleable, and enriched home literacy and instructional environments make a difference. The lesson of the neurobiological studies is that in order to address the genetic and neural influences, much more intensive and effective instruction is needed. In addition, this intensity must be provided early before the child falls behind in order to provide the explicit exposure to print the brain needs to become specialized for reading. When intense instruction is delayed, the child struggles to develop the capacity for rapid orthographic processing that is essential for reading experience. Adults who have no formal exposure to instruction can be taught reading skills, but their levels of proficiency are not as strong as when exposed at an early age and even the neural system is less well developed. In a child at risk for WLRD, this problem is not just a matter of exposure to print; rather, the child or illiterate must be taught how to extract language from vision (Dehaene, 2009). This depends heavily on the capacity to develop a metacognitive understanding that words have internal structures based on speech, that is, phonological awareness.

Underlying these brain studies is a "chicken-and-egg" question: To what extent are differences in brain structure precursors to the reading difficulties or consequences of a lack of reading experience because of an inability to read words, which, in turn, limits access to print and thus stimulation of neural circuits undergirding reading development. In the end, the chicken-and-egg question is logically unresolvable and really the wrong question. It's not what comes first; the question is how neurobiological and environmental factors act jointly to create a complex cognitive skill like reading that is not evolutionarily derived, but built upon other evolutionary-based systems to permit the development of a reading brain that extracts language from print. Whether the chicken or egg came first is unknowable; what is knowable is how evolutionary systems conspired to create both.

INTERVENTIONS FOR WLRDs

Over the past 30 years, considerable scientific knowledge has accrued regarding understanding and preventing reading difficulties in young children. *Despite this research showing clear efficacy for early interventions and improved outcomes in word reading, fluency, and comprehension in the early grades (K–2), intervention outcomes for remedial efforts—a point at which LDs are typically identified—are much weaker.* Neurobiological studies help us understand why this is the case. One factor is the need to access print in an *explicit* fashion early in reading instruction to build the capacity for rapid orthographic processing in the ventral stream of the neural network mediating fluent reading (Dehaene, 2009) so that the words themselves have significance to the person. This is essentially an experiential factor that stems from the inability to access print early in development when the ventral system, especially the fusiform gyrus, requires considerable exposure to print to develop into an automatic orthographic pattern recognizer that gives immediate access to the meaning of a word. To paraphrase Dehaene (2009), *learning to read is unlocking language from vision; Seidenberg (2017) characterized it as "language at the speed of sight."* There may be issues with the plasticity of the perceptual learning capacity of the ventral system that have not been adequately studied, but the malleability of the dorsal and ventral systems in development and in response to intervention is established. Another factor is that the genetic correlation of reading increases with age, reflecting the homogenizing influence of schooling (Olson et al., 2014). Neither of these potential neurobiological constraints indicates a lack of malleability, but do reflect a need for highly intensive intervention at older ages, where struggling readers begin to appear less responsive to intervention.

The research base demonstrating these findings points to two essential policy issues that are the major messages from these two decades of research (see Chapter 5): (1) *early intervention and an emphasis on beginning reading through explicit, comprehensive core reading instruction (Tier 1) is essential for preventing dyslexia. Core instruction should be supplemented with opportunities to extend instructional time in small groups (Tier 2), with differentiation addressing individual children's weaknesses in reading development.* Such an approach will prevent subsequent emergence of WLRD in many children and likely reduce the number of children who need remedial services in special education or Tier 3; (2) *For students in grade 3 and beyond who do not receive or benefit adequately from early intervention, intense, differentiated instructional approaches are needed (Tier 3)* if the goal is to accelerate the child's or adult's reading proficiency and narrow the gap relative to typically developing peers (see Chapter 5; Spear-Swerling, 2015). Note that in both these scenarios, the student is maintained in core reading instruction as much as possible to maximize the amount of instruction received in reading: *supplement, don't supplant*. As we discuss in Chapter 11, these policy issues have been at the forefront of the messages from the research community, but not consistently understood or implemented. This produces a significant discrepancy between what we know from research and the nature of evidence-based practices implemented in schools.

Empirical Syntheses across a Broad Age Range

Considerable evidence supports the use of specific instructional methods addressing word recognition accuracy and fluency difficulties in poor readers. This research parallels studies conducted at a classroom level demonstrating the importance of *explicit* instruction in the alphabetic principle as a component of any reading program. Two broad meta-analyses synthesized earlier research studies. These meta-analyses focus on children broadly defined with unspecified LDs, but it is likely that most were identified with WLRD.

Students with LDs

One of the first meta-analyses was conducted by Mastropieri and Scruggs (1997), who synthesized interventions involving reading comprehension for K–12 students identified with LDs. Across 68 studies conducted from 1976 to 1996, the mean effect size was 0.98 for unstandardized measures and 0.40 for standardized tests. Variation in effectiveness was not related to grade level. Larger effects emerged from studies employing self-questioning intervention strategies as opposed to text enhancement or providing more general instruction on reading comprehension.

In a comprehensive meta-analysis of intervention studies for students identified broadly with LDs, Swanson, Hoskyn, and Lee (1999) grouped intervention studies into four instructional models: direct instruction, strategy instruction, direct instruction combined with strategy instruction, and other approaches not categorized as direct instruction or strategy instruction. The use of the term “direct instruction” should not be confused with the family of programs described under the Direct Instruction family, examples of which are discussed below. Direct instruction included interventions involving breaking tasks into smaller steps, administering probes, using feedback and diagrams, modeling of skills and behaviors, and related interventions. Strategy instruction included attempts at student collaboration, teacher modeling, reminders to use strategies, multiprocessing instructions, dialogue, and other interventions related to teaching students strategies. Studies that included strategy instruction resulted in larger effect sizes than those that did not (0.84 vs. 0.67); direct instruction produced larger effect sizes than those without direct instruction (0.82 vs. 0.66). Combining direct instruction and strategy instruction yielded larger effect sizes (0.84) compared to direct instruction alone (0.68) or strategy instruction alone (0.72). Note that these effects are in the moderate-to-large range, showing that remedial reading interventions across a variety of different methods improve reading outcomes. Significant effects were observed in word recognition, comprehension, and fluency.

National Reading Panel Report

Commissioned by the National Institute of Child Health and Human Development (NICHD), the National Reading Panel (NRP; NICHD, 2000) conducted a meta-analysis of 96 studies designed to improve phonemic awareness skills. The analysis yielded effect sizes that were in the large range immediately after intervention (0.86) and remained strong over the long term (0.73). There was evidence of generalization to reading and spelling in the moderate range (0.53–0.59). The NRP found that phonemic awareness instruction was most effective when it included a letter component, when instruction focused on one or two types of phonemic manipulations as opposed to multiple types, and when students were taught in small groups.

Programs lasting less than 20 hours were typically more effective than longer programs, with single sessions lasting about 25 minutes. There was little difference in effectiveness between classroom teachers and computers.

Similar findings were apparent in the NRP meta-analysis of data derived from studies of the effectiveness of phonics instruction on a variety of reading outcomes, most often word recognition. Seventy-five studies were screened, and 38 were retained for meta-analysis. The overall effect size of phonics instruction was in the moderate range (0.44). Programs that included phonics instruction were more effective than comparisons that provided either implicit or no phonics instruction. Programs in which phonics was taught more "systematically" were more effective than programs that taught less systematically. Phonics instruction was effective in individual tutorial programs (0.57), small-group programs (0.42), and whole-class programs (0.39). It was much more effective when introduced in kindergarten (0.56) or first grade (0.54), as compared with grades 2–6 (0.27). Phonics instruction was more effective in kindergarten (0.58) and grade 1 (0.74) for students at risk for reading problems. It tended to be less effective for students who were defined with LDs in reading (0.32) and had a negligible effect size in low-achieving readers in grades 2–6. As suspected, word recognition skills were most significantly affected in younger students (effect size = 0.60–0.67), with effects on spelling (0.67) and reading comprehension (0.51). Again, gains were smaller in all domains after grade 1.

The NRP report was widely criticized, mostly because it was misinterpreted as advocacy for phonics methods, which fed into pedagogical disputes. For example, Camilli, Wolfe, and Smith (2006) reanalyzed the results of the NRP report and concluded that the effects of "systematic" phonics instruction (relative to "unsystematic" or no phonics instruction) were overestimated. In a reanalysis of Camilli et al. (2006), Stuebing et al. (2008) showed that Camilli et al. used different methods and asked different questions. They replicated both the Camilli et al. and the NRP conclusions depending on how the results were synthesized. Stuebing et al. argued that the term "systematic" was misinterpreted and that the pedagogical disputes were overly dichotomized, suggesting that the NRP report supported instruction that is (1) sufficiently comprehensive (addresses word recognition, fluency, and comprehension; Mathes et al., 2005); (2) differentiated (organizes instruction according to strengths and weakness in these three domains; Connor, Morrison, Fishman, Schatschneider, & Underwood, 2007); and explicit (the instructor consciously tells the student what he or she needs to learn, organizes the material in advance, provides guided and monitored practice, and checks for maintenance; Vaughn et al., 2012).

At this point in the development of reading interventions, the issue is not whether to provide explicit phonics instruction; rather, the question is how to integrate phonics instruction with instruction on other components central to learning to read. Individuals who argue that the solution to reading difficulties is simply to introduce more phonics instruction in the classroom, without incorporating instruction in other critical reading skills (e.g., fluency, vocabulary, comprehension), are not attending to the NRP findings or the converging scientific evidence. This is true for programs that attempt to enhance the reading abilities of all students in the classroom, as well as programs that attempt to enhance reading in students with LDs.

Prevention of Reading Disabilities

As we discussed in Chapter 5, prevention programs typically include screening assessments to identify students with difficulties in acquiring foundational skills in word recognition and fluency and target interventions to address specific deficits. After screening, students identified as at risk are monitored for progress in reading using CBMs. Intervention programs usually address problems acquiring phonological awareness and word recognition skills, but many programs also address academic needs in the area of vocabulary and comprehension; some involve writing activities. Studies designed to assess the capability of specific approaches to prevent reading disabilities have accumulated because of the increased ability to predict which students will develop such difficulties as they enter and proceed through school (Foorman et al., 2004). Thus, these studies largely target students who are at risk for reading difficulties because of early phonological processing and/or word recognition difficulties. In this section, we distinguish studies that intervene at a classroom level from those that identify students who are at risk and pull them out for intervention. We review only studies that begin in kindergarten or grade 1, but note that preschool interventions are also demonstrably effective (Lonigan et al., 2015), including implementations from an MTSS framework (Lonigan & Phillips, 2016).

Classroom Studies

As we discussed in Chapter 5, the most important component of any intervention program for children with LDs is the core instructional program. In reading, this translates to programs that are explicit, comprehensive, and provide opportunities for practice. The teacher must have data on reading strengths and weaknesses, and the ability to form small, more homogeneous groups based on instructional needs. Because so much has been written about effective classroom instruction, we will refer to some general texts (Adams, 1990). In the context of MTSS, the National Center for Learning Disabilities has an excellent summary of effective core instruction (Hughes & Dexter, 2008). A rubric for evaluating instructional programs was recently developed by Foorman, Smith, and Kosanovich (2017).

Classroom studies either attempt to introduce comprehensive reading programs into the classroom with an accompanying emphasis on professional development, or offer a classroom-level intervention that the teacher provides or directs. Introducing reading curricula into the classroom, with professional development linked explicitly to the curriculum, typically results in improved reading scores for the classroom as a whole, as well as accelerating reading development in students who are at risk for reading difficulties (Snow, Burns, & Griffin, 1998). However, few classroom programs have been empirically validated. As D. Fuchs, L. S. Fuchs, and Vaughn (2014b) noted, Tier I programs often are based on instructional principles derived from empirical research or at least a pedagogical theory. They are aligned with these principles, but are rarely validated through empirical study. We present four examples of empirically validated Tier 1 programs: (1) Direct Instruction, (2) the University of Texas–Houston classroom intervention study, (3) the efficacy of differentiated classroom instruction (Connor & Morrison, 2016), and (4) Peer-Assisted Learning Strategies.

Direct Instruction

Direct Instruction refers to the family of programs representing the methods of classroom and remedial instruction in reading, math, and writing developed by Engelmann and colleagues (e.g., Engelmann, Becker, Hanner, & Johnson, 1978). Direct Instruction includes an extensive professional development component that helps teachers understand the rationale for this approach to reading instruction, lesson plans, methods for error correction, and grouping strategies. The curriculum extends beyond phonics into fluency and comprehension. In the classroom reading program, which is called Reading Mastery, lessons are typically fast-paced and follow a prescribed lesson plan. The lessons usually last 35–45 minutes and contain 12–20 tasks. These methods are based on task-analytic and behavior management systems, but line up with an emphasis on explicit instruction in word recognition and teaching of self-regulation strategies. The programs include opportunities for practice using individualized workbooks that match the content in the group lesson. Like other strong reading programs, there is an emphasis on the use of multiple modalities, including writing, and an articulatory component where students carefully pronounce sounds and words accurately and quickly.

There are several meta-analyses of Direct Instruction that have been carried out over a period of decades (e.g., Adams & Engelmann, 1996; Adams & Carnine, 2003). In the most recent comprehensive meta-analysis, Stockard, Wood, Coughlin, and Khoury (in press) completed a meta-analysis of over 393 articles (328 studies and 3,999 effects) that involved a complete Direct Instruction program that met inclusion criteria for methodological rigor. There were 226 studies of reading. The overall effect size for studies of reading was d = 0.51 (95% confidence interval = 0.44, 0.57). Moderator variables involved publication sources of variability (e.g., publication year); design (e.g., randomized control trial); assessment (e.g., norm-referenced assessment); sample characteristics (e.g., ethnicity, urban), and intervention characteristics (e.g., duration). Most effects were calculated from continuously distributed scores, with about half controlling for pretest data. About one-third of the students were identified with special needs or low reading abilities. Despite substantial variability, the effects were robust across these potential moderators, indicting similar effects across the range of reading ability, sociodemographics, study origin, year of publication, and other moderators. These effects are in the medium range, educationally significant, consistent with previous meta-analyses, and accumulated over a period exceeding 50 years.

More hours of instruction and stronger fidelity to the program were associated with higher effect sizes, with evidence of maintenance of gains. Even in large scale implementations of Direct Instruction as a Tier 1 core reading program, Kame'enui, Fien, and Korgesaar (2013) reported an effect size of d = 0.21 based on 49 studies, 38 of which were done by external evaluators. This effect size is significant, especially in the context of large-scale implementation.

Explicit Instruction in the Classroom

Foorman et al. (1998) contrasted the effects of reading curriculums that varied in the explicitness of instruction in word recognition for at-risk students receiving Title I services in eight schools in grades 1 and 2. The students were taught by one of three approaches: (1) explicit code—a basal curriculum that provided explicit instruction in word recognition, along with instruction in comprehension strategies; (2) embedded code—a phonics program that emphasized the learning of phonics concepts within the context of whole words; and (3) implicit code—a curriculum that stressed (a) contextual reading; (b) responses to literature; (c) writing, spelling, and phonics in context; and (d) integration of reading, writing, listening, and speaking, with no decontextualized instruction in phonics. All students received the same amount of time in the respective programs, and student–teacher ratios were comparable. The teachers received professional development and support for implementation. Each approach was compared to the others as well as a standard instruction condition.

Figure 6.3 shows an example of the results using a word-reading task administered four times during the school year. Across a variety of literacy outcomes, students in the explicit code group improved at a faster rate than students who received implicit code instruction, and had significantly higher scores in word reading, phonological processing, and spelling. A significantly higher percentage of students in the implicit and embedded code groups than in the explicit instruction group showed inadequate improvement in word reading over the year. In addition, Foorman et al. (1998) found that the relation between phonological analysis and word reading was stronger for explicit code students than for implicit code students, suggesting that the effects of explicit instruction on word reading stemmed from its effects on phonological awareness.

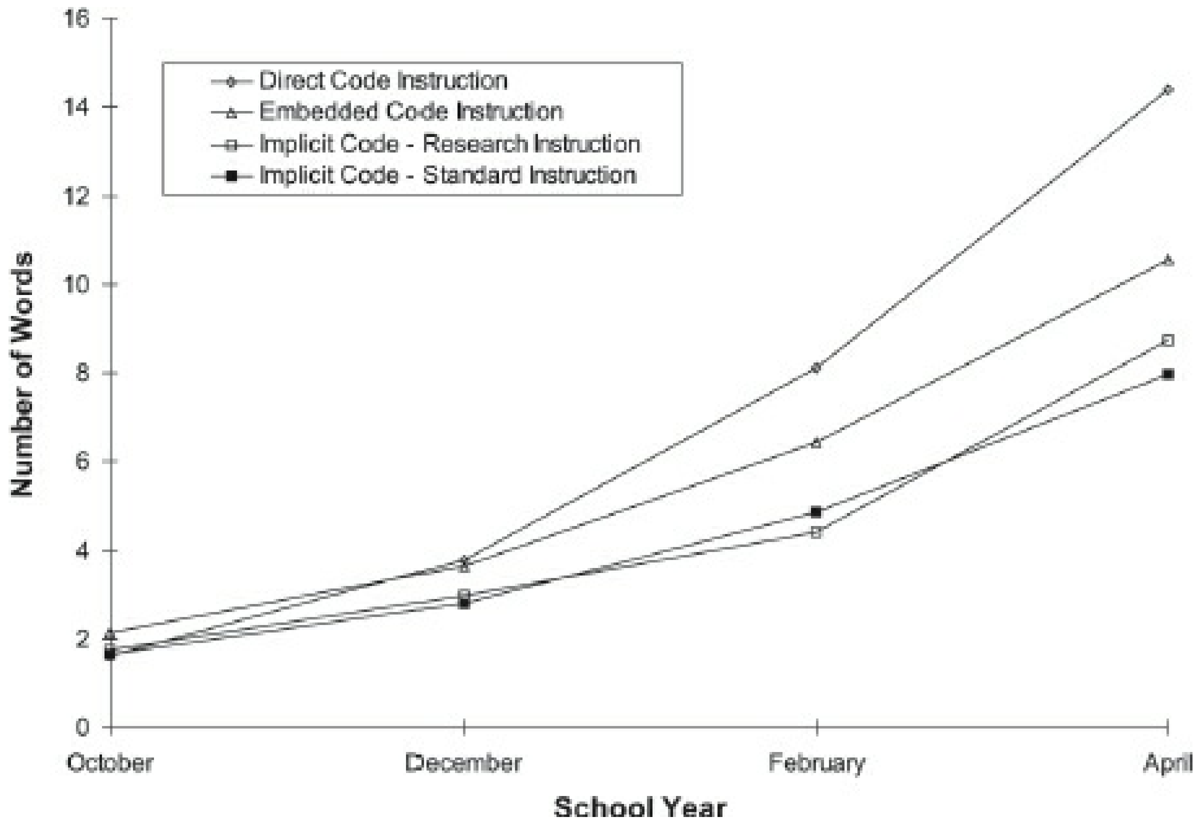

FIGURE 6.3. Growth in word reading raw scores at four time points during the school year by curriculum. Students who participated in the direct (explicit) code condition showed more rapid growth and higher end-of-year performance. From Foorman, Francis, Fletcher, Schatschneider, and Mehta (1998, p. 46). Copyright © 1998 American Psychological Association. Reprinted by permission.

Differentiated Instruction in the Classroom

Foorman et al. (1998) provided correlational evidence showing that the more explicit program yielded greater gains because it was more effective with children who had weaker phonological awareness skills. This may represent an interaction of the child's profile of strengths and weaknesses in reading and different kinds of instruction. In a series of experimental studies, Connor and associates demonstrated clear evidence of interactions of child characteristics and different instructional emphases, supporting the value of differentiated instruction in the classroom (Connor et al., 2007, 2009; Connor & Morrison, 2016). For example, Connor et al. (2007, 2009) helped teachers measure child attributes involving reading decoding and comprehension. They then helped teachers differentiate the amount of instruction in code-based versus meaning-based instruction based on weaknesses in decoding versus comprehension. Across grades 1–3, classrooms in which teachers varied the amount of instruction to address the child's reading profile resulted in better outcomes compared to classrooms in which this assistance was not provided.

Connor and Morrison (2016) described three essential elements of their approach: context, content, and instructional attention management. *Context* refers to the instructional environment, including whole-class, small-group, and individual instruction. In effective classrooms, teachers use flexible grouping strategies and move smoothly from whole-class instruction to smaller groups based on learning needs. Understanding the child's reading strengths and weaknesses allows for the formation of homogeneous learning groups and differentiation of instruction in the classroom. *Content* is the amount of instruction devoted to learning decoding skills versus vocabulary and comprehension. Decoding skills are foundational and involve phonological awareness, phonics, fluency, and spelling. Meaning-based instruction involves vocabulary, comprehension strategy instruction, and listening comprehension. *Instructional attention management* addresses who controls the focus of the instruction: child or teacher.

This three-dimensional model can describe the range of activities in which a teacher might engage the child in whole-class and small-group activities in the classroom. One major finding is that while meaning-based instruction is effective at the whole-class level, code-based instruction is four times more effective if done in small groups led by the teacher or the child. To assess reading strengths and weaknesses, brief, reliable, and easily administered formative assessments are used. The assessment data are used to compute the amount of time the teacher should devote to code and meaning-based instruction in teacher-led versus child-led activities. Lesson plans can be downloaded from the school's literacy curriculum that is aligned with individual student needs. Online professional development and coaching is provided to facilitate the teacher's ability to implement this differentiated instruction.

Peer-Assisted Learning Strategies

An alternative and cost-effective classroom-level intervention has been developed based on collaborative learning (Jenkins & O'Connor, 2003), one of the approaches to small-group instruction also found in the work of Connor and associates. *Collaborative learning* refers to a set of practices in which students work together in pairs or small groups in prescribed ways on structured learning activities. Such activities emerge from a number of the models reviewed above, often integrating cognitive and explicit instructional principles. As a set of practices, cooperative learning has a large empirical base that provides strong support for its use at a classroom level (O'Connor & Jenkins, 2013). This is partly because such practices facilitate classroom management and differentiated instruction delivered in small groups within the classroom. Peer mediation, in which peers collaborate with one another through tutoring, was identified as an effective practice is a recent meta-analysis (Wexler, Reed, Pyle, Mitchell, & Barton, 2015). These authors reported moderate-to-large effects favoring peer-mediated learning, particularly if a peer feedback component was included.

In the reading area, a well-developed form of collaborative learning is represented by Peer-Assisted Learning Strategies (PALS; L. S. Fuchs & D. Fuchs, 2000; D. Fuchs & L. S. Fuchs, 2005). PALS is a classroomwide intervention in which students with stronger academic skills are paired with students who have weaker academic skills for about 30 minutes of instruction three times per week. In each session, students work in highly structured ways on decoding, fluency, and comprehension activities. The distribution of activities across these domains is determined by grade level, and teachers are encouraged to differentiate the reading material and the level of scaffolding for each activity to match the instructional needs of the lower-performing student in each pair. There is an extensive literature on the efficacy of PALS, which has been developed for reading and math and used in research from preschool through secondary school (L. S. Fuchs & D. Fuchs, 2000; D. Fuchs & L. S. Fuchs, 2005).

D. Fuchs and L. S. Fuchs (2005) reported evidence that PALS improved not only word recognition skills, but also reading fluency and comprehension. In summarizing the results of several studies involving PALS in first grade, Denton and Mathes (2003) found that PALS resulted in 69–82% of the poorest readers in the classroom progressing to the average range by the end of the intervention, based on a criterion of word reading above the 25th percentile. Extrapolating this reduction to the total school population indicates that PALS potentially reduces the population base rate for reading difficulties from 25% to 5–6%, results that are similar to those reported by Foorman et al. (1998).

PALS is designed as an adjunct to Tier 1 instruction, but could also be considered a cost-effective method of embedding Tier 2 methods within the classroom. PALS is a Tier 1 intervention if it is does not significantly extend the amount of time the student spends in reading instruction. It is a Tier 2 intervention if it supplements Tier 1 instruction and significantly increases the amount of instructional time in reading.

Tutorial Studies

In the next section, we review studies that are largely independent of classroom-level interventions. These studies typically rely on a one-to-one or small-group intervention model in which at-risk students are pulled out of the classroom for supplementary instruction. Although initial studies focused on the one-to-one format, more recent implementations incorporate small groups of two to five students. We begin with two meta-analyses, then turn our attention to a highly visible program and examine specific studies of empirically validated practices.

Meta-Analyses of Tier 2 Interventions

Wanzek et al. (2016) identified 72 articles addressing the effect of Tier 2 reading interventions in grades K–3. They reported moderate effects on standardized (0.54) and nonstandardized (0.62) measures of word reading and phonological awareness (0.54); smaller effects on standardized language and reading comprehension measures (0.36); and larger effects on nonstandardized language and reading comprehension measures (1.02). Effect sizes are usually larger on nonstandardized measures developed by the interventionist and more aligned with the taught curriculum, where greater transfer would be expected. There were no moderating effects of group size (1:1 to small groups of 2–5), grade level, instructional specialist, number of hours of intervention, or type of intervention (foundational vs.

multicomponent). However, the bulk of the interventions were in kindergarten and grade 1, with a clear trend for larger effects in earlier grades. Intervals ranged from 10 to 60 minutes, but most were 20–30 minutes. Group sizes were almost equal for 1:1 versus small group. The authors concluded:

> Overall, the research demonstrated moderate, positive effects of less extensive interventions on both standardized and not-standardized measures of foundational reading skills such as phonemic awareness, decoding, word identification, decoding fluency, word identification fluency, and text reading fluency. Smaller effects were noted for less extensive interventions on standardized measures of language/comprehension, with the majority of the standardized measures assessing reading comprehension. (p. 567)

The results of this meta-analysis were consistent with Wanzek and Vaughn (2007), who synthesized 18 studies of more intensive intervention lasting at least 100 sessions for students in grades K–5. Although some may consider 100 hours to be more like a Tier 3 intervention, this amounts to daily 30-minute lessons over 40 weeks, not different from what has been implemented in many supplemental Tier 2 intervention studies (e.g., Mathes et al., 2005; Vellutino et al., 2006). The meta-analysis reported moderate effects on foundational skills and reading comprehension. Effects were higher for interventions serving smaller groups (but not necessarily 1:1) and for interventions at kindergarten and grade 1. There were no differences in outcomes for highly scripted versus less scripted interventions, paralleling Mathes et al. (2005).

Reading Recovery

A popular early intervention program for first-grade students reading in the lower 20% of their classes is Reading Recovery (RR; Clay, 1993). This intervention provides daily, individual 30-minute lessons to first graders who are identified as being at risk on the basis of a survey of reading skills. A complete RR program includes 20 weeks of lessons, although the actual duration of the program varies from student to student. The RR program stresses that basic decoding and phonics skills should be taught in the context of authentic reading and writing activities and encourages the use of multiple teaching strategies (use of context clues, word attack, etc.) to identify words, rather than focusing on only one strategy, such as “sounding out” words. The RR teacher is responsible for selecting texts for each individual student to challenge, but not frustrate, the student. A major emphasis is on the teacher's observational skills and judgment.

Shanahan and Barr (1995) provided a comprehensive review of the effectiveness studies conducted to date with RR, reporting that the program results in substantial gains in reading for approximately 70% of participating students. However, they noted that many of the studies reviewed were methodologically flawed. Another meta-analysis, which also found that RR was effective for many grade 1 students (D'Agostino & Murphy, 2004), disaggregated RR outcomes by whether the outcomes involved standardized achievement tests or the Observation Survey, an assessment developed by the program authors (Clay, 2002) and highly aligned with the RR curriculum. It also separated results for students who successfully completed RR (i.e., met program criteria) versus those who were unsuccessful or left the program before receiving 20 lessons and according to the methodological rigor of the studies. When the comparison group was composed of low-achieving students, average effect sizes on standardized achievement tests for all discontinued and not discontinued students were in the small range (0.32) and higher for discontinued (0.48) than not discontinued (–0.34) students. This finding was consistent with that of Elbaum, Vaughn, Hughes, and Moody (2000), who reported that RR was less effective for students with more severe reading problems.

More recent studies continue to show that RR is effective for many students, but much less effective for those at risk for WLRD and who need more explicit instruction in the alphabetic principle. The What Works Clearinghouse (2013), a component of the Institute for Educational Sciences responsible for evaluating the effectiveness of educational interventions, identified 79 studies that exposed students to RR. Selecting only five studies that involved rigorous experimental designs, with samples of only 74–227 students, the effect size estimates for general reading achievement was 0.75. This is large, with strong effects reported for word reading, fluency, and comprehension. These effect sizes are much higher than those reported by D'Agostino and Murphy (2004).

In a subsequent large-scale evaluation study of several thousand students followed for up to 4 years, May et al. (2015) reported an effect size on overall reading achievement of 0.69, which is in the upper moderate range. Over 4 years of implementation with 6,888 students (May, Sirinides, Gray, & Goldsworthy, 2016), the effect size on the Iowa total reading score was 0.48 and 0.99 for the Observation Survey. Unfortunately, there were issues with the completion rate, fidelity of implementation, and the dropping of students who were not successful from the program (Chapman & Tunmer, 2016). Although completion rates and fidelity improved over the 4 years of the study, differences between those exposed and not exposed to RR were not statistically significant after grade 1 in a relatively small sample powered to detect differences of about 0.33. Most estimates of effect size fell well below these detectable thresholds.

There are extensive criticisms of RR, especially from New Zealand where the program has seen widespread implementation (Chapman, Greaney, & Tunmer, 2015). Concerns about the effectiveness of RR revolve around two issues: (1) whether RR is successful with the lowest-performing students and (2) whether RR is cost-effective. In terms of the first concern, RR has typically targeted students who perform in the lowest 20% of their classes. The actual performance level of participants varies from school to school. Although research conducted by the developers of RR continues to indicate efficacy for about 70% of the students, its reported effects are much weaker when students who do not meet the program's exit criteria are included in the analyses of outcomes. This seems especially true for students with significant phonological awareness difficulties. Gains for the poorest readers were often minimal, which may be related to the need for more explicit instruction in decoding (Chapman et al., 2015; Elbaum et al., 2000).

Some adaptations of the RR program have been shown to improve efficacy. Tunmer, Chapman, and Prochnow (2003) modified an RR program in New Zealand to include phonological awareness and explicit phonics instruction, implementing it with economically disadvantaged minority students. Comparing the modified program in seven schools with a historical control cohort from the same schools revealed that students who received the modified program scored higher than the historical controls on all phonological awareness and reading measures, including standardized measures of reading achievement and measures like those employed in RR. These gains persisted through grade 2. The Responsive intervention used in Mathes et al. (2005; see Chapter 5) had many features of RR, but with more explicitness of the alphabetic principal and the group size, with excellent results.

The second issue with RR involves its cost-effectiveness (Hiebert, 1994). The professional development component is expensive and, because RR requires one-on-one tutoring, many schools find it difficult to implement on a long-term basis. The question, however, is whether any reading intervention in elementary schools needs to be provided on a 1:1 basis. Versions of RR administered in small groups have not been shown to be less efficacious. For example, Iversen, Tunmer, and Chapman (2005) developed a version of RR for small groups, observing no differences in outcomes for students taught in 1:1 and 1:2 formats. Mathes et al. (2005) taught in group sizes of 1:3.

Other Tutorial Studies

In this next section, we select studies that demonstrate experimentally important principles of early reading instruction.

Timing of Instruction. The meta-analyses tend to show mixed results for the timing of instructional delivery. We have implied throughout this book that early intervention is more effective than remediation after children fall behind. Two studies show strong evidence in support of this hypothesis. Connor et al. (2013) recruited and followed a large group of students from grades 1 to 3. At each grade, teachers were randomized to provide either Connor's differentiated instructional methods described above or to offer a comparison condition where math instruction was provided. In the design, students could receive 1, 2, or 3 years of differentiated instruction. Students who received 3 years of instructions showed the strongest reading skills. First-grade instruction was effective, but results were not sustained. Even so, without first-grade instruction, third-grade outcomes were poorer. Thus, first-grade instruction was necessary, but not sufficient.

Lovett et al. (2017) examined the impact of grade level of implementation on early intervention outcomes. A multicomponent reading intervention program with established effectiveness was provided in a small-group format to students in grades 1–3 at risk for or with identified WLRDs. Each student received about 125 hours of instruction over 7 months and was followed for

1–3 years depending on entry to the program. Collapsing across grade and 14 outcomes, this quasi-experiment showed that students in intervention outperformed comparison students not exposed to the intervention by an average effect size of 0.99. On measures of word reading, students who received intervention in grades 1 and 2 showed about twice as much growth relative to the comparison group, who had been wait-listed to and received intervention in grade 3. This acceleration was even larger in a 1-year postintervention follow-up. Moderate-to-large effects were observed on three standardized measures of reading comprehension that ranged from 0.6 to 0.90, which did not interact with grade level. Although meta-analytic studies have not found that IQ predicted intervention outcomes, this study found interactions of lower IQ scores and outcomes, such that students who entered the study with lower IQ scores showed more growth than students with higher IQ scores who entered the study. Although not what might be expected based on common conceptions of IQ scores as aptitude tests, IQ is correlated with reading, so this finding does make sense. The intervention was more effective with students who were more severely impaired.

Importance of the Alphabetic Principle. Many prevention studies have shown that programs providing explicit instruction in the alphabetic principle are more effective than programs that provide implicit or incidental instruction (Torgesen, 2000; Torgesen et al., 1999; Vellutino et al., 1996; Vellutino, Scanlon, & Jaccard, 2003; Vellutino et al., 2006). We highlight one experimental study (Denton et al., 2014) because it used Guided Reading (Fountas & Pinnell, 1996), a widely implemented reading program used as a supplement to language arts instruction or as a tutorial program in the elementary grades. This approach provides extended time reading texts leveled for difficulty facilitated by a teacher who helps the child become engaged in reading. It does not emphasize explicit instruction in the alphabetic principle and like RR, teaches a multiple cuing strategy when a child cannot recognize a word that may involve guessing from context.

Denton et al. (2014) evaluated first-grade tutorial interventions that compared a Guided Reading intervention with an intervention defined as "explicit." First graders at risk for reading difficulties based on knowledge of letter sounds, phonological awareness, and an absence of sight-word reading skills, supplemented by teacher nomination, were randomly assigned to

Guided Reading, Explicit Instruction, or a business-as-usual comparison group (often a school-based implementation of Guided Reading). Each child received 45 minutes of instruction four times weekly for 23–25 weeks. Guided Reading as implemented by Denton et al. taught multiple cuing systems for identifying words and had little decontextualized phonics instruction. It used texts leveled for difficulty, with considerable time spent in extended reading and writing activities and text discussion. Explicit Instruction used only one strategy for word identification (sounding out); had daily phonics instruction as part of the lesson plan, which was often decontextualized; used decodable and authentic texts; and employed explicit instruction in comprehension strategies.

The results revealed that both intervention groups performed better than the comparison group. Although the differences in word reading, fluency, and comprehension were not statistically significant, effect sizes favored Explicit Instruction. An analysis of expected growth based on initial level of performance showed that the children in the Explicit Instruction group showed greater accelerated development of decoding, fluency, and comprehension skills. Notably, Denton et al. (2015) provided a Tier 3 intervention to students who did not respond adequately to first-grade instruction and thereby accelerated gains in about half the inadequate responders relative to the business-as-usual group. The best predictor of growth was the level of phonological awareness skills at baseline, which has been reported to be the best single predictor of inadequate response in a series of studies (Fletcher et al., 2011; Denton et al., 2013, 2014; Miciak et al., 2014b). Hulme et al. (2012) found that the positive effects of an early reading intervention for kindergarten children with weak oral language skills on word-level skills were fully mediated by gains in phonological awareness and letter sound knowledge.

Group Size. In a meta-analysis, Elbaum et al. (2000) found that larger groupings of one teacher to three students were just as effective as 1:1 groupings across a range of interventions. Vaughn et al. (2003) systematically manipulated group size to compare interventions delivered 1:1, 1:3, and 1:10. Across a variety of reading assessments involving word recognition, fluency, and comprehension, outcomes were comparable for 1:3 and 1:1 interventions, and both of these were better than interventions in group sizes of 1:10.

Wanzek and Vaughn (2007) and Suggate (2016) also did not find differences between 1:1 and small-group instruction, although smaller groups were associated with better outcomes. These findings support conclusions reached by the NRP (NICHD, 2000) that small-group instruction (1:3) is likely sufficient for many young struggling readers at Tier 2.

Unit of Instruction. It is often believed that learning the alphabetic principle requires a structured phonics program that teaches letter–sound correspondences in a rule-based format. However, there is little evidence supporting one method of teaching phonics over another. The key is helping the child access internal units of words and connecting them to speech sounds. Thus, a variety of synthetic phonics approaches are effective. These approaches teach letter sounds as a set of rules that must be blended from single letter to larger units of speech. However, approaches based on analytic phonics, in which isolated letter sounds are less emphasized, and the child is encouraged to recognize phoneme–grapheme correspondences across words, as in word families, is also effective. This finding was clearly apparent in Mathes et al. (2005) and Torgesen et al. (2001).

In addition, a meta-analysis of teaching strategies based on morphological knowledge (derivational, inflectional, and compounding of words) found 17 studies addressing morphological awareness (Goodwin & Ahn, 2010). Across studies that spanned a wide range of grades and reader types, morphological awareness was associated with positive gains in phonological awareness (0.49), morphological awareness (0.40), and vocabulary (0.40), with lower effects on reading comprehension (0.24), and spelling (0.20). For some children, one particular approach to teaching phonics may not be effective. If progress is monitored, alternative approaches may be tailored to student needs to increase success. The importance of teaching multiple units of code was clearly evident in the nature of the interventions in Lovett et al. (2017), which used a basic phonological program from Direct Instruction approaches, but added morphological instruction and other approaches to extend the unit of instruction.

Persistence of Effects. It is common to question whether the effects of early intervention persist. In a review of 71 studies with intervention and comparison groups, but ranging in grade, Suggate (2016) found posttest effect sizes of 0.37 that dropped to 0.22 after an average of 11 months following intervention (outcomes were collapsed across learner types). For students identified as at risk, low readers, or reading-disabled, average immediate effect sizes were 0.54, 0.40, and 0.37, respectively. These effects dropped to 0.35, 0.32, and 0.30, respectively, all in the small-to-moderate range. There were no differences by grade level, except for more of a drop-off in kindergarten and grade 1. Group size (1:1 vs. small group) did not moderate the effects. More maintenance was apparent with stronger experimental designs and higher dosage. As Lovett et al. (2017) suggested, Suggate's results differ from many individual studies and may be related to how effect sizes were computed. Suggate's effect sizes may be lower because younger comparison children are more likely in an active treatment group. There were negative correlations of grade and reading scores in the control group.

Suggate excluded Blachman et al. (2014) because the length of follow-up (10 years) was an "outlier." This study is worth considering because of the length of the follow-up interval. The follow-up study involved the cohort who received intervention in Blachman et al. (2004). The sample was second- and third-graders with poor word recognition ability randomized to an individualized tutorial or business-as-usual groups. The individualized tutoring intervention lasted 8 months (average of 105 hours), with an emphasis on explicit instruction in phonological and orthographic connections in words as well as text-based reading, of narrative and expository text to enhance fluency, comprehension, and engagement, along with other writing activities and games. These students also participated in an fMRI study before and after intervention (see B. A. Shaywitz et al., 2004).

Figure 6.4 shows representative scores on measures of word-reading accuracy, comprehension, and fluency of text reading at baseline, end of the school year (about 8 months), 1 year after completion of the intervention (20 months from baseline), and over a decade later when the majority of students were no longer in secondary school (from Blachman et al., 2014). After the intervention (8 months) across multiple outcomes, students who received the intervention had greater gains in word recognition, fluency, comprehension, and spelling compared to students who received the interventions provided in their schools. These gains were maintained in a 1-year follow-up. The effect sizes were generally in the moderate-to-large range across reading domains, ranging on standardized tests from 0.55 for reading comprehension to 1.69 for word recognition. Interestingly, in the fMRI study, B. A. Shaywitz et al. (2004) found that prior to the intervention, students with reading difficulties exhibited much less activation of brain areas in the left hemisphere commonly associated with reading difficulties. After the intervention, students who received the experimental intervention showed greater activation of bilateral inferior frontal gyri, the left superior temporal sulcus, and the occipitotemporal region (middle and inferior temporal gyri, middle occipital gyrus). B. A. Shaywitz et al. interpreted these results as showing normalization of left occipitotemporal regions associated with efficiency in reading, but noted compensatory changes involving the right frontal region.

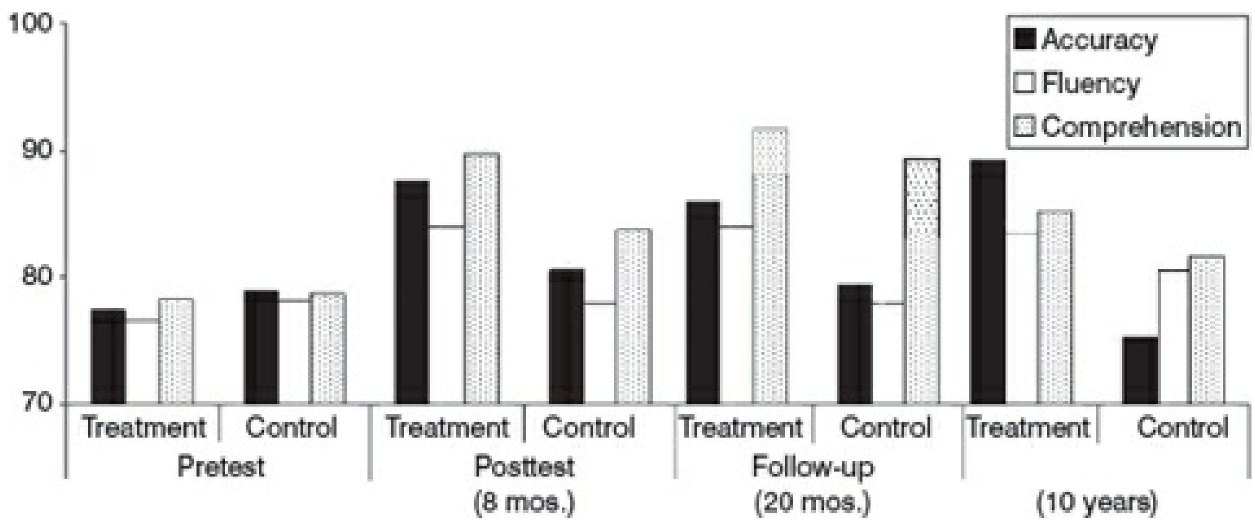

FIGURE 6.4. Intervention results at pretest, 8 months, 20 months, and 10 years for treated and control subjects on the accuracy, fluency, and comprehension measures of the Gray Oral Reading Test (GORT). Clear differences in treatment and controls are apparent. Data from Blachman et al. (2004, 2014).

In the 10-year follow-up, Blachman et al. (2014) identified 84% ($n = 69$) of the original participants (Figure 6.4). The treatment group continued to outperform the comparison group on measures of word reading (0.53–0.62) with smaller, but meaningful, effects on fluency and spelling. Comprehension differences were minimal. Although the gains diminished over time, the main effects on word reading were still statistically significant after 10 years with no intervention from the researchers. There are few such long-term follow-up studies.

Summary: Prevention Studies

The classroom and tutorial studies reviewed in this section show that early intervention may reduce the number of students at risk for reading difficulties, including those who might eventually be characterized with LDs in reading as well as those who are economically disadvantaged and may be poorly prepared to read. Intervention studies that address the bottom 10–25% of the student population may significantly reduce the number of at-risk students. Both classroom and small-group tutorial programs are effective (i.e., successful intervention programs do not require 1:1 tutoring). In addition, the most effective programs are comprehensive, integrated with distributed instructional emphasis on the alphabetic principle, teaching for meaning, and opportunities for practice. Earlier intervention is associated with greater gains and the results may persist.

Coupled with the results from studies that layered classroom and tutorial interventions reviewed in Chapter 5 as well as preschool interventions (Lonigan et al., 2015), this body of research provides strong support for the efficacy of early intervention. When this layering occurred in an MTSS framework, such as in the provision of tutorial instruction for at-risk students from classrooms with a well-structured, classwide, dyadic intervention such as PALS (which provides for a high level of structured practice on critical curricular content); or where the core reading program is already strong and supplemented with supplementary Tier 2 intervention, the number of at-risk students appears to fall below 2% in some studies. Moreover, outcome studies show that these changes are effective through grade 5 (and longer based on Blachman et al., 2014) and that domains involving word recognition, fluency, and comprehension are positively impacted.

Reading Remediation Studies

Empirical Syntheses

It is difficult to bring students with WLRD up to grade level if the intervention begins after grade 2. In the previous edition (Fletcher et al., 2007), we indicated that remedial studies tend to yield effect sizes that are comparable to those of early intervention studies. More recent and better controlled studies show that effect sizes are generally much lower than previous studies had suggested, although there is variation if the program is sufficiently intense. The fundamental problem is that student access to print is considerably delayed, reducing print exposure and reading experience—essential for building fluency (Torgesen et al., 2001) and organizing the ventral stream of the neural network as an orthographic pattern analyzer, which provides immediate access to meaning (Dehaene, 2009).

100 Years of Reading Intervention Research

This pattern of diminishing effect sizes is clearly apparent in a series of meta-analyses of the effects of reading intervention for students in grades 4–12 published by Scammacca and colleagues. In the first study, Scammacca et al. (2007) synthesized 31 studies from 1980 to 2004, mostly involving research conducted on grades 4–8. Collapsing across outcomes involving decoding, fluency, and comprehension, the average effect size was 0.95, indicating large differences between adolescents exposed and not exposed to an intervention in a randomized or quasi-experimental study. The effects on standardized measures was, however, much smaller (0.42), but still in the moderate range. Interventions focused on comprehension strategies and vocabulary, and multicomponent programs, were more effective than word-study interventions. Fluency interventions yielded negligible effects, but study sample sizes were small for the word-study and fluency interventions. Not surprisingly, most interventions focused on comprehension. In examining outcomes for students specifically identified with LDs, reading comprehension gains were similar for students with and without LDs, paralleling results from Swanson et al. (1999) reviewed above. Scammacca et al. noted that it was difficult to determine the extent to which struggling readers were closing the gap relative to typically developing adolescents.

In a follow-up, Scammacca, Roberts, Vaughn, and Stuebing (2015) extended the meta-analysis to 2011. Across 82 unduplicated effect sizes, the mean effect size was 0.49, much lower than in the first analysis. In addition, the mean effect size for standardized tests dropped from 0.42 to 0.21. Comparing studies from 1980 to 2004 with those from 2005 to 2011 revealed statistically significant evidence of a decline in effect sizes related to increased use of standardized outcome measures, more rigorous research design, and improvements in the business-as-usual comparison group's performance, implying better school practices.

Partly to address the issue of declining effect sizes, Scammacca et al. (2016) attempted to synthesize 100 years of reading intervention research, encompassing studies from 1914 to 2014. The review was largely narrative because studies before 1980 were not presented or conducted using research designs compatible with empirical synthesis. Specifically, outcome pre-1980 results were usually reported using grade or age-equivalent scores, which can be averaged.

The effectiveness of reading interventions can be demonstrated across the decades involved in this research. However, there is a clear tendency for increasing rigor in the designs. Between 2010 and 2014, the number of studies exceeded any single decade, with average sample sizes three times larger than the averages during 2000–2009. Half the studies include standardized outcome measures and the focus expanded to include increasingly large numbers of students identified as struggling readers but not with LDs. In many instances, comparisons were made to groups of struggling readers exposed to alternative active interventions (e.g., Morris et al., 2012). The number of randomized controlled trials dramatically increased. As a result of this increased rigor, Scammacca et al. concluded, "experimental and quasi-experimental studies from the 1980s through the 2010s produced effect sizes that have declined sharply and consistently over the decades" (p. 781). Shown in Figure 6.5, the decline is especially apparent for reading comprehension measures, where Scammacca et al. (2015) found effect sizes of 0.19 for standardized tests.

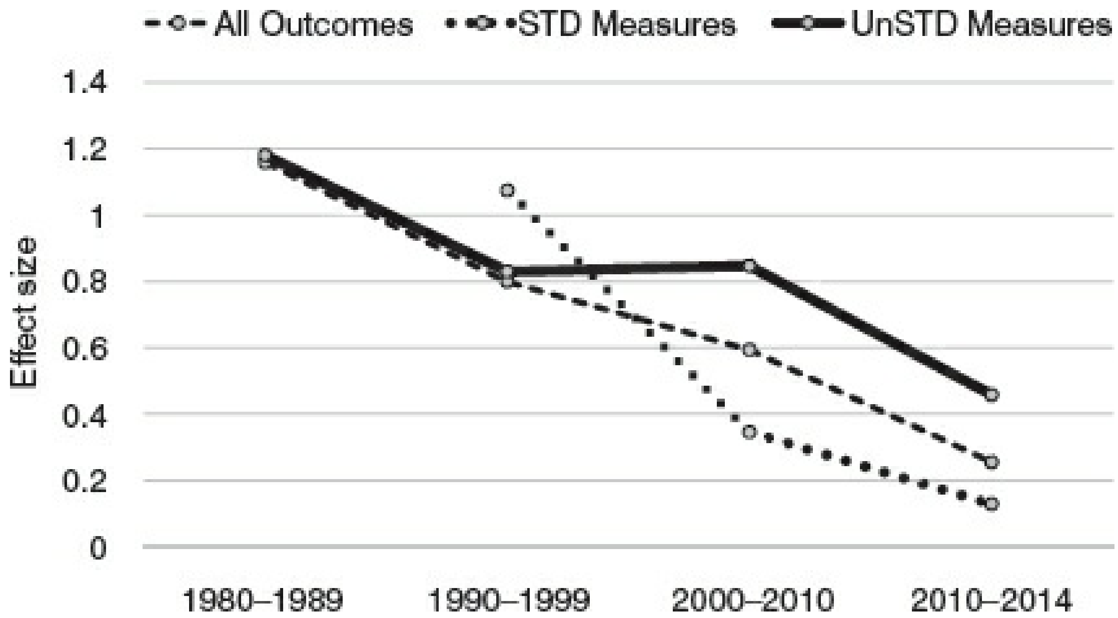

FIGURE 6.5. Declines in effect size estimates from intervention studies spanning 1980–2014 for all reading measures, standardized reading tests, and unstandardized tests tied to the curriculum. Standardized tests usually show lower effect sizes than unstandardized tests, but the overall pattern in this meta-analysis is for a decline in effect size estimates. Data from Scammacca et al. (2015).

Unfortunately, as the more recent meta-analyses demonstrate, studies investigating the efficacy of interventions for improving literacy outcomes with older readers (grades 4–9) with reading difficulties often yield findings with no or low impact for reading comprehension (James-Burdumy et al., 2012; Kemple et al., 2008). In the next sections, we highlight interventions proposed to be highly effective for remediation of students with word-level difficulties and then turn to specific intervention studies that demonstrate more robust effects. The overriding message from these studies is that remedial interventions need to be intense, explicit, long in duration, and responsive to progress during the intervention, such that programs are adjusted in response to progress-monitoring data to increase the students' chance for success.

Multisensory Methods

Historically, prominent remediation approaches used with people with WLRD were used to develop spelling and writing skills as well as reading skills. An early example of this type of method was the Fernald approach (Fernald, 1943), which incorporated principles of language experience and whole-word (not whole-language) instruction in the teaching format. In essence, the reading material to be learned was provided by the students through the dictation of their own stories. The Fernald approach emphasized learning words as wholes and discouraged teaching students how to "sound out" new words. Given what is now known about the importance of decoding skills in the learning-to-read process, especially the importance of "sounding out" words, it is not surprising that the Fernald method was not substantiated by research evidence (Myers, 1978).

Other programs considered "multisensory" were derived from the early work of Samuel and June Orton under the general rubric of "Orton–Gillingham" (O-G) approaches. Early versions of these programs emphasized the need for instruction to all sensory modalities. These approaches required the student to learn associations between letters and sounds. Students were taught to see a letter (visual), hear its sound (auditory), say its sound (auditory), trace the letter (tactile), and write the letter (kinesthetic). Words mastered were eventually inserted into sentences and passages to promote text reading and reading comprehension. There was an emphasis on understanding the structure of language and sounding out words.

These early efforts were reformulated by Anna Gillingham and Betsy Stillman in the 1960s and have continued to evolve. Many of the remedial approaches reviewed in this chapter, including approaches used in research by Blachman, Berninger, Wolf, and others, which emphasize the importance of explicitly and systematically teaching students about the structure of language, reflect the influence of these earlier remedial approaches (Moats & Farrell, 1999). Similarly, commercial programs such as the Lindamood Sequencing Program for Reading, Spelling, and Speech (Lindamood & Lindamood, 1998) and Phono-Graphix (McGuiness et al., 1996) reflect the influence of O-G instruction. In response to the students of interest, these programs initially focused primarily on word recognition, but have expanded to incorporate activities related to reading fluency and comprehension, writing, and oral language development under the rubric of multisensory structured language education.

As outlined in Birsh (1999), the content of multisensory structured language instruction involves six components: (1) phonology and phonological awareness, (2) sound–symbol association, (3) syllable instruction, (4) morphology, (5) syntax, and (6) semantics. This content is embedded in five principles of instruction: (1) simultaneous, multisensory teaching to all learning modalities (visual, auditory, kinesthetic) to enhance memory and learning; (2) systematic and cumulative organization of material; (3) direct teaching through continued teacher–student interaction; (4) diagnostic teaching involving continued assessment of individual needs; and (5) both synthetic (putting parts of language together to form a whole) and analytic (presenting the whole and breaking it down into constituent parts) instruction. With the exception of the multisensory component, which remains controversial, principles 2–5 characterize many effective approaches to reading remediation for students with word recognition and fluency difficulties, along with the focus on explicit teaching of the structure of language.

Older versions represented as O-G approaches have received little research attention (Hallahan, Kauffman, & Lloyd, 1996). Unfortunately, evidence in support of newer generation programs, widely characterized as "research-based" and specific for students with dyslexia, remains scant. The NRP found only four studies with adequate methodological quality that involved variations of older multisensory O-G programs. Two of these programs yielded positive effect sizes and two did not. For example, Oakland, Black, Stanford, Nussbaum, and Balise (1998) implemented the Dyslexia Training Program, an adaptation of the widely employed Alphabetic Phonics program developed at the Texas Scottish Rite Hospital, for 2 years of daily instruction in small groups. In relation to a comparison group of students who were served in "regular practice" classrooms, effect sizes associated with the Dyslexia Training Program were not regarded as significant (NICHD, 2000). Two years of instruction resulted in changes from about the 3rd percentile of word recognition ability to the 10th percentile. Given the severity of the readers' difficulties, and the meta-analytic evidence for remediation programs in general, this judgment may be harsh, for the outcomes were in the moderate range for comprehension and decoding (0.65–0.80), comparable to other reading remediation programs.

In another study, students with identified reading disabilities in grades 2 and 3 who were provided services in public school special education resource rooms received one of two programs in which phonics was taught explicitly, one of which was an alphabetic (synthetic) phonics program based on an O-G method and the other an analytic phonics method (Recipe for Reading). Students in these two groups were compared with a group that received an intervention involving teaching sight-word recognition skill (Foorman et al., 1997). Although there was a clear tendency for students who received the alphabetic phonics program to show better gains in phonological analysis and word-reading skills at the end of 1 year of intervention, these differences were not apparent when verbal intelligence scores—higher in this group—were controlled in the analysis. Foorman et al. (1997) also noted that the size of the instructional groups was too large to promote adequate implementation of any of the programs.

Ritchey and Goeke (2006) identified 12 experimental and quasi-experimental evaluations of O-G methods in schools and clinics, and at elementary, secondary, and college levels. Across the 12 studies, there was a mean effect size of 0.82 for word attack skills, which was lower for word recognition skills (0.42). The effect size for comprehension was 0.76. When the outcomes of students exposed and not exposed to O-G methods were compared, nine studies found outcomes in favor of O-G methods and three did not. Results should, however, be interpreted cautiously because of the small number of studies generating effect sizes and because, in some instances, the range of effects was less than 0. This indicates a wide range of effects and suggests that the true effect may prove negligible in a well-controlled study. Ritchey and Goeke concluded, "Despite the wide acceptance and enthusiasm for OG and OG-based programs, not all studies reported them to be superior, and caution should be taken when attempting to generalize any of the reviewed results" (p. 181). They added, the review "was challenging due to the small number of extant studies that employed experimental or quasi-experimental designs" (p. 181). Eleven of the 12 studies were conducted using quasi-experimental designs and not the randomized control designs that Scammacca et al. (2016) noted have proliferated since 2010 and result in reduced effect sizes.

An important issue is what is meant by the multisensory component, which is often treated as central despite the focus on teaching about the structure of language. It is difficult to provide a precise definition of this term, which will vary across programs. Some studies that compare instruction with and without the traditional multisensory components do not indicate differences in outcomes (Clark & Uhry, 1995; Moats & Farrell, 1999). Wise, Ring, and Olson (1999) did not find that a multisensory articulatory component, as in the Lindamood program, was a necessary component of their own intervention. At the same time, most reading programs require students to pronounce sounds and words accurately and quickly for automaticity. Recent guidelines from the Center for Effective Reading Instruction (CERI; 2018) have tried to demystify the multisensory component by refocusing on the idea of teaching to multiple modalities in the context of reading and writing (saying a word, reading a word, writing a word). There is more emphasis on the intense, systematic, and explicit approach to instruction, the link with specific types of struggling readers, and explicit attention to the structure of language. These are characteristics of any strong approach to reading (e.g., Direct Instruction) and not specific to children with dyslexia, who generally should be taught as other children are taught to read, but with more intensity, time on task, and differentiation. Although there is limited evidence for the efficacy of programs under the multisensory rubric when it refers to teaching to different senses, other programs, reviewed below, which have multimodality components, do show positive effects.

Tier 3 Interventions

Although O-G methods are often implemented as core or supplemental programs, their primary implementation is as an intensive intervention, usually for people with severe WLRD, and usually in a 1:1 teacher:student ratio. As such, we would consider them examples of Tier 3 interventions because of their intensity, explicitness, duration, and group size. In the next section, we identify specific studies that generate strong effects.

Multicomponent Methods

The longest continuing program on reading remediation research in North America is directed by Maureen Lovett at the Hospital for Sick Children in Toronto. The current program, which evolved from a series of studies, is now called the PHAST (Phonological and Strategy Training) reading program (Lovett, Barron, & Frijters, 2013). PHAST includes components of Direct Instruction and strategy training to promote generalization of word recognition strategies. It has a meta-cognitive component that includes text comprehension and writing that promotes generalization to all aspects of reading proficiency, but targeting people with WLRD.

In the initial phase of this research, children with severe word reading disabilities were randomly assigned to either an intervention that is a modification of Reading Mastery, a Direct Instruction program, called Phonological Analysis and Blending/Direct Instruction (PHAB/DI), or to a program with a metacognitive focus that teaches word recognition through the application of different strategies called Word Identification Strategy Training (WIST). Both programs recognize the importance of decoding instruction that helps children break apart words and the importance of instruction that maximizes transfer of learning. The PHAB/DI program emphasizes letter sound units, and the WIST program focuses on larger subsyllable units.

These programs resulted in different patterns of transfer of learning, thus showing treatment-specific effects. For example, PHAB/DI was associated with stronger results specifically on phonological decoding, such as with pseudowords; the WIST program resulted in generalization to regular and exception words in English. These programs did not normalize reading skills, and 35 hours of instruction did not seem adequate. Yet, the students in these interventions were largely in upper elementary and middle school classes when they began the intervention and entered with very severe reading difficulties, often below the 5th percentile. Combining PHAB/DI and WIST into PHAST was subsequent validated as an effective reading intervention (e.g., Morris et al., 2012), PHAST has continued to evolve as Empower Reading (Lovett et al., 2013).

The PHAST reading program has also been employed in multisite intervention studies involving collaboration by Lovett's group, Wolf's group in Boston, and Morris at Georgia State University (Morris et al., 2012). In the initial 5 years, a group of 279 students received different combinations of the interventions in schools in Toronto, Boston, and Atlanta. The samples were carefully constructed to control for variations in SES, ethnicity, and intellectual levels, all involving students in second and third grade. Half the children at each site and in each group were from lower SES backgrounds, and within lower and middle SES levels, half were white or African American.

Four randomized treatment groups were compared. One group received the original PHAST reading program (PHAB + WIST). The second group received a combination of PHAB and Wolf's RAVE-O (Retrieval, Automaticity, Vocabulary Elaboration, and Orthography) program (Wolf, Miller, & Donnelly, 2002), which is described in Chapter 9 in the section on reading automaticity.

These two programs, which combined explicit instruction methods emphasizing the alphabetic principle with different forms of language focus and strategy instruction, were compared against two comparison groups, one that was taught Direct Instruction math and study skills and one that received PHAB/DI along with study skills training. Results showed that students who received either combined condition achieved higher levels of word recognition and comprehension performance than students who received only PHAB/DI; all three groups performed at higher levels than the math comparison group. This intervention, which involved approximately 70 hours of instruction, resulted in changes of about 0.50 standard deviations across reading domains. Approximately 50% of students who received the two combined interventions showed word recognition ability that approximated the average range. The significant gains in word reading, reading fluency, and comprehension using a variety of measures were not only maintained at a 1-year follow-up, but also showed evidence of continued acceleration. *Equally noteworthy, intervention outcomes did not interact with race, IQ level, or SES.* This study reveals that multidimensional programs yielded a gain for children with lower IQs equivalent to that for those with higher IQs at entry, and a benefit for children from lower SES environments equal to that for those from more advantaged circumstances.

Intense Focus on Decoding

Other intervention studies parallel these results, generally showing that the nature of the program is less important than its comprehensiveness and intensity. In a frequently cited study, Torgesen et al. (2001) enrolled students reading below the fifth percentile in word recognition ability in grades 3–5 in an intense 8-week program in which the students received 2 hours of instruction per day, 5 days per week (about 67 hours over the 8-week period). The interventions involved either the well-known Lindamood–Bell Auditory

Discrimination In-Depth program or a program called "Embedded Phonics" developed for this study. A time-by-activity analysis showed that the Lindamood–Bell program involved about 85% time in instruction in phonological decoding, about 10% in sight word instruction, and 5% time in reading and writing connected text. The Embedded Phonics program, in contrast, involved about 20% instructional time in phonological decoding, 30% in sight word instruction, and 50% in reading or writing connected text.

There was little difference in the relative efficacy of the two interventions, so Figure 6.6 collapses across the two interventions. As depicted in Figure 6.6, the results showed significant improvement of about one standard deviation in word recognition, slightly less than one standard deviation in comprehension, and little change in fluency. The gains in word recognition and comprehension persisted for 2 years past the intervention (Figure 6.6). About 70% of the students who received one of these interventions were able to read in the average range, defined as word recognition scores above the 25th percentile, after the intervention and, most remarkably, 40% exited special education.

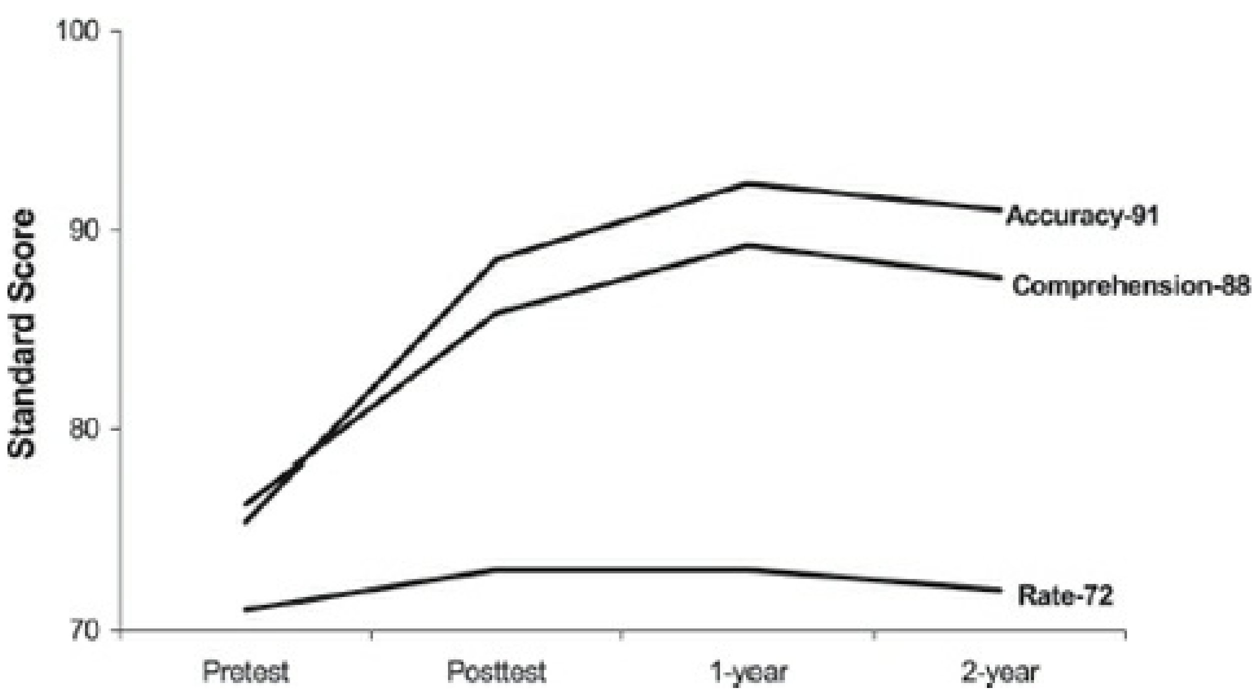

FIGURE 6.6. Growth in accuracy, rate, and comprehension scores from the GORT collapsed across treatment at pretest, posttest, and 1- and 2-year follow-ups. There is more growth in the accuracy of word reading and in comprehension as compared with reading fluency. Data from Torgesen et al. (2001).

Disappointing, however, was the absence of changes in fluency. In explaining disappointing effects for fluency, Torgesen et al. (2001) suggested that reading rate was limited because the number of words in grade-level passages that the students could read "on sight" was much smaller than the number that could be read by average readers. Thus, when comparing fluency rates on stories that were at the student's instructional level, there were no rate differences. However, grade-level passages reduced the fluency differences because there were too many words the students did not have as part of their sight-word vocabulary. There is a strong relation between reading fluency and practice, so that if students are not able to access print for 3–5 years, it would be very difficult to close this gap. Torgesen (2002) estimated that students in the interventions would have to read for 8 hours per day for a year in order to close the gap created by the delay in the students' access to print.

In one of the earliest intervention-imaging studies, Simos et al. (2002a) employed MSI before and after children with severe dyslexia participated in an intense phonologically based intervention based on the explicit program in Torgesen et al. (2001). The children received intervention for 2 hours a day, 5 days a week, over an 8-week period, for approximately 80 hours of intensive phonologically based instruction per child. As Plate 5 shows, before intervention the eight children with dyslexia uniformly displayed the aberrant pattern of activation in the right hemisphere that has been reliably identified with MSI. After intervention, the children's word-reading accuracy scores improved into the average range. In addition, in each case, there was significant activation of neural circuits in the dorsal temporoparietal regions of the left hemisphere. There was also a tendency for reduction in right hemisphere activity.

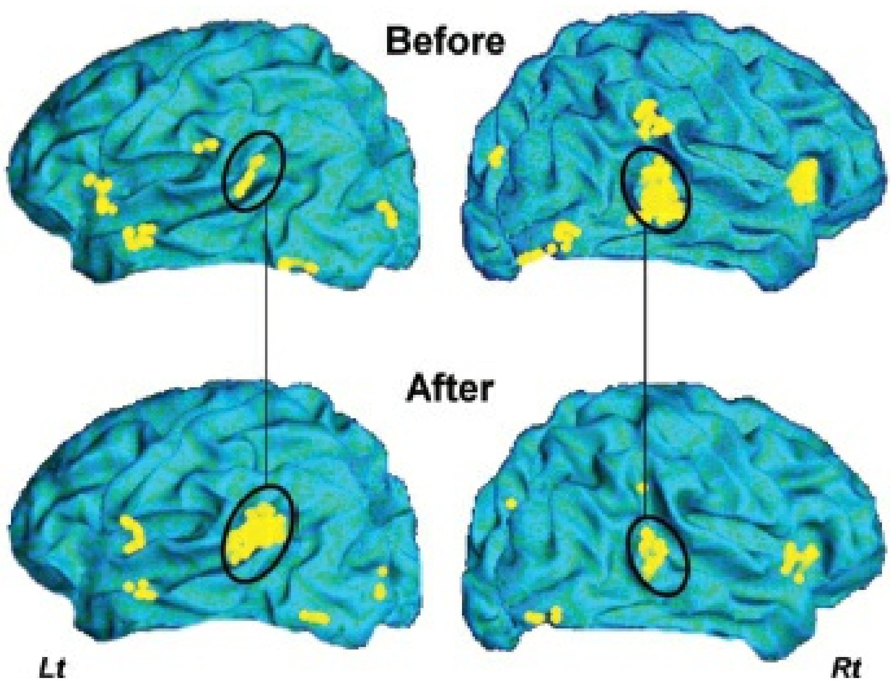

PLATE 5. MSI activation maps from a poor reader before and after intervention. Note the dramatic increase in left temporoparietal activation associated with the significant improvement in phonological decoding and word recognition ability. From Fletcher, Simos, Papanicolaou, and Denton (2004, p. 273). Copyright © 2004 The Guilford Press. Reprinted by permission.

In another MSI study after three tiers of intervention (see Chapter 5; Mathes et al., 2005; Denton et al., 2006b) that included an intensive Tier 3 intervention like Torgesen et al. (2001), Simos et al. (2007b) found changes in brain activity after intervention were primarily normalizing in adequate responders. The changes consisted of increased duration (and degree) of neural activity in the left temporoparietal region and a change in the relative timing of activity in both temporoparietal and frontal regions. These changes were apparent in individual scans involving 12 of the 15 participants. At the end of the intervention, students could be differentiated by the degree of activation in the temporoparietal (dorsal) regions; see Figure 6.7).

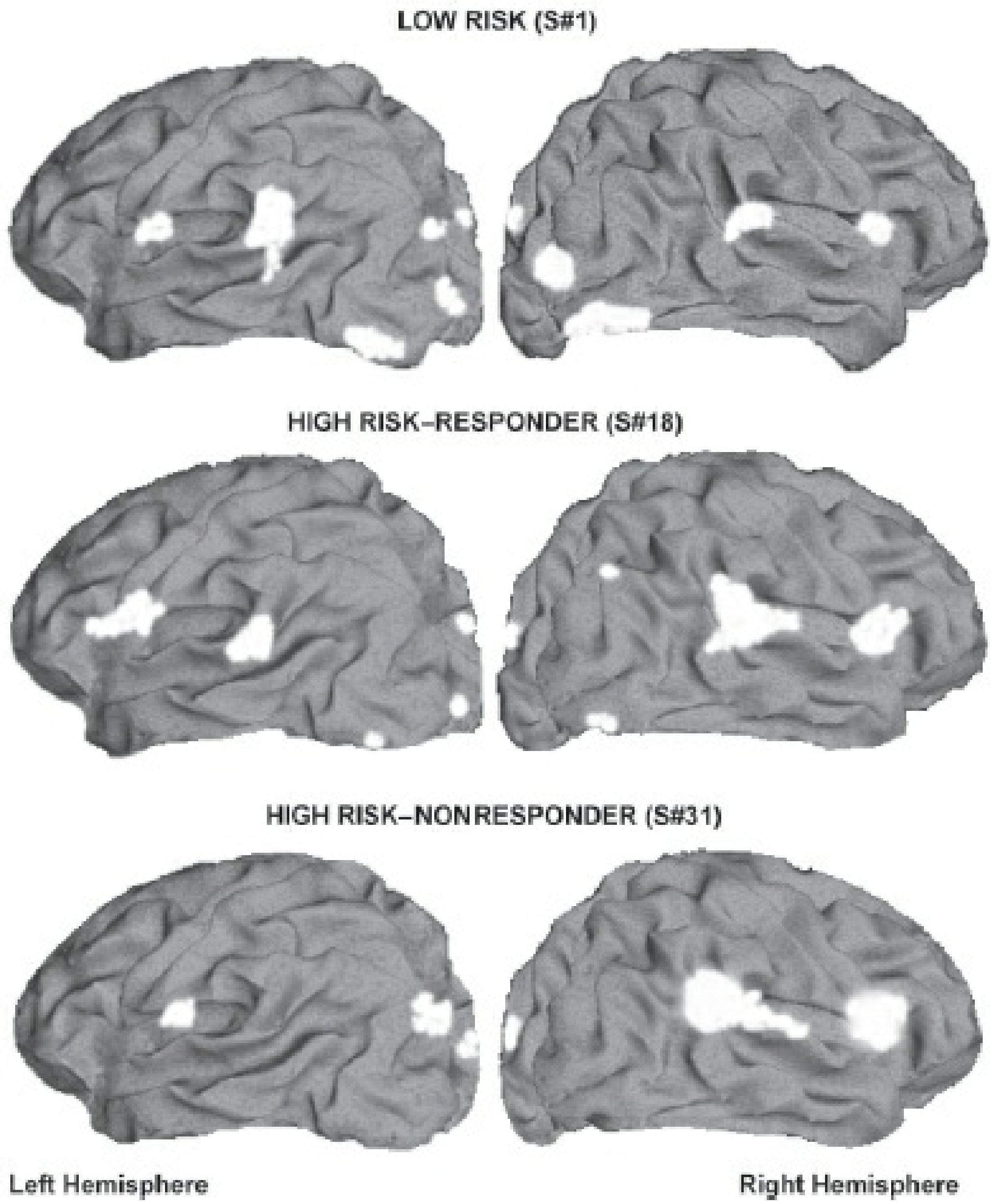

FIGURE 6.7. Activation maps from students who at the end of first grade were at low risk for reading problems; at high risk and responded to intervention; and at high risk but did not respond to intervention. Note the difference in activation involving the left temporoparietal area. From Simos et al. (2005). Copyright © 2005 American Psychological Association. Reprinted by permission.

The approach developed by Torgesen et al. (2001) has not been evaluated experimentally against comparison groups of alternative interventions, nor have the parameters of service delivery been evaluated against alternatives, such as the provision of daily instruction for 1 hour per day for 16 weeks. We do not know whether it is the intensity of the intervention or the number of hours that is important, especially given the findings in the previous section supporting the efficacy of approximately 70 hours of explicit instruction in a small-group, school-year-long format.

Comorbidity

A major question is whether Tier 3 interventions are effective with students who have comorbid behavior difficulties, including ADHD. A meta-analysis (Benner, Nelson, Ralston, & Mooney, 2010) examining the effects of reading interventions for children identified with broadly defined behavior disorders utilizing a variety of reading interventions, often with concurrent behavior interventions, yielded average (but highly variable) effect sizes that were moderate to large across 13 group comparison studies and a set of single-case studies. Most interventions were supplemental to general education classroom instruction. The primary conclusion was that reading interventions were effective with children who have concurrent behavior problems.

A recent randomized trial of children with comorbid WLRD and ADHD provides convincing evidence for the effectiveness of reading interventions as well as the value of interventions for ADHD (Tamm et al., 2017). Children in grades 2–5 (n = 216) were randomly assigned to 16-week trials of reading intervention, ADHD treatment, and combined treatment. The reading intervention was provided to one or two students at a time for 45 minutes 4 days per week. This program included explicit phonics instruction as well as instruction in spelling, fluency, and text comprehension. The ADHD treatment included medication (mostly stimulants, depending on the child's response) and parent training. Intervention outcomes were only reported for word-reading skills and ADHD symptoms. Children in either the single modality reading intervention or the combined treatment showed significant gains in reading compared to the group that received only ADHD treatment with effects in the small-to-moderate range (0.23–0.39). Similarly, children in the ADHD intervention group or the combined group improved in ADHD behavior relative to the reading intervention-only group, with effect sizes in the moderate range (0.36–0.87). Note that combined treatment did not produce crossover effects to reading, but word-reading outcomes were reported. Due to the effects of stimulants on frontal-striatal brain circuits identified as impaired in ADHD (Nigg, 2009), crossover effects demonstrating stronger improvement in the combined treatment group on reading comprehension, which involves higher-order self-regulation skills, may emerge in subsequent analyses of the data.

In addition, a study that randomized children broadly defined with dyslexia, ADHD, and comorbid dyslexia/ADHD to receive atomoxetine, a nonstimulant used for ADHD, reported significant effects on word reading in the group with dyslexia and the comorbid group (S. E. Shaywitz et al., 2017). There were no effects on reading comprehension and no effects on reading in the comorbid group or the group with only ADHD. The reading component was uncontrolled, so the basis for these changes, especially their specificity to the group without ADHD, is not clear. Further study is warranted. In two reviews, Gray and Climie (2016) noted mixed effects on reading for stimulant and nonstimulant medications, with Hutchinson, Ghuman, Ghuman, Karpov, and Schuster (2016) concluding that evidence for effects of atomoxetine on reading in dyslexia was limited and at best mixed.

Treatment of Secondary Students and Adults

In a series of studies, Vaughn and colleagues treated adolescents in grades 6–8 identified because of poor performance on the Texas state accountability measure of reading comprehension. As we noted above, over 80% of these adolescents had problems with decoding and fluency (Cirino et al., 2013). Three studies were conducted over a 3-year period representing Tier 2, 3, and 4 levels of intervention. In the first 2 years, a Tier 1 professional development and coaching model was also implemented for all students in these grades. In all studies, three groups of students were identified and included the following: (1) *typical readers*—students meeting grade-level expectations in reading; (2) *struggling readers*—these students were provided an additional reading intervention for 50 minutes daily in reading implemented by teachers trained by the researchers; and (3) *struggling reader comparisons*—who were not treated by the researchers, but many of whom received additional support (tutorials, afterschool reading groups) typically oriented to test preparation. The interventions always taught word reading, fluency, and comprehension, with a major focus on building background knowledge and vocabulary, but were differentiated by forming homogeneous groups based on their reading profile.

In Year 1, the treatment condition was provided to groups that varied in size from 5 to 15 students (Vaughn et al., 2010a), with specific comparisons of either small-group instruction (four or five students) or large-group instruction (eight to 12 students; Vaughn et al., 2010b).

In the second year, students now in grades 7 and 8 who met benchmarks (passed the Texas state accountability test in reading comprehension) were discharged from intervention and those who did not respond adequately to intervention were randomly assigned to either "individualized" intervention" or to "standardized" intervention. Inadequate responders in the comparison group were also followed (Vaughn et al., 2011b). This contrast permitted comparison of interventions that were intended to align with individual student needs versus a condition in which the same intervention was provided to all students, but with differentiation of groups according to reading strengths and weaknesses. The interventions were provided to small groups of two or three children for 50 minutes each day across a school year. After this Tier 3 intervention, a small group of students who began the study in grade 6 were still inadequate responders. This group received an individualized intervention using a variation of the program implemented in the previous year, in groups of two to four students for 50 minutes daily for a school year.

The results indicate that a year of Tier 2 intervention was not sufficient for many adolescents. The average effect size of differences in reading outcomes relative to the business-as-usual comparison was 0.16. There was no effect of group size. In Year 2, the Tier 3 intervention was associated with significant improvements in decoding, fluency, and comprehension, all in the moderate range. There was no difference in individualized versus standardized interventions except for students identified for special education, where the standardized intervention was more effective.

This finding may seem counterintuitive to long-held beliefs about the importance of individualizing intervention for struggling readers, but virtually none of the meta-analyses reviewed above employed individualized interventions (see Vaughn & Fletcher, 2012). Moreover, the individualized condition did not employ protocols employed in a previous research program conducted with CBM (D. Fuchs et al., 2013). As we reviewed in Chapter 5, data-based design of interventions that make strong use of progress-monitoring data enhance the capacity for individualization and show clearly beneficial effects of individualized intervention based on CBM results (D. Fuchs et al., 2013). So additional research is clearly needed, which employs previously validated approaches to data-based individualization.

The Tier 4 intervention (Year 3) did yield large effects on reading comprehension, with an effect size of 1.20. Figure 6.8 shows slight acceleration in the interval scores of the group that received 3 years of intervention, but less acceleration in the reading comprehension scores in the comparison group. This is not because the students fail to show improvement in reading comprehension. If standard scores were presented (see Figure 5.2 for standard scores), the slopes would be markedly different in both groups. But the adolescents in the comparison group are not making the same amount of progress as their peers, so each cross-sectional comparison shows reduced age-adjusted standard scores (Figure 5.2).

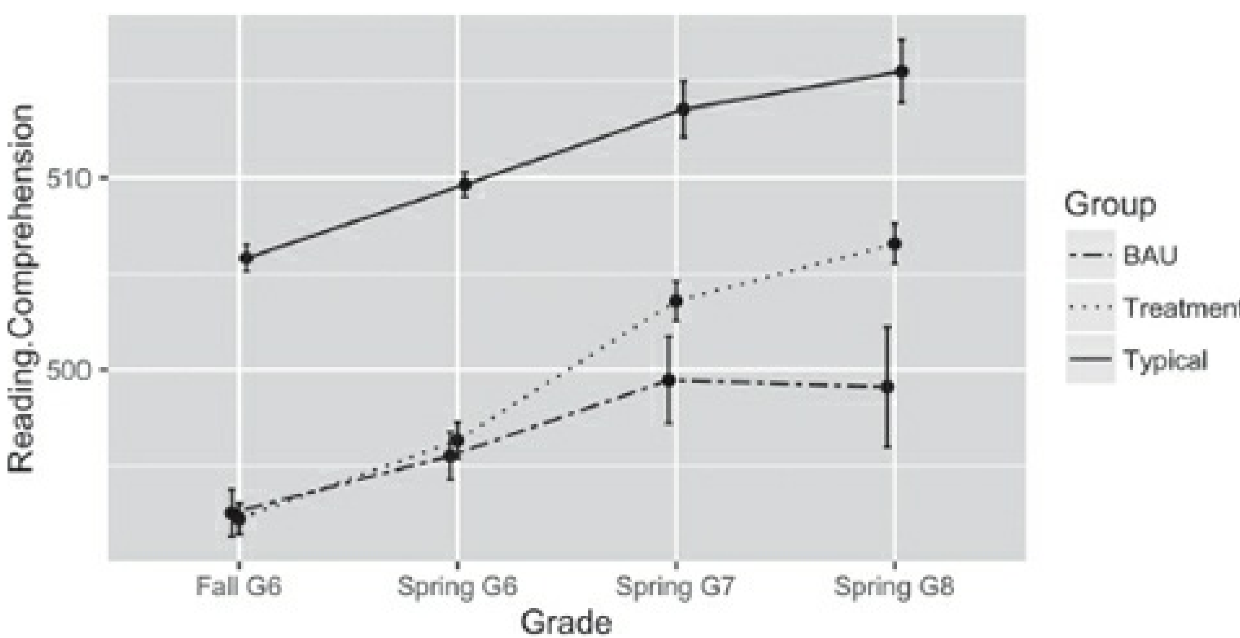

FIGURE 6.8. Changes in reading comprehension interval scores by adolescents in a business-as-usual comparison group, a treatment group that received 3 years of intervention, and typical readers. Students who received intervention showed accelerated gains in reading comprehension and closed the gap relative to typical readers, but adolescents who did not receive the research-based intervention plateau. Data from Roberts et al. (2015) based on Vaughn et al. (2012). Courtesy Greg Roberts and Jamie Quinn.

These sobering results suggest that remedial interventions for adolescents may stabilize reading skills and prevent adolescents from falling further behind. Functional levels of literacy are improving. But without sustained intensive intervention, students fall further behind. This lack of progress may figure into issues like dropout and graduation rates. After Year 3, almost all of the students who received Tier 4 intervention passed the state accountability test.

The level of intensity and duration of intervention may seem surprising. However, the U.S. Department of Education funded 17 randomized control trials to examine the effects of interventions on older students' reading comprehension outcomes, yielding a mean effect size of 0.11 (Boulay, Goodson, Frye, Blocklin, & Price, 2015). Of 10 interventions investigated, only four had positive or mixed results, with the remaining associated with negligible effects. In a less intense study of high school students, Lovett, De Palma, and Frijters (2012) implemented a version of PHAST that also included an emphasis on knowledge of text structures and reading comprehension strategies. A quasi-experimental comparison was performed of one semester of intervention for 268 treated high school students severely impaired in word reading (average word recognition score around the 3rd percentile) and 83 waiting list controls comparable in reading skills. The treatment group demonstrated significant gains in word reading and reading comprehension, with an average effect size relative to the controls of 0.68 (moderate). There was evidence of closing of the gap because of accelerated gains in the treatment group: students who received the modified PHAST program achieved an average gain of about 8 standard score points on word attack skills relative to the control gains of about 2 points. However, after the intervention, acceleration of gains was not apparent for most reading domains, with growth flattening except for slight acceleration of reading comprehension skills. Lovett et al. concluded that while gains in reading can be achieved with adolescents with severe WLRD, one semester is not sufficient for many students.

Even in adults, intervention gains using strong experimental designs and standardized measures are hard to achieve in a year of intervention. Greenberg et al. (2011) compared four instructional approaches in 198 adults with word-reading grade-equivalent scores ranging from 3.0 to 5.9 (< 5th percentile). The interventions involved Decoding and Fluency instruction;

Decoding, Comprehension, and Fluency instruction; Decoding, Fluency, Comprehension, and Extensive Reading; and a control condition involving a standard adult literacy curriculum. Decoding and Comprehension were based on Direct Instruction methods; Fluency was based on repeated reading methods (see Chapter 9). Each participant attended a class for 2 hours, four times weekly, for a total of about 100 hours. It was expected that the more comprehensive condition would lead to greater gains. However, the only statistically significant difference was better word attack skills in the treated groups relative to controls. Effect size differences were negligible. Many other relatively controlled studies of adult literacy programs have not shown robust effects (Torgerson, Porthouse, & Brooks, 2005).

Plate 6 may help us understand why remediation of adolescents and adults is so difficult. This plate (from Rezaie et al., 2011b) shows MSI brain activation in response to a reading task in adolescents prior to beginning the multiyear study of Vaughn et al. (2010a). In the upper panel is an adolescent who is a typical reader; in the middle is an adolescent who responded to the intervention; the bottom panel shows an adolescent who did not respond to the intervention. The absence of left hemisphere activation at baseline in the inadequate responder is striking.

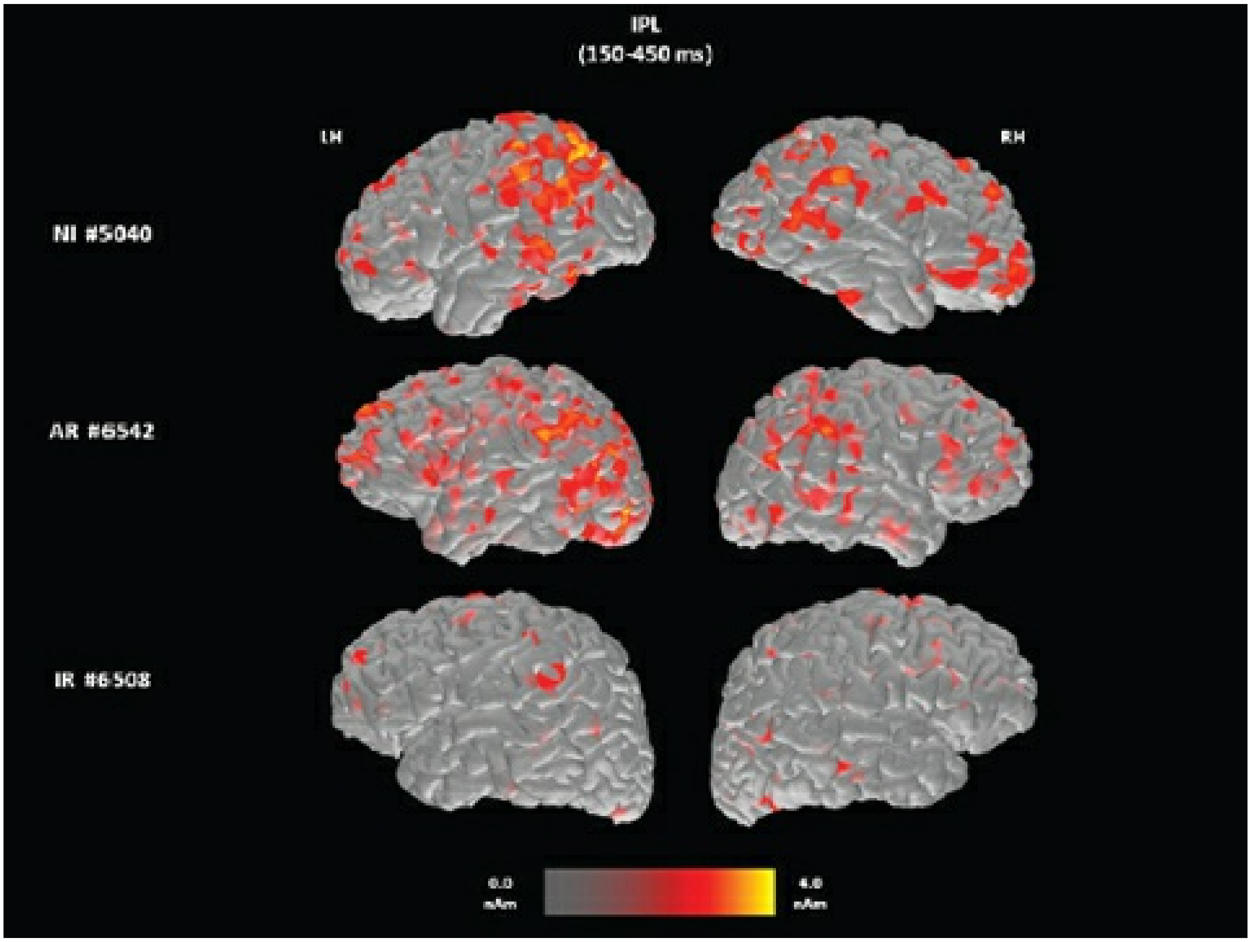

PLATE 6. Brain activation maps using magnetic source imaging from three representative participants in Vaughn et al. (2010): a nonimpaired reader (NI; top row), a student who later showed adequate response to intervention (AR; middle row), and a student who did not show adequate response (IR; bottom row). The relative intensity of activated voxels is shown at the bottom of the figure. Each set of images presents activity in the angular and supramarginal gyri, where significant group differences were observed, between 200 and 500 milliseconds, in the left (LH) and right (RH) hemispheres. From Rezaie et al. (2011b, p. 880). Copyright © 2011 Cambridge University Press. Reprinted with permission.

Summary: Remedial Studies

The severity of the word-reading difficulties of the older samples in Lovett et al. (2012) and Greenberg et al. (2011) is of great concern. There remains a strong faction in the reading education community who minimize the importance of explicit instruction on foundational skills, focusing on implicit and incidental learning of word recognition strategies that are ineffective, such as guessing from context. Remedial studies show that foundational skills can be improved in students with LDs in reading, typically characterized by word recognition difficulties. The effects are most apparent in word recognition, but also show transfer to comprehension if the intervention is sufficiently intense. Fluency gains are often smaller, but vary across studies and may reflect the age and the severity of reading difficulties of the students addressed by the study.

For example, Blachman et al. (2004) and Morris et al. (2012) obtained stronger gains in fluency than Torgesen et al. (2001), but many of the students were younger and their difficulties less severe than those in Torgesen et al. (2001). A variety of approaches are associated with improvement, including commercial programs that were incorporated in different studies (Lindamood–Bell, Phono-Graphix), research-based approaches (Reading Mastery, PHAST, RAVE-O, PASP), and programs that were not reviewed (e.g., Spell-Read PAT; Rashotte, MacPhee, & Torgesen, 2001; see Florida Center for Reading Research, 2005).

The specific program is less important than how it is delivered, provided it is explicit, differentiated, and increases time on task. (Torgesen et al., 2001; Vaughn & Fletcher, 2012). There are also associations with the length of instruction; many hours are required to accelerate reading development in older students (grade 3 and beyond). To reiterate a critical finding, programs that are explicit, oriented to academic content, teach to mastery, provide scaffolding and emotional support, and monitor progress while introducing programmatic adjustments to ensure intervention is addressing individual student needs are more effective. Outcomes are specific to the content of instruction, so that more comprehensive programs yield better outcomes. Future development of remedial programs must involve more attention to reading fluency, which seems least responsive to intervention (see Chapter 9).

Summary: Interventions for WLRDs

Intervention studies demonstrate that dyslexia can be most successfully treated when it is identified early in development and before formal diagnosis. Most impressive are the results of studies that attempt to prevent the development of the sizeable achievement gaps associated with dyslexia: to prevent early deficits from becoming a disability. Although prevalence estimates for WRLD remain high, and always depend on the criteria used to designate a reading problem, there is reason for optimism in terms of reducing the number of students who have intractable reading problems and who require long-term remediation. A key for all research efforts is to focus on clearly defined phenotypes, which for dyslexia we suggest should stem from the assessment of academic skill deficits.

Table 6.1 summarizes important principles from this research review, building on the general discussion of effective intervention principles for LDs in Chapter 5. This table highlights the importance of explicit, differentiated, and multicomponent instruction. Maximal effectiveness is apparent when risk for WLRD is identified early in development as part of a general screening program for reading difficulties. We reiterate the greater difficulty that emerges when intervention is delayed, and the greater constraints introduced on learning because the child has not had access to print and the opportunity to program the neural systems that must be in place in order to support reading.

TABLE 6.1. Intervention: Fundamental Principles for WLRDs

1. Teach phonics explicitly in the context of a multicomponent, integrated instructional program that includes sight-word recognition, spelling, fluency, vocabulary knowledge, and comprehension. Differentiate according to student strengths and weaknesses.
2. Rely on explicit instruction to firmly establish associations between phonemes and graphemes; to address the broad range of phonics patterns and teach these patterns in an orderly way; include cumulative, mixed review so that previously taught patterns receive review and continued practice to develop automaticity with associations and patterns; help learners understand how and why there are exceptions to those associations and patterns; and ensure transfer from word-level competence to text reading by repeated exposure to words and word patterns in text.
3. Teach morphology and larger units of orthography in reading and spelling.
4. Teach using multiple modalities to enhance learning: see the word, say the word, write the word, use the word in text.
5. Engage learners in reading instructional-level material.
6. Prevent word recognition and spelling problems early because later remediation is difficult and requires considerable intensity, especially to develop automaticity.

CONCLUSIONS: WRLDs

This extensive review of research on WLRD, or dyslexia, illustrates the research advances that have been made over the past 30 years around the world in understanding dyslexia. What is especially impressive about the research is not only the growth within domains of inquiry, such as cognitive processing, brain function, genetics, and intervention, but also the integration across domains. Research involving children and adults with dyslexia is linked and is producing an integrated, coherent view of dyslexia. The starting point for any coherent theory is a classification that is reliable and valid, and that yields identification criteria indicating the presence or absence of the class of interest. In this respect, dyslexia is unique among LDs in terms of generating definitions that are inclusionary and that specify how to go about identifying people with dyslexia.

This research shows that the primary academic skill deficits that lead to identification of dyslexia involve problems with the accuracy and fluency of decoding skills, and spelling. Cognitive research identifies reliable correlates and predictors of these marker variables, the most robust involving phonological awareness. Additional cognitive processes involve rapid naming of letters and digits as well as working memory for phonological material. Dyslexia has reliable neurobiological correlates, with a burgeoning evidence base on the neural correlates of word recognition and dyslexia, including the triangle network of brain regions that emerges in development through instruction and experience with reading. There is also substantial research indicating that word reading is a moderately heritable trait and which is beginning to identify specific genetic markers of dyslexia that involve several different genes.

Prevention studies are very promising, but the translation issue continues to loom (see Chapter 11). The data on remedial outcomes in adolescents and adults is sobering. Remediation is possible, but the effects at the secondary level seem to stabilize deficits rather than narrow the achievement gap. This does not mean that the need for intervention should be reconsidered. Instead, intervention intensity must be ensured for these older students. This means longer duration interventions in smaller groups with data-based individualization to promote differentiation.

To identify dyslexia and establish that it is a disability, it is imperative that the concept of instructional response be incorporated into definitions of dyslexia and other LDs (Chapter 3). In the absence of a definition that includes instructional response, prevalence assessments will remain difficult. It is noteworthy that other medical disorders that are essentially dimensional, such as obesity and hypertension, are defined in relation to intervention outcome. Imagine specifying criteria for obesity or hypertension in the absence of data-based guidelines about the point at which treatment is indicated to reduce risks for strokes, heart attacks, and diabetes (Ellis, 1984; S. E. Shaywitz, 2004).

CHAPTER 7

Text-Level Reading Disabilities (Specific Reading Comprehension Disability)

In this chapter, we focus on disabilities involving reading comprehension, which we refer to as a text-level LD to differentiate it from the WLRD reviewed in Chapter 6. There are clearly people with LDs in reading who do not have difficulties with the accuracy and fluency of word-reading skills. These LDs are often described as a specific reading comprehension disability (SRCD) and should not be labeled with dyslexia to avoid confusion because dyslexia is defined by a conspicuous problem with word reading and spelling. Referring back to Figure 1.1, in SRCD, the nature of the academic skills deficit is different (low reading comprehension and intact decoding); the cognitive correlates vary, with little involvement of phonological processing and much greater involvement of oral language and cognitive control skills; and the neurobiological correlates also vary. This difference is easily understood from the simple view of reading (SVR), which was formulated by Gough and his colleagues (Gough & Tunmer, 1986; Hoover & Gough, 1990) In the SVR, reading comprehension is the product of word-reading and linguistic comprehension; people with SRCDs have problems with the listening comprehension side of the SVR. SRCD has been studied less frequently than WRLDs, but research has been flourishing over the past decade.

DEFINITION AND CLASSIFICATION

Most definitions of LDs, such as in IDEA (2004; U.S. Department of Education, 2004), DSM-5 (American Psychiatric Association, 2013a), and the ICD-10 (World Health Organization, 2013) identify reading comprehension as a specific form of LD. The specificity of SRCDs emerges because they are not due to word-level problems. Therefore, definitions are based on the discrepancy of intact word reading and poorer reading comprehension. Definitions of SRCD are difficult because they are based on discrepancy scores and the degree of discrepancy and decisions about average word reading and poor reading comprehension are somewhat arbitrary. Little research has been completed to assess the influence of variability in definitional criteria, but the phenotype clearly exists.

Word reading and reading comprehension are highly related and the co-occurrence of word-reading and reading comprehension disabilities is high. With disability defined as below the 10th percentile in both domains, co-occurrence was 60% in the large Colorado twin study (Willcutt, 2014). Given that readers must be able to read and know the meanings of 90% of words in a text to understand what they are reading (Nagy & Scott, 2000), this high degree of overlap is not surprising. However, other factors also account for the strong relations between word reading and reading comprehension, including partial overlap in the developmental and genetic pathways for these two aspects of reading. Despite this high degree of overlap, word reading and reading comprehension are also separable. Listening comprehension and reading comprehension, for example, are more highly correlated than word reading and reading comprehension, particularly for older children (Catts, Adlof, Hogan, & Weismer, 2005; Christopher et al., 2012; García & Cain, 2014). Comprehension in both skilled and less-skilled comprehenders is similar whether materials are presented through written text, aurally, or even through pictured sequences (Bishop & Adams, 1992; Gernsbacher, 1990; Kendeou, Bohn-Gettler, White, & van den Broek, 2008).

Disabilities in reading comprehension overlap with those in other areas of academics. In the Colorado twin sample (Willcutt, 2014), the correlation of reading comprehension and mathematics was .62. It is difficult to obtain estimates of the overlap of SRCDs with other types of disabilities and neurodevelopmental disorders (e.g., mathematics, writing, ADHD) from the literature because reading disability is often defined through composite scores on decoding, spelling, and comprehension measures (e.g., Wadsworth, DeFries, Willcutt, Pennington, & Olson, 2015). In the Colorado twin study sample, using cut points at the 10th percentile for each academic domain, 45% of children with SRCD had math disabilities and 45% had writing disability (Willcutt, 2014). Data from the Quebec Twin Newborn Study showed modest relations of inattention and reading comprehension even after accounting for those due to decoding (Plourde et al., 2015).

Given the empirical database, we do not argue for the existence of a comprehension disability that is *specific to reading*. Indeed, Kamhi (2009) has argued for the "narrow view of reading." From this view, because of the close link of reading and listening comprehension, only WLRD represents a "specific" form of LD and specific reading or listening comprehension problems should be considered a language problem. Perhaps the most important thing to understand about SRCD is that comprehension for oral language and written language are generally comparable.

Prevalence

Prevalence estimates for SRCD range from 1 to 15% depending on the exclusionary criteria and cut points (for both reading comprehension and word decoding) used to define the groups as well as on age- or grade-level criteria (Clarke, Snowling, Truelove, & Hulme, 2010; Cutting et al., 2013; Keenanet al., 2014; Spencer et al., 2014a; Stothard & Hulme, 1996). Furthermore, the use of single indicators to assess reading comprehension likely introduces considerable variability in prevalence estimates across studies given that different reading comprehension tests tap decoding and language comprehension to different extents and at different ages (see Chapter 4; Keenan et al., 2008; Keenan & Meenan, 2014). With these caveats, studies that report incidence rates in their samples are discussed below.

In a longitudinal study of 8-year-olds (Nation, Cocksey, Taylor, & Bishop, 2010), 8.7% of the sample were adequate decoders/poor comprehenders. Children's reading comprehension had to be below the 25th percentile and their decoding above the 25th percentile with at least 10 standard score points separating the two domains of reading. In a study of over 425,000 children in first to third grades, SRCD was defined as at or below the 5th percentile on the reading comprehension test with nonword decoding fluency at or above the 25th percentile (Spencer et al., 2014a). Less than 1% of children in the first grade and about 2% of children in the second grade met criteria for SRCD. In the first grade, only about 10% of children with reading comprehension difficulties did not also have word-reading difficulties, whereas by second grade, over half of those with reading comprehension difficulties had adequate word reading.

To the extent that estimates of SRCDs may be more accurate in large epidemiological samples, these findings illustrate two points. First, in the early primary grades, reading comprehension deficits without decoding deficits are relatively rare; however, as early as second grade, the proportion of children with reading comprehension difficulties who do not have concurrent decoding difficulties increases considerably.

Do prevalence estimates of SRCD continue to increase with increasing age? Leach et al. (2003) found that most children with specific comprehension difficulties were identified after second grade. Catts, Compton, Tomblin, and Bridges (2012) estimated prevalence in 493 children followed from kindergarten to grade 10, reporting increases in the number of children with SRCDs over time. About 13% of the sample showed late-emerging reading difficulties (i.e., the reading problems emerged after grade 2). Of these children, 52% met criteria for SRCD, 36% met criteria for WLRD, and 12% met criteria for both. Note that 32% could be identified as reading-impaired in any single grade. About 52% were persistently poor in reading, with much of the fluctuation likely due to the measurement issues discussed in Chapter 3.

A large study of 1,748 middle school students overselected for and tested all students with reading difficulties (N = 1,025) based on a state reading comprehension assessment (Cirino et al., 2013). Most of the struggling readers in the sample had difficulties in more than one aspect of reading. Using a threshold of the 20th percentile, about 85% of these older struggling readers had difficulties in reading comprehension, but less than 15% had isolated difficulties in passage-level reading comprehension: Most students' comprehension difficulties co-occurred with difficulties in decoding, word-reading fluency, or both.

In another large-scale twin study of about 1,500 children and youth from 8 to 19 years, the prevalence of SRCD ranged from 1 to 13% depending on the threshold and method used to determine disability status (Keenan et al., 2014). This study is of particular relevance because it did not overselect for poor decoders or poor comprehenders and because comprehension and decoding were measured as latent variables, meaning that several measures of decoding and comprehension were used and the estimate was based on the overlap in variance across measures. The prevalence rates for SRCD varied with how comprehension was measured, the method for selecting poor comprehenders, and the age of the child. When reading comprehension and word decoding were measured at the latent level (average score for decoding had to be one standard deviation higher than the average score for reading comprehension and reading comprehension had to be below the 25th percentile), 7.5–10% of the sample was identified with SRCD depending on the latent variable (real word vs. pseudoword reading) used to define adequate word-level skills.

Similar to the Spencer et al. (2014a) study of younger children, the lowest prevalence rates in the Keenan et al. (2014) study were obtained for the younger children (mean age of 9.5 years) using reading comprehension measures associated with greater reliance on decoding abilities at younger ages (1%). In contrast, when the latent reading comprehension factor was based on tests shown to draw more strongly on general language comprehension abilities, the prevalence rate was 7% in this younger subsample. The rate was 11% when the latent comprehension factor was listening comprehension rather than reading comprehension.

To contextualize these prevalence rates, it is important to ask whether these different means of determining SRCD produce substantial overlap in the individuals who are identified as having reading comprehension disabilities. The answer to this question is “No.” Some reasons why this might be the case are discussed in Chapter 3, reflecting the use of rigid thresholds and the measurement error of the tests.

Sex Ratio

Badian (1999) found male:female sex ratios of about 2.4:1 using an IQ–

achievement discrepancy definition and 1.6:1 for a low achievement definition of reading comprehension difficulties. To our knowledge, this is the only published study examining sex ratios for comprehension. In the Colorado longitudinal twin study, using a range of cut points on a comprehension composite score or on individual comprehension measures, there is a slightly higher significant rate of reading comprehension difficulties in males versus females; however, the ratio is only 1.1–1.2:1 (Willcutt, 2014). This large twin study tends not to show sex differences in any reading domain. Interestingly, one of the largest epidemiological studies of oral language disorders in young children observed slight male preponderance: boys (8%) and girls (6%) (Tomblin et al., 1997).

Developmental Course and Outcomes

There are few data on outcomes related to SRCD. However, longitudinal studies of reading comprehension help illustrate how reading comprehension develops in typical and struggling readers and line up with the studies below on core cognitive processes. In this next section, we make use of longitudinal studies of typical and atypical development of reading comprehension to identify the "pressure points" for reading comprehension (Perfetti & Adlof, 2012). Many reading and reading-related cognitive skills and sources of knowledge have been correlated with reading comprehension, but not all of them are equally important for comprehension. Using time precedence, longitudinal studies can help to determine the important developmental precursors associated with later reading comprehension.

Typical Development

First we consider findings from longitudinal studies of reading comprehension in children unselected for reading comprehension difficulties. As we argued in Chapters 2 and 3, research evidence is highly compatible with the idea that the attributes of LDs are dimensional, representing the lower end of the normal distribution of an academic skill or other attribute (e.g., instructional response).

Longitudinal studies of reading comprehension are more recent than those pertaining to the development of word reading; most have been conducted within the last 15 years. Most follow children over relatively short developmental time windows, with a few exceptions. One study that followed typically developing readers from 7 to 11 years found that vocabulary, grammatical knowledge, and knowledge of story structure, as well as inference and comprehension monitoring all contributed unique variance to reading comprehension across this age range even after controlling for the autoregressive effects of reading comprehension from earlier assessment points (Oakhill & Cain, 2012).

These researchers also assessed the role of domain-general cognitive processes. They found that working memory also predicted unique variance in reading comprehension and did not fully account for the relation of two text-level abilities—inference and comprehension monitoring—to reading comprehension (Cain, Oakhill, & Bryant, 2004a). Similarly, de Jong and van der Leij (2003) found that vocabulary skills at age 7 predicted reading comprehension 3 years later, even after controlling for prior reading comprehension and decoding.

In a longitudinal study across grades 1 to 6, Verhoeven, van Leeuwe, and Vermeer (2011) found a reciprocal relationship between vocabulary development and reading development. Across these grade levels, stability in vocabulary ability was high. At the beginning of school in this Dutch sample (grade 1), vocabulary predicted later reading decoding and comprehension, but decoding starting in second grade also predicted vocabulary development through to sixth grade. In contrast, a study from first to fourth grade determined that while vocabulary knowledge from earlier time points was a leading indicator for later reading comprehension, the opposite relationship did not hold (Quinn, Wagner, Petscher, & Lopez, 2015). Interestingly, early vocabulary knowledge may be a better predictor of later reading comprehension than early listening comprehension (Verhoeven et al., 2011). Findings from all of these studies are consistent with the lexical quality hypothesis (Perfetti, 2007), which stresses the importance for reading comprehension of acquiring and integrating information about both word form and meaning during reading; what these longitudinal studies show is that word meaning is accounting for *growth* in reading comprehension across development.

Longitudinal studies of children beginning in the preschool years show the increasing importance of oral language abilities for reading comprehension over time. While oral comprehension and decoding are strongly related in preschool and at the beginning of schooling, this relationship begins to diminish over the early primary grades (Kendeou, van den Broek, White, & Lynch; 2009; Storch & Whitehurst, 2002). In the Storch and Whitehurst study, word decoding was the strongest predictor of reading comprehension in the early grades, whereas by grades 3 and 4, oral language skills were more predictive of comprehension. Kendeou et al. (2009) found that oral comprehension and decoding skill at age 6 each made a unique contribution to reading comprehension at age 8. It is likely that the relatively greater strength of word decoding as a predictor of reading comprehension earlier versus later in development is partly due to differences in the requirements of reading comprehension measures at different ages (see Chapter 4).

These longitudinal studies of typically developing school-age children are informative for understanding what uniquely contributes to the development of reading comprehension. However, they do not tell us about how early developing comprehension-related abilities are or are not associated with later reading comprehension. Unlike code-based skills, components of comprehension-related factors such as vocabulary and world knowledge, grammatical knowledge, and knowledge of text structure, particularly for narratives, are developmentally unconstrained. That is, they begin to develop early in life and undergo considerable change in the preschool years. Thus, it is also important to understand the developmental course of reading comprehension by understanding the relation of these early developing comprehension-related abilities in infants and preschool children to their later reading comprehension.

Findings from the few longitudinal studies that follow children from the early preschool years through to school age and that measure reading comprehension outcomes are generally consistent with the school-age longitudinal studies of comprehension discussed above. Early language (a latent construct of receptive and expressive vocabulary) at 16–24 months predicted a range of school-age reading and language outcomes including phonological awareness, word-reading accuracy, vocabulary, and reading comprehension (Duff, Reen, Plunkett, & Nation, 2015). In another longitudinal study, inference skills, literal narrative comprehension, and grammatical knowledge at 4–5 years of age were longitudinal predictors of reading comprehension 1 year later. The relation of vocabulary knowledge and reading comprehension was mediated by both inference and literal comprehension (Silva & Cain, 2015).

Atypical Development

There are few longitudinal studies of children with SRCDs. The paucity of studies is likely related to the challenges of identifying this population in prospective studies, particularly the subgroup of children with adequate decoding, but poor comprehension. Because adequate decoding is a prerequisite for reading comprehension, SRCDs are not typically identified until after third grade when word recognition and reading fluency have become more consolidated (Leach et al., 2003).

Longitudinal studies of children with SRCD have retrospectively identified less skilled comprehenders at later ages from their larger longitudinal samples and looked at what abilities at earlier developmental time points are related to later difficulties in reading comprehension. Catts, Adlof, and Weismer (2006) identified 57 poor comprehenders in eighth grade from a larger sample of children who had been participating in a large-scale epidemiological study and were assessed in kindergarten, grade two, and grade four. This allowed for retrospective analyses of students who had developed specific comprehension impairments. The group with SRCD differed from typically developing readers and poor decoders in vocabulary, grammatical knowledge, and oral discourse comprehension in kindergarten, in grades two and four, and in grade 8, where inference making also emerged as a source of group differences. While there was no difference in phonological processing skills between the good and poor comprehenders in grades two, four, or eight, the poor comprehender group had weaker phonological awareness skills in kindergarten compared to the controls.

Using a similar retrospective methodology by testing a large sample of children at ages 5, 5.5, 6, 7, and 8, Nation et al. (2010) identified 15 poor comprehenders at age 8. A comparison group of same-age good comprehenders was also selected from this sample. These authors found that, in general, poor comprehenders did not show early deficits in the areas of word-reading accuracy or fluency. Both groups achieved age-appropriate levels of letter knowledge, early word reading, and reading fluency at all time points, including kindergarten. The children with comprehension disabilities differed at each earlier time point from their peers with better comprehension in vocabulary and grammatical knowledge and listening comprehension. In contrast to Catts et al. (2006), Nation et al. (2010) found no statistical differences on most measures of phonological awareness at any of the time points, with the exception of the sound-matching task at age 5. However, some moderate effect sizes did emerge when comparing early phonological awareness skills between these two groups, which in light of the small sample size suggests that poor comprehenders may have some difficulty with this skill earlier in their development.

What do we know about how language abilities assessed in infancy and the early preschool years predict later difficulties in reading comprehension? Using similar cut points as Catts et al. (2006) to define their study groups, Justice, Mashburn, and Petscher (2013) compared receptive and expressive language abilities from 15 months to 54 months of age in fifth-grade children with low reading comprehension, but adequate decoding, to peers with adequate decoding and comprehension. Statistically significant differences between the groups emerged in language comprehension and expression by 54 months. Large effect sizes but nonsignificant contrasts for language measured at earlier time points could reflect small sample size or less stability in early language measures, particularly when measuring language using single indicators (Bornstein, Hahn, Putnick, Suwalsky, 2014).

In a study that used a longitudinal mediation framework, the developmental precursors of reading comprehension were measured beginning at 36 months in typically developing children and in children with spina bifida myelomeningocele (SBM), a neurodevelopmental disorder associated with poor comprehension but good word reading (Barnes et al., 2014; Pike, Swank, Taylor, Landry, & Barnes, 2013). Working memory, narrative retelling using words and gestures, and listening comprehension at age 36 months mediated or partially accounted for the group differences on a measure of inferential reading comprehension at 9.5 years (Pike et al., 2013). Consistent with the findings from the study by Catts et al. (2006) reported above, even though the groups did not differ in word-reading ability at 9.5 years, phonological awareness at 5 years mediated the effect of group on reading comprehension at 9.5 years (Barnes et al., 2014).

Findings from these longitudinal studies of children with SRCD are highly consistent with longitudinal studies of children without comprehension disabilities, reflecting the dimensional nature of LDs. They suggest that children who go on to have SRCDs have early deficits in several aspects of knowledge and oral language, including vocabulary knowledge, grammatical knowledge, knowledge of text structure, discourse/listening comprehension, and inference making. There is some evidence that working memory might be a unique contributor to reading comprehension beyond these other sources of knowledge and skills.

Interestingly, the longitudinal findings also suggest that children who go on to become poor comprehenders may be *slower to develop* phonological awareness skills even though these early difficulties are no longer discernable by the early primary grades. One hypothesis is that early phonological deficits may be indicative of a processing bottleneck in reading acquisition that contributes to reading comprehension impairments (Catts et al., 2006). Another hypothesis is that phonological abilities contribute both to the development of word-level semantics (i.e., vocabulary) and to the development of word decoding (Metsala & Walley, 1998).

Given the importance of vocabulary knowledge and oral language across development for predicting reading comprehension ability and disability, it is tempting to ask whether early language measures provide an avenue for early identification and intervention for children prior to entering school. However, large population-based longitudinal studies suggest considerable instability in the language abilities related to later reading comprehension. When core language abilities are measured at a latent level (i.e., when measurement is not dependent on a single test), there is considerable stability between 4 years of age to 10 and 14 years, but less stability between 20 months and 4 years (Bornstein et al., 2014). These findings are generally consistent with those from large population-based studies in which language in the first 2 years of life does not explain much of the variance in language in the preschool and early school years (e.g., Ghassabian et al., 2013). In the Duff et al. (2015) study of children followed before 2 years of age out to 4–9 years of age, early language (parent report) explained 11% of the variance in later reading decoding and 18% of the variance in later reading comprehension. These findings suggest that at the level of the individual, our current methods for measuring language in infants and toddlers is not technically adequate for the purposes of early identification of children at risk for later disabilities in reading comprehension.

ACADEMIC SKILL DEFICITS

The primary academic skill deficits used for defining SRCD involve, simply put, deficits in the variety of abilities that allow the reader to abstract meaning from text. This is a complex set of processes closely linked with listening comprehension, noting also the differences in language systems by eye and by ear (Gough & Tunmer, 1986). For example, comprehension of reading can be no stronger than the comprehension of language, a clear example being vocabulary: a child may be able to decode a word, but if he or she does not know the word's meaning, comprehension of the text will be impaired. Although there are a small number of reading, language, and domain-general cognitive abilities that emerge below as core deficits that characterize individuals with disabilities in reading comprehension, we do not propose that these abilities are relevant to the assessment of reading comprehension for the purposes of identification (see Chapter 4). In addition, we emphasize the need to assess reading comprehension using multiple procedures because of variations in determining impairments based on the nature of the test (Keenan et al., 2008).

CORE COGNITIVE PROCESSES

Reading comprehension requires the coordination of multiple cognitive processes directly and indirectly involved in understanding spoken and written text. Prior to discussing specific cognitive processes, a brief review of reading comprehension frameworks is needed to illustrate how these core cognitive processes are linked.

Models of Reading Comprehension

Three main frameworks guide research in reading comprehension development and disability (Barnes, 2015). One focuses on describing *component skills* that contribute to variance in text comprehension (e.g., the direct and inferential mediation [DIME] model of Cromley & Azevedo, 2007; the SVR of Hoover & Gough, 1990). Figure 7.1 depicts the version of the DIME model tested in Cromley, Snyder-Hogan, & Luciw-Dubas (2010). Historically, the component skills framework comes from the SVR. Some researchers have further articulated this framework by adding a reading fluency component (Kirby & Savage, 2008) and by parsing linguistic comprehension into component skills such as word and world knowledge, inference making, and comprehension monitoring/strategies, as in the DIME model (Figure 7.1; Cromley et al., 2010).

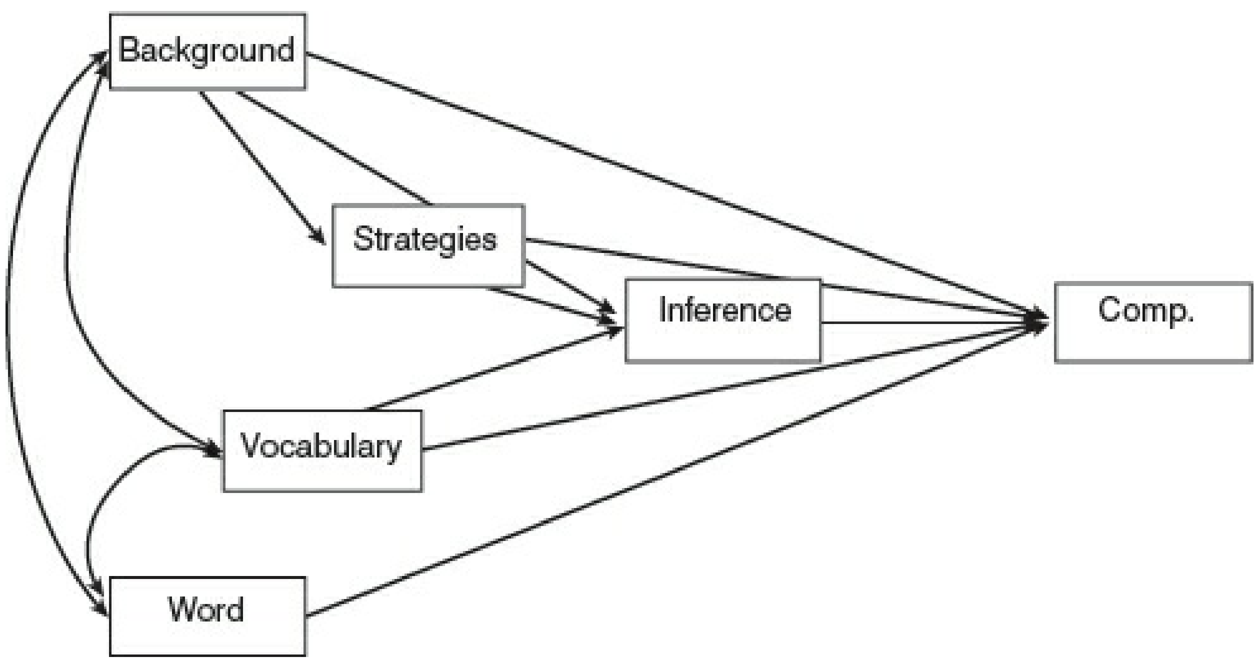

FIGURE 7.1. DIME model with the five components of reading comprehension: background knowledge, vocabulary, word reading, strategies, and inferencing. From Cromley, Snyder-Hogan, and Luciw-Dubas (2010, p. 688). Copyright © 2010 the American Psychological Association. Reprinted by permission.

Another framework describes the *process* of comprehension in which lower- and higher-level reading and cognitive processes interact as the text unfolds over time to result in the construction of a representation of the real-world situation that the text describes (e.g., the structure-building framework of Gernsbacher, 1990; the construction integration model of Kintsch, 1988; the landscape model of van den Broek, Rapp, & Kendeou, 2005; the situation models framework of Zwaan & Radvansky, 1998). This framework is less focused on cognitive components and more so on cognitive processes and the interaction of text characteristics and the reader.

A third approach is the reading systems framework (Perfetti, Landi, Oakhill, & Snowling, 2005), which combines aspects of component skills and process models. It posits the interaction of various sources of knowledge (lexical–semantic, orthographic, syntactic, general knowledge including knowledge about text structure) with comprehension processes such as inference involved in the construction of the text-based representation (the meaning of the propositions in a text) and the situation model (the real-world situation described by the text). This framework also assumes the influence of cognitive processes, including limited-capacity working memory (see Figure 7.2; Perfetti & Stafura, 2014).

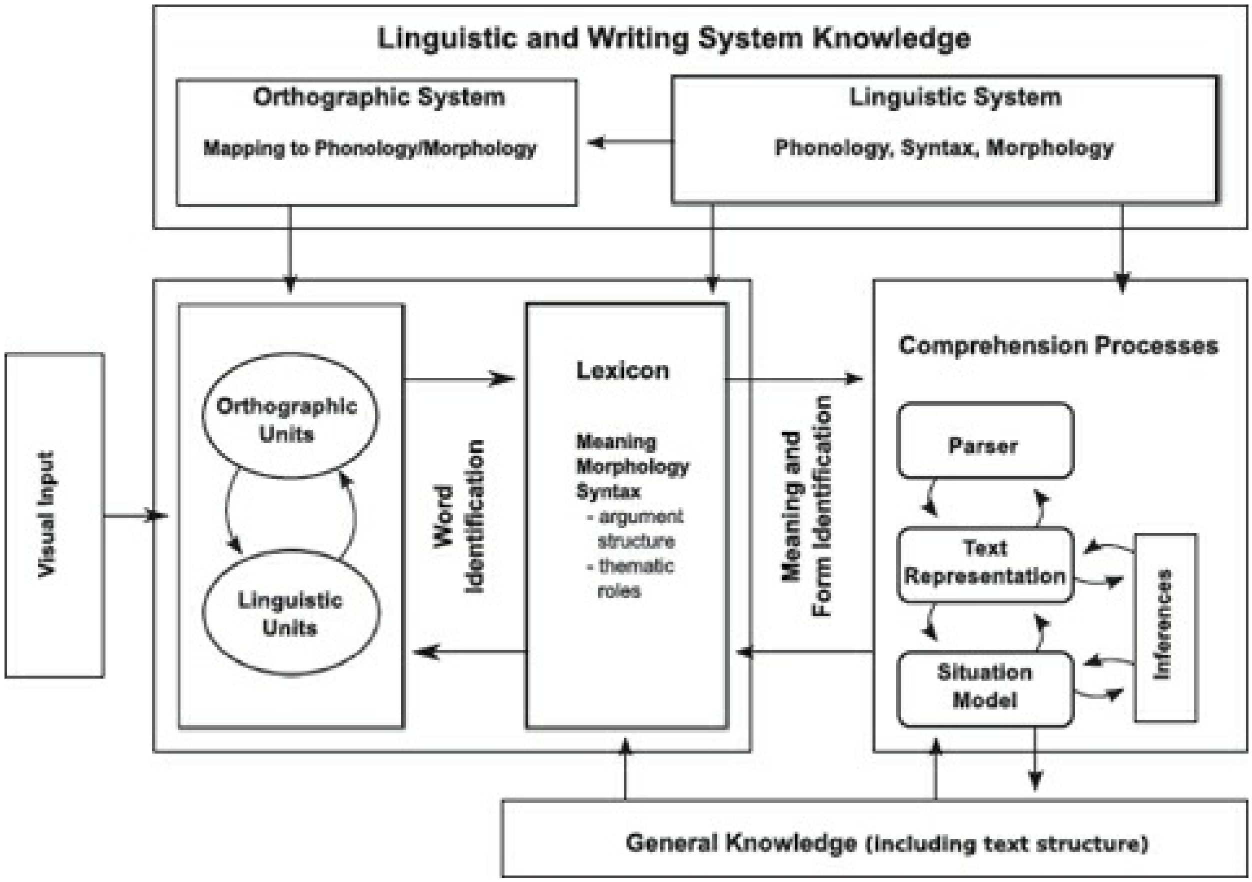

FIGURE 7.2. The reading systems framework. The most important components are sources of knowledge, cognitive and language processes, and their interaction. The lexicon mediates word identification and comprehension systems. Adapted from Perfetti and Stafura (2014). Copyright © 2014 Taylor & Francis. Reprinted by permission.

To some extent, these approaches represent different levels of explanation and investigation rather than alternative models (Barnes, 2015). Given the global nature of theoretical frameworks for reading comprehension, it is perhaps not surprising that we do not have well-articulated models of the *development* of reading comprehension. One exception is Connor's (2016) lattice model (Figure 7.3), which posits reciprocal effects across early and middle childhood of text-specific, linguistic, and social–cognitive processes that both influence and are influenced by learning to read and comprehend. Given the recent emergence of this model, it will be of interest in future studies to see whether and how these putative reciprocal sources of influence account for reading comprehension in children who are identified with disabilities in reading comprehension. This model includes instruction, which should be an important component of developmental models of academic skill acquisition, even though it is often not considered. Cain and Barnes (2017) suggest that constructing developmental models of reading comprehension will depend, in part, on understanding how situation models develop out of children's very early and ongoing experience with their linguistic, social, and physical environments.

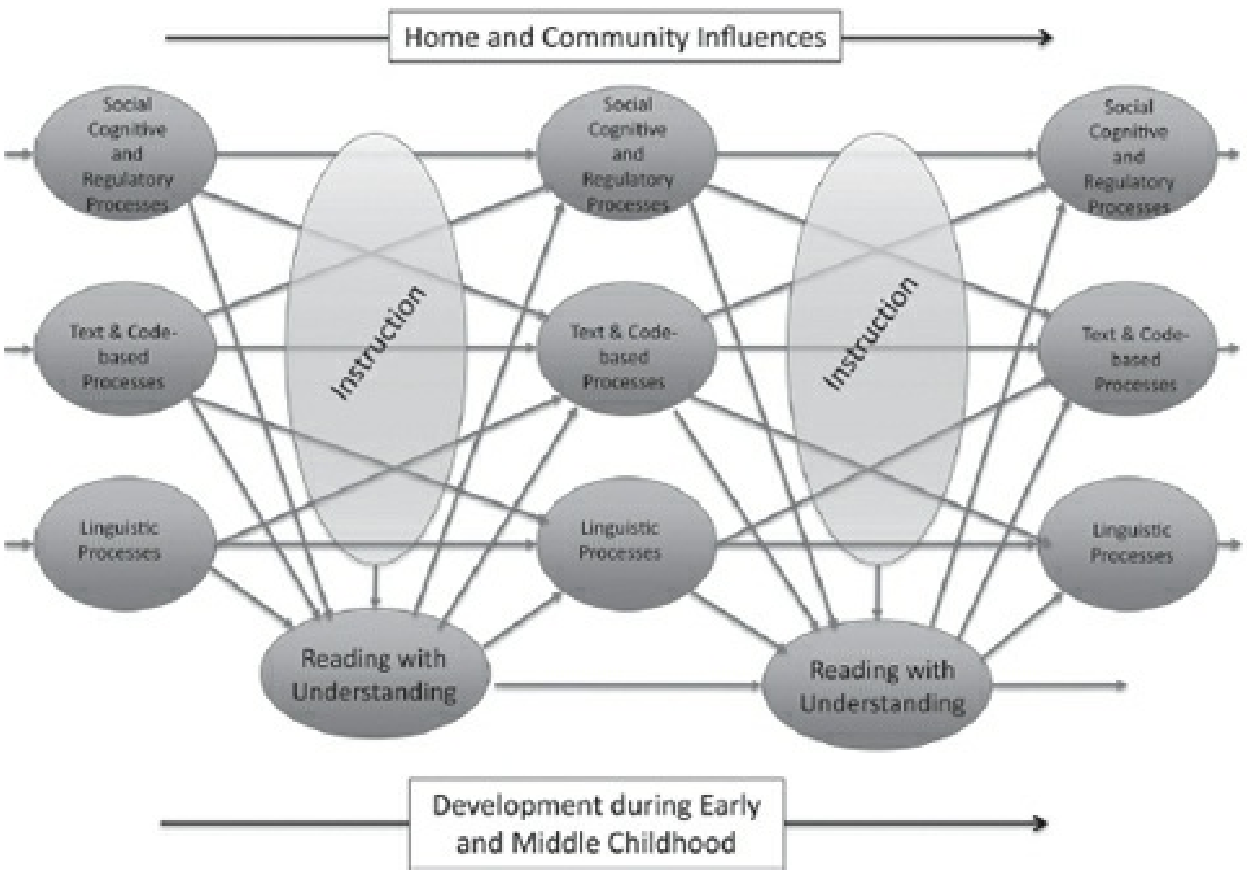

FIGURE 7.3. The lattice model of reading comprehension, showing reciprocal effects of language, textual, and cognitive processes. These relations are shown in the context of social, environmental, and school influences. In development, these relations change, but help determine how teachers should make decisions about instruction for individual students. From Connor (2016, p. 270). Copyright © 2016 John Wiley & Sons, Inc. Reprinted by permission.

Findings from descriptive studies of the correlates of reading comprehension difficulties in individuals with SRCD generally align with what might be expected from these various reading comprehension frameworks. Studies have identified core cognitive correlates of SRCD involving foundational language abilities, listening comprehension, working memory, and a variety of processes that support meaning construction at the level of text and discourse, such as inferencing.

Foundational Language Abilities

Children with SRCDs often have more basic deficits in vocabulary and general semantic or world knowledge, morphology, and understanding of syntax that impair reading comprehension (Nation, Clarke, Marshall, & Durand, 2004; Stothard & Hulme, 1996). The language deficits of these children are typically not severe enough to classify them as speech and language-impaired (Nation et al., 2004); furthermore, their phonological skills are typically not deficient by school age (Cain & Oakhill, 1999; Oakhill, 1993). Indeed, vocabulary knowledge is closely related to background knowledge and has direct and substantial influences on reading comprehension regardless of comprehension skill level (e.g., Ahmed et al., 2016).

Even when vocabulary knowledge is controlled in studies comparing skilled to less skilled comprehenders, poor comprehenders have more subtle, but significant, semantic deficits. These deficits are apparent on measures of language involving semantic judgment and fluency (Nation, Adams, Bowyer-Crane, & Snowling, 1999), as well as on those measuring rapid access to less familiar meanings of common words (e.g., the river meaning of "bank"; Henderson, Snowling, & Clarke, 2013).

Ultimately, language development is at the heart of reading comprehension. In studies by Catts and colleagues that look specifically at language skills in poor comprehenders, and comprehension skills in children with oral language impairments, the overlap is high and problems with vocabulary and syntax are common links (Catts et al., 2006). These early developing core language abilities are also highly predictive of later reading comprehension. Foundational oral language abilities in kindergarten (i.e., vocabulary, grammar, narration) add significant prediction to reading comprehension in third grade even after controlling for kindergarten prereading skills (i.e., phonological awareness, rapid naming) and word decoding in second grade (Catts, Nielsen, Bridges, & Liu, 2016).

Listening Comprehension

Language comprehension and listening comprehension are sometimes both used to refer to receptive language skills. However, in the field of reading comprehension, listening comprehension means more than just receptive language skills. It includes discourse-level processes such as inference making and knowledge of text structure as well as general cognitive resources such as working memory that impact both reading and listening comprehension. Thus, *listening comprehension* is a construct that needs greater delineation and that has similar measurement challenges as reading comprehension. We address some of the processes underlying both reading and listening comprehension below.

Just focusing on the term as it is used in the field of reading comprehension, it is well established that difficulties in listening comprehension parallel problems with reading comprehension (Stothard & Hulme, 1996). Studies comparing reading and listening comprehension in normative samples show high levels of overlap (e.g., Barnes, Dennis, & Haefele-Kalvaitis, 1996). Children cannot understand written language any better than they can understand the same information presented to them orally. Dissociations of listening and reading comprehension may occur in some cases. For example, comprehension of complex text might be better during reading versus listening where speed of processing is under the reader's control in the former modality. However, little research demonstrates these dissociations. Any language or cognitive difficulties that hinder oral language comprehension also affect the person's ability to understand text that he or she reads or even to comprehend text read to him or her. Despite these relations, Carretti, Caldarola, Tencati, and Cornoldi (2014) found that teaching listening comprehension has less impact on reading comprehension than teaching reading comprehension directly using similar programs, although there was improvement associated with the listening comprehension program.

Working Memory

Working memory is commonly identified as a source of cognitive deficit that affects comprehension in poor comprehenders. Both listening and reading comprehension make demands on working memory because they require that words and sentences be simultaneously processed and stored to facilitate the integration of words and ideas within the text as well as the integration of text with prior knowledge. Meta-analysis has documented moderate-to-strong relations of verbal working memory and comprehension (Daneman & Merikle, 1996); however, working memory is as strongly related to word reading as it is to reading comprehension (Peng et al., 2017), and explains some of the overlapping variance between word reading and reading comprehension (Christopher et al., 2012).

Studies of individuals with reading comprehension difficulties typically find that working memory is impaired (Stothard & Hulme, 1996; Nation et al., 1999). Cain, Oakhill, and Lemmon (2004b) found that learning of novel vocabulary from context (i.e., incidental word learning) was impaired in poor comprehenders when the context was not adjacent to the new word, and in this study, working memory capacity, but not immediate memory span, was related to the successful inferring of meanings of novel words from context. Cain, Oakhill, and Bryant (2004a) found that working memory, as assessed by a sentence span test, contributed unique variance to inference making, comprehension monitoring, and story structure knowledge even when decoding ability, Verbal IQ, and vocabulary were controlled. However, the exact nature of the relation between working memory and comprehension is not well understood, particularly in individuals with comprehension difficulties (Savage, Lavers, & Pillay, 2007). For example, although most reading comprehension models presume that working memory is a general cognitive resource that facilitates comprehension processes, there is some evidence that reading per se may account for developmental changes in verbal memory (e.g., Nation & Hulme, 2011). The relation of storage/integration and inhibitory processes in working memory to individual differences in various comprehension processes have been proposed (Pimperton & Nation, 2010), but require further study (Barnes, Stuebing, Fletcher, Barth, & Francis, 2016b). However, even in studies that involve working memory, assessments of higher-order text-level processes contribute unique variance to comprehension outcomes, as shown in the next section.

Higher-Order Text-Level Processes

Reading comprehension cannot be explained solely on the basis of word recognition, foundational oral language abilities, and working memory. Even when these skills are controlled, deficits in reading comprehension still arise (Cain, Oakhill, & Bryant, 2000; Cain et al., 2004a). This is due to difficulties with discourse-level skills including: inferencing, comprehension monitoring, text integration, and other metacognitive skills related to comprehension that are partly, but not completely, explained by variability in working memory (Cain et al., 2004a, 2004b).

Inferencing

A substantial body of research shows that even when poor comprehenders understand literal or stipulated meanings provided by the surface code of the text, they have difficulty making a variety of inferences that require integration of text or integration of text with knowledge, and these difficulties can occur even at very local levels of text such as anaphoric reference (Yuill & Oakhill, 1991). The difficulties are apparent even when working memory demands and differences in background knowledge are controlled (Cain, Oakhill, Barnes, & Bryant, 2001). It may be that the inferencing problems do not reflect a fundamental inability to make an inference, but rather an inability to do so in the context of text comprehension. This would suggest a strategic deficit, for example, a failure to monitor and repair comprehension by making an inference necessary to maintain semantic coherence, a resource-capacity deficit (i.e., where ongoing integration of information at the text level exceeds working memory capacity), or both. Cain and Oakhill (1999) found that prompting poor comprehenders to engage in a strategy that would support making an inference led to improved inferencing. Similarly, reducing both working memory and metacognitive demands during reading also resulted in improved inferencing, but did not eliminate inference-making differences between skilled and less skilled comprehenders (Cain et al., 2001).

Comprehension Monitoring

Several metacognitive processes are used to control and check comprehension when reading (and listening). Successful comprehension monitoring requires the reader to identify inconsistencies in the text, gaps in understanding, or the need to seek information from other parts of the text or from one's own store of world knowledge. These have been referred to as cognitive control processes involved in the reader's conscious and strategic *search after meaning* (van den Broek et al., 2005). Nation (2005) summarized multiple studies indicating that poor comprehenders have difficulties with monitoring. Thus, a focus on comprehension monitoring is a common part of strategy instruction in reading comprehension interventions.

Text Structure Sensitivity

As a final example of a higher-order process important for reading comprehension, consider the child's sensitivity to the nature of the text he or she is reading. Texts have different genres. They can represent narrative stories, expository text, poems, directions, hypertext, and other genres. Each genre carries a distinct linguistic style and is often laid out in ways that vary. Understanding this variation facilitates comprehension. Although some narrative structures (e.g., setting, characters, plot, solution) common to stories are found in both written and oral texts, some text structures, such as those for expository text (e.g., compare-contrast, cause-effect; Meyer, 1987), are primarily found in written language, suggesting one way in which knowledge of text structure may sometimes be specific to reading comprehension.

In addition to effects of genre, other aspects of the structure of text provide important information that facilitates comprehension, including the title of the story, the first sentence of the paragraph, beginning and ending paragraphs, and related aspects of story structure. Children who struggle with comprehension are less aware of genre and story structure variation. They do not attend to this type of information, but do respond to efforts that attempt to teach them about text features and how attending to these features facilitates comprehension (Perfetti et al., 2005).

Component Skills Studies

Several studies have tested the component skills models discussed above to determine the relative importance of hypothesized component skills to reading comprehension, including whether these relations change with age and level of comprehension skill. One large-scale study of adolescents took a latent variable approach to modeling reading comprehension outcomes in students from seventh to 12th grades (Figure 7.1; Ahmed et al., 2016). Vocabulary and world knowledge were indistinguishable at a latent level, suggesting that vocabulary may be a more general proxy for semantic memory. The effects of word/world knowledge on reading comprehension were direct and substantial in size (also see Cromley & Azevedo, 2007). There were also significant indirect effects of word/world knowledge on reading comprehension through the effect of knowledge on inference making, which is consistent with the longitudinal study of Silva and Cain (2015) with much younger children in which the relation of preschool vocabulary knowledge and early reading comprehension was mediated by inference skill (also see Cromley & Azevedo, 2007; Cromley et al., 2010).

Inference making had the largest direct effect on reading comprehension in older students taking method variance into account, which essentially removed the requirements for reading shared across different measures (Ahmed et al., 2016). Other studies of ninth-grade and college-age samples have also reported large direct effects of inference making on reading comprehension (Cromley & Azevedo, 2007; Cromley et al., 2010). In a study that further investigated the nature of inference-making skills using the same students in the Ahmed et al. (2016) study, those with difficulties in reading comprehension were less efficient than their better comprehending peers in integrating world knowledge while they were reading in order to maintain semantic coherence. This was the case even though the students had the knowledge needed to make the inferences, and the effect held even controlling for word-reading fluency, vocabulary and general world knowledge, and working memory (Barnes, Ahmed, Barth, & Francis, 2015).

There is also ample evidence from many studies of younger and older school-age children that inference making, whether involving the ability to link up words and ideas within the text or between knowledge and text, is deficient in less-skilled comprehenders regardless of whether such abilities are assessed using reading-time measures (e.g., Barnes et al., 2015; Barth, Barnes, Francis, Vaughn, & York, 2015), inference-eliciting questions (e.g., Cain et al., 2001, 2004b) or think-aloud procedures (Denton et al., 2015).

The findings for word/world knowledge and inference making in these studies of adolescents did not vary by comprehension skill level, suggesting that the factors that are important for reading comprehension in older children and adolescents are largely invariant across the distribution of comprehension ability. A dominance analysis study of reading comprehension in third, seventh, and 10th grades (Tighe & Schatschneider, 2014) also showed the importance of verbal knowledge as a unique and robust predictor of reading comprehension, particularly in the later grades. In the Spencer et al. (2014a) study of 425,000 children in grades 1–3, virtually all of those with reading comprehension difficulties had low vocabulary knowledge—again underlining the importance of the relation of word/world knowledge to reading comprehension across development.

Although working memory emerged as a separate factor predicting reading comprehension in all grades in the Tighe and Schatschneider (2014) dominance analysis, it was the least predictive of the factors in this study. This dominance analysis also showed that word decoding factors are dominant or strong direct predictors of reading comprehension in the earlier elementary grades. Although less of a factor in the Ahmed et al. (2016) study, which excluded students with extremely low decoding skills, a latent decoding accuracy/fluency measure had a small, but significant direct effect on reading in adolescents. This effect was more prominent in the middle school versus high school grades.

A meta-analysis of component skills and their relation to reading comprehension in struggling adult readers, found that six component skills showed the largest effects ($r \geq .5$): morphological awareness, language comprehension, fluency, oral vocabulary knowledge, word recognition, and working memory (Tighe & Schatschneider, 2016). There were not enough studies with other putative components skills such as inference, comprehension monitoring, and understanding of text structure to include in the meta-analysis.

Summary: Cognitive Processes

Whether we consider findings from longitudinal studies that follow typically or atypically developing children over the school-age years or from the preschool to school-age years, or findings from large-scale studies that model reading comprehension cross-sectionally, a coherent story emerges that is consistent with the reading systems framework discussed earlier (Perfetti & Stafura, 2014). First, knowledge about words and the world are critical for reading comprehension, showing large effect sizes in both longitudinal and cross-sectional studies and moderating the effects of reading instruction on reading comprehension outcomes. Second, other sources of linguistic knowledge such as grammar and story structure are also predictive of growth in reading comprehension, particularly for younger children. Third, text-level abilities such as comprehension monitoring and inferencing also predict growth in reading comprehension, the latter especially robust at the secondary school level.

Fourth, even in the presence of text-level skills such as comprehension monitoring and inference making, limited capacity in cognitive systems involving working memory appear to play a role in growth in reading comprehension, although their effects may be less strong than for other factors. To the extent that discourse-level skills such as inference making and comprehension monitoring require working memory and perhaps other executive functions, it may not be particularly surprising that working memory provides little or less unique prediction of reading comprehension with discourse-level skills in predictive models. There is less evidence for a strong role of other executive processes such as inhibition in reading comprehension (Barnes et al., 2016b; Christopher et al., 2012).

The final point is that word-reading accuracy and fluency are important for reading comprehension, but carry more weight in the early elementary school years. As reviewed above, word-reading difficulties limit reading comprehension, particularly in younger readers, with smaller effects for older readers. There has been a tendency to assume that older students with reading disabilities are largely those with comprehension problems rather than decoding problems or both (e.g., Biancarosa & Snow, 2006). It is worth underlining some important facts about older individuals with reading disabilities, which sometimes get lost when comparing the *relative* roles of word decoding and the other comprehension-related processes discussed above for younger versus older children. Many, but not all, children with WLRD also have more general difficulties with oral language comprehension. Given the SVR model of reading comprehension as the product of word decoding and listening comprehension, children who have limitations in both word decoding and language comprehension are likely to have significant difficulties in reading comprehension across both the elementary and secondary school years (Brasseur-Hock, Hock, Kieffer, Biancarosa, & Deshler, 2011; Cirino et al., 2013). Also, some children's word-reading deficits may first only emerge after the early elementary school years, and these older children and adolescents may struggle with word reading, reading comprehension, or both (Catts et al., 2012; Leach et al., 2003). Recall that deficits in word reading characterize a significant proportion of older school-age children with reading comprehension disabilities, as discussed in the "Prevalence" section above. That is, children identified with only word-reading disability or only SCRD represent smaller subgroups of adolescents with reading disabilities than do children with disabilities in *both* word reading and reading comprehension (e.g., Catts et al., 2006; Cirino et al., 2013). The implication for assessment is that it is important to know, even in older students, about word-reading skills, especially for differentiating reading comprehension intervention.

NEUROBIOLOGICAL FACTORS

Neurobiological studies of the neural substrates involved in reading comprehension have increased significantly over the past decade although there are still few imaging studies that address SRCD. Because text-level processing requires the integration of orthographic, phonologic, and semantic representations at the word level, imaging findings for reading comprehension overlap to a considerable extent with those of word reading (see Chapter 6).

As the review of core cognitive processes indicated, reading comprehension involves more than the compiling of words into sentences and sentences into passages; it involves the integration of sources of knowledge (word/world; grammar; text structure), text-level processes such as inference and monitoring, and cognitive resources (e.g., working memory) to construct a contextually accurate interpretation of the text. Thus, we included studies using single sentence comprehension in Chapter 6 because these studies focus on people identified with WLRD. In this chapter, we review what is known about some of the neural networks involved in different aspects of text comprehension. Most of these studies have been conducted with adults although a few have compared how adults and children process the same text stimuli. We then review the handful of neuroimaging studies that have been conducted to look at longitudinal aspects of reading comprehension and studies of text processing in individuals with SCRD.

Functional Neuroimaging Studies of Text Comprehension

A meta-analysis of functional neuroimaging studies of text comprehension (Ferstl, Neumann, Bogler, & von Cramon, 2008) reported findings in line with the notion that there are some areas of brain involved in text comprehension generally and others that are relatively more task- or process-dependent. Ferstl et al. (2008) included 23 studies of "higher-level comprehension," that is, comprehension of connected text above the single sentence level, and pragmatic comprehension such as understanding metaphors or deriving the "moral" of the story. Several different types of contrasts were used, but regardless of the type of analysis, bilateral activation of the anterior temporal lobes—generally considered areas dealing with semantic processing—was a consistent feature of text comprehension (see Figure 6.2). Because the temporal lobes are multimodal association areas of brain, it makes sense that this brain region is implicated in the integration of several types of information and sources of information during reading. Left frontomedial and parietomedial cortices were co-activated and implicated in inference making regardless of the types of materials and task used to measure inference. Although the right hemisphere has been hypothesized as a neural correlate of nonlanguage processing including inference making (Gernsbacher & Kaschak, 2003), in this meta-analysis the right hemisphere was only associated with metaphor processing, possibly reflecting the processing of novel stimuli.

Another type of functional imaging work of interest to theories of reading comprehension is based on models of situation model construction and related theories of embodied cognition (Glenberg, 1997; Zwaan & Radvansky, 1998). Readers construct mental representations of the text that simulate aspects of real-world events and that include information about character motivation/goals/emotions, spatial locations of characters and objects, temporal features of events, and causal relations between objects and human interactions with objects. In these studies changes in brain activation are measured at key points in a text such as when a character changes his or her physical location, when a character interacts with an object such as swings a tennis racket, or reacts to the actions of another character or an event. In general, these studies suggest that discourse and text comprehension is content-specific, in that areas of brain that are activated during real-world events or viewing of real-world events are also activated when reading text, that is, emotional information activates ventromedial prefrontal cortex and amygdala (e.g., Ferstl, Rinck, & von Cramon, 2005) and premotor cortex is activated when reading about a character's physical actions (e.g., Speer, Reynolds, Swallow, & Zacks, 2009).

Given findings in adult readers that situation model construction involves activating areas of brain associated with the real-world referents described by text, what do these processes look like in children with or without reading comprehension difficulties? To our knowledge, there are only two such studies conducted with typically developing children. These studies are important for what they reveal about the protracted development, even for typical readers, of the ability to automatically associate printed words with their real-world referents. In a cross-sectional fMRI study (Dekker, Mareschal, Johnson, & Sereno, 2014), 7–10-year-old children, like adults, activated regions of brain functionally associated with *pictures* of animals and tools (i.e., sensorimotor areas of brain involved in grasping for pictures of tools and occipital regions associated with biological motion and shape for pictures of animals). The *written* words for these objects activated similar regions of brain to the pictures, but only in adults. For children, there were no brain areas that selectively responded to written animal versus tool words even though pictures of these words showed such category-specific selectivity in neural response.

Further evidence that developing strong linkages between the phonological, orthographic, and semantic representations requires significant experience with reading comes from a study comparing 6- and 9-year-old typical readers. Only at 9 years of age was the visual word-form area recruited during listening tasks (Monzalvo & Dehaene-Lambertz, 2013). These findings suggest that significant experience with the printed word is needed to link up orthographic–phonological–semantic representations of words. These findings are consistent with the lexical quality hypothesis of Perfetti (2007), which suggests that high-quality, highly integrated word-level representations are critical for efficient reading comprehension. To the extent that the acquisition of high-quality lexical representations has a protracted developmental course related to reading experience, children with less exposure to print (those with difficulties in word reading, reading comprehension, or both) might be inefficient not only in reading the words, but in accessing their real-world referents important to constructing a semantically coherent representation of the text. Note that even though children with specific reading comprehension disabilities can read the words, they do not read as much as their typically developing peers (Cutting et al., 2013).

Longitudinal Studies of Typically Developing Children

Findings from the Cincinnati MRI of Neurodevelopment (C-MIND) study in which longitudinal behavioral and imaging data are available are relevant to questions about the neural markers of the development of reading comprehension. After controlling for SES, preschool children whose parents engaged in more shared book-reading experiences showed more neural activation (fMRI) in left temporal–parietal–occipital cortex, a region of brain associated with semantic processing (Hutton et al., 2015). In another study, greater frontal and supramarginal gyrus activation during narrative listening comprehension at 5–7 years was associated with reading comprehension at 11 years (Horowitz-Kraus, Vannest, & Holland, 2013). Looking over an even longer developmental time window, Horowitz-Kraus et al. (2015) found that greater activation in frontal and anterior regions of brain during a listening comprehension task at 5–7 years of age was related to higher American College Testing reading comprehension scores at 18 years of age. Sufficiently powered longitudinal studies that combine serial neuroimaging with developmentally sensitive and comprehensive behavioral assessments of reading comprehension and the core comprehension-related processes identified above are missing from the literature.

Neuroimaging Studies of SRCDs

Few neuroimaging studies focus specifically on SRCD. Cutting et al. (2013) used a lexical decision task ("Is this a real word?") to compare neural processing of pseudowords, high-frequency words, and low-frequency words in adolescents with WLRD, specific reading comprehension disability, and typically developing controls. The logic behind the use of these three types of stimuli was as follows: to the extent that individuals with SRCD have intact phonological–orthographic processing, their neural response to pseudowords (and high-frequency real words) should be similar to the typically developing group, but different from the group with WLRD. In contrast, the group with SCRD would be expected to differ from the typically developing group in their response to low-frequency words, which "stress" their semantic processing systems.

The regions-of-interest analyses revealed that typically developing and SRCD groups looked similar to each other and different from the group with WLRD in the classic dorsal and ventral pathways involved in phonological and orthographic processing of words (see Figure 6.2). Both groups had higher signal response in these areas than the group with WLRD. In contrast, the group with SRCD had greater deactivation of the left versus the right angular gyrus compared to typically developing individuals and those with WLRD.

Cutting et al. (2013) also conducted psychophysiological interaction analyses, which allow determination of whether the correlation between two brain areas differs as a function of the psychological context (i.e., low-frequency versus high-frequency words). These analyses showed a higher correlation between left IFG and several cortical and subcortical regions including left hippocampal and parahippocampal gyri in the group with

SRCD compared to their typically developing peers. This finding was interpreted to reflect an “access” deficit in which effort was required to connect phonological and orthographic representations to their semantic referents. This difference in IFG–hippocampal connectivity was specific to the group with SRCD. Because the WRLD and SRCD groups were similar on measures of Verbal IQ, the difference in connectivity for the SRCD group is not due to group differences in general verbal abilities.

These findings are consistent with those from behavioral and electrophysiological studies of individuals with SRCD. Even when these readers know something about word meanings, they have difficulties rapidly and effortlessly accessing a broad range of semantic connections about and between words. For example, children with SRCD have difficulty quickly relating semantically related words like *brother-father,* which have a more distant semantic relationship than high semantic associates such as *brother-sister* (Nation & Snowling, 1999). In electrophysiological studies, Perfetti and his colleagues have found that adults who are good decoders but less-skilled comprehenders show a smaller neural effect of this type of semantic relatedness between words, which is consistent with the idea that lexical representations for and between semantically related words are of lower quality for these individuals than they are for more skilled comprehenders (e.g., Landi & Perfetti, 2007; Perfetti & Stafura, 2014).

In contrast to studies of children and adults with WLRD (Chapter 6), far less is known about whether SRCDs are associated with differences in brain structure. Bailey, Hoeft, Aboud, and Cutting (2016) used multivariate pattern analysis to study whether there are gray matter anomalies associated with SRCDs. They compared gray matter volume in regions of interest for typically developing readers, adolescents with WLRD, and adolescents with SRCD. These regions of interest involved those found to be anomalous in individuals with WLRD, but also regions associated with two potential sources of difficulty in reading comprehension, namely, IFG areas involved in semantic processing (e.g., posterior middle temporal or dorsal IFG), and those involved in executive functions (e.g., dorsolateral prefrontal cortex).

As expected, the group with WLRD differed from the typically developing group in volumes of the left fusiform and supramarginal gyrus. The group with SRCD was differentiated from the other two groups on the basis of reduced gray matter volume in two areas within the temporal lobe (segment of inferior temporal gyrus and left superior temporal gyrus) and with multiple clusters of reduced gray matter volume in the right frontal cortex. These findings fit with early behavioral studies of less skilled adult comprehenders where similar patterns of semantic deficits were found regardless of whether the material to be comprehended was presented as text or in pictures (Gernsbacher, 1990).

Given the overlap in areas of brain that are activated during both word reading and the reading of extended text, one question that arises is how regions of brain associated with word reading and reading of passages interact and whether these interactions vary as a function of reading comprehension skill. Aboud, Bailey, Petrill, and Cutting (2016) examined functional connectivity from regions showing overlapping activation during word reading and reading comprehension. Participants were children and adolescents varying in their comprehension levels, tested on a word-level processing task (sequential presentation of one to six words at a time that were scrambled from actual short expository texts) and on a text-level processing task (sequential presentation of one to six word phrases from expository texts). During both the word-reading task and the text-reading task, typical areas associated with rapid word recognition (left occipital–temporal ventral regions), semantic processing (left IFG, middle temporal gyrus, and temporal pole), and verbal working memory (dorsolateral prefrontal cortex) were activated. The default mode network, a network of brain regions co-activated when the person is not actively engaged in processing information, was uniquely activated during text reading, suggesting lack of focus or mind wandering.

Of greater interest were findings suggesting that overlapping areas were deployed somewhat differently depending on the task and the skill of the reader. During the word-level reading task, semantic and working memory nodes showed connectivity with dorsal and ventral components of the word-reading network. In contrast, during the text-reading task these same nodes showed connectivity with the left angular gyrus, a region of brain implicated in conceptual processing and the integration of multiple sources of information for sense making. These patterns of connectivity were moderated by skill level. With respect to text-level reading, lower comprehension skill was correlated with lower connectivity of semantic and working memory areas with the left angular gyrus.

Genetic Factors

A small corpus of studies on the behavior genetics of reading comprehension has emerged in the last decade (reviewed in Olson et al., 2014). Molecular genetic studies of individuals with language impairment (not just a phonologic deficit or speech sound disorder) may be relevant for reading comprehension, but there is not enough work in this area to apply it to reading comprehension (see review in Smith et al., 2010). These studies take a dimensional perspective and generally do not identify participants with SCRD.

In the earliest study in the behavior genetics of reading comprehension, Keenan, Betjemann, Wadsworth, DeFries, and Olson (2006) found substantial genetic influences on both reading comprehension and listening comprehension. Furthermore, word recognition and listening comprehension each contributed independent genetic influences on reading comprehension; to illustrate, the genetic correlation between word recognition and listening comprehension was only .37, whereas shared environmental influences between word recognition and listening comprehension were much higher, and the genetic influences on word reading and listening comprehension were largely independent of each other. These findings have since been replicated with larger and different samples (Harlaar et al., 2010).

Recent behavior genetic studies have inquired about the longitudinal stability of genetic and environmental influences on reading comprehension and whether the genetic and environmental influences on reading comprehension change with development. Soden et al. (2015) found, similar to those who have studied the genetics of word reading, that the shared environment contributed to significant variance in early reading comprehension (grades 1 and 2), but not to later reading comprehension (grades 3–6), with steadily increasing genetic influences from the early to later elementary grades. In another study using the longitudinal TEDS sample at 7, 12, and 16 years of age, Tosto, Asbury, Mazzocco, Petrill, and Kovas (2016) found almost complete genetic overlap between oral language and reading comprehension by 12 years of age and increasing heritability in oral language between 7 and 12–16 years of age.

Christopher et al. (2015) used an international twin sample to test which prereading abilities prior to kindergarten (latent measures of phonological awareness, rapid naming, print awareness, vocabulary, and verbal memory) predicted first- and fourth-grade reading comprehension. They also evaluated the role of genetic and environmental influences on these relations. There was phenotypic and genetic stability in findings relating prereading skills to word reading and spelling in that these relations were similar at post-grade 1 and post-grade 4. Furthermore, the etiology of the longitudinal relations of most of the prereading skills with word reading and spelling across the two grades was primarily due to genetic factors. For reading comprehension, the findings varied developmentally. Phenotypically, the relation of verbal memory to reading comprehension increased in importance post-grade 4. In contrast to word reading and spelling, the relation of prereading skills to reading comprehension post-grade 4 was due to both genetic and shared environmental influences. The relation of reading comprehension with vocabulary, phonological awareness, and print knowledge was largely accounted for by shared environmental influences, while the relation of verbal memory to reading comprehension was primarily due to common genetic influences.

These behavior genetic findings are consistent with component skills models of reading comprehension such as the SVR (Hoover & Gough, 1990); even at the behavioral genetic level, reading comprehension is the product of word recognition and linguistic comprehension, and the contribution of linguistic comprehension to reading comprehension increases over time. The behavior genetic findings also converge with those from longitudinal studies of word reading and reading comprehension in that somewhat different abilities are associated with the *development* of reading comprehension versus word-reading (e.g., Oakhill & Cain, 2012).

Research looking at the genetic and environmental sources of the relation of various cognitive processes to academic skills is a relatively recent endeavor. Christopher et al. (2016) used the Colorado twin sample to investigate the genetic and environmental contributions to the overlap in working memory, inhibition, processing speed, and naming speed with word reading, reading comprehension, and listening comprehension. The relations between reading and cognitive processes were driven largely by hereditable factors. Although the genetic influences on reading were mostly shared across the cognitive measures, there was additional independent genetic variance from working memory and processing speed for both word reading and reading comprehension.

Summary: Neurobiological Factors in Reading Comprehension

In many respects, the findings from neuroimaging studies parallel the SVR, showing overlap with the triangle network involved in word reading (see Figure 6.2) and semantic areas involved in language comprehension. There is also involvement of frontal regions associated with cognitive control. These patterns are clearly apparent in the few imaging studies of SRCD, which at the very least show that WLRD and SRCD can be differentiated at the neural level. Genetic studies are mostly behavioral, but support the heritability of reading abilities. However, like the cognitive studies, they also demonstrate that word reading and reading comprehension are separable, with the latter closely linked with listening comprehension.

INTERVENTIONS FOR READING COMPREHENSION AND SRCDs

General Instructional Approaches

General approaches to reading comprehension instruction are often classified into two different types: specific skills instruction and strategy instruction (Clark & Uhry, 1995; Swanson et al., 1999). Cooperative learning has also influenced reading comprehension instruction.

As the name suggests, specific skills instruction focuses on teaching some of the component skills and sources of knowledge that can be applied to texts as discussed earlier, such as vocabulary, inference, and text structure.

Vocabulary can be taught through explicit instructional approaches or relying on more contextual approaches (NICHD, 2000). Skills such as making inferences can be taught by having children *read* short passages and answer inference-eliciting questions with or without first teaching and/or activating background knowledge, or by teaching children to find "clue" words in text that help the reader make an inference to fill in gaps in understanding. Text structures can be taught for both narrative and expository text using a variety of techniques. However, for such skills-based approaches to be effective, the teacher must provide the instruction in an explicit and systematic manner. One key is that this type of training always occurs in the context of reading.

In contrast to specific skills instruction, strategy instruction is instruction in "cognitive processes requiring decision making and critical thinking" (Clark & Uhry, 1995, p. 107). Strategy instruction in reading comprehension is an outgrowth of several cognitive psychology theories and concepts, notably schemas, metacognition, and mediated learning. For example, schemas involve the idea that a reader brings certain psychological frameworks, or "mental schemas," to a text. During reading, in order for the reader to comprehend, facts must be added or adjusted to the reader's mental schema. The study of metacognition has also had considerable influence on reading comprehension research. Wong (1991, pp. 239–240) stated that "good readers who possess meta-cognitive skills in reading are aware of the purpose of reading and differentiate between task demands. They actively seek to clarify the purposes or task demands through self-questioning prior to reading the given materials . . . [and] evaluate their own comprehension of materials read." The teaching of metacognitive strategies is beneficial to poor comprehenders even though metacognition is not causally related to comprehension skill, but is an essential part of comprehension (Perfetti et al., 2005).

Students with LDs in reading, perhaps through a history of reading texts that are difficult for them to comprehend, may have lower thresholds in their "search after meaning" (van den Broek et al., 2005). Over time, these children may have lower expectations that what they read will be semantically coherent such that they do not monitor for breakdowns in understanding and do not attempt to repair such breakdowns. The explicit teaching of metacognitive strategies addresses these issues by deeply engaging the reader in making sense of text. Again, this teaching is in the context of reading.

Finally, cooperative learning, which involves the effects of student–teacher interactions (and student–student interactions) on the student's later ability to solve problems independently, has also influenced reading comprehension theory and instruction. For example, Maria (1990) conceptualized reading instruction as an interaction between reader, text, and teacher. The reader brings decoding ability, oral vocabulary, and background knowledge to the text. The text is no longer perceived as having a single meaning for all students. Rather, meaning is constructed through this interaction. The teacher is viewed as a manager and facilitator who provides direct instruction in strategies, but who also encourages independence (Clark & Uhry, 1995). Figure 7.3, Connor's (2016) lattice model, outlines the nested context in which access to the meaning of text occurs and is facilitated by methods such as cooperative learning.

Other intervention methods based on these types of cognitive strategies have been developed to teach reading comprehension. For example, Palinscar and Brown (1985) developed a teaching method called "reciprocal teaching" found to enhance reading comprehension skills. In addition, Pressley and his colleagues have developed interventions based on "transactional strategies" to increase reading comprehension skills that are based in part on Vygotskian concepts (Pressley, 2006). In this method of instruction, students are "provided with direct instruction in a number of comprehension strategies and are encouraged to talk about and choose a strategy for understanding what they read. . . . Students are provided with positive feedback when a strategy is successful" (Clark & Uhry, 1995, p. 111). Instruction also involves teacher modeling of different comprehension strategies.

Bos and Anders (1990) developed an interactive teaching model, which is similar to Pressley's (2006) transactional teaching method, and that is also based on Vygotskian principles. This model incorporates six teaching–learning characteristics: (1) activating prior knowledge; (2) integrating new knowledge with old knowledge; (3) cooperative knowledge sharing and learning; (4) predicting, justifying, and confirming concepts and text meaning; (5) predicting, justifying, and confirming relationships between concepts; and (6) purposeful learning. Initially, a teacher models these strategies for the students, but gradually moves away from being an instructor to being more of a facilitator.

The Carnegie report (Biancarosa & Snow, 2006) made 15 recommendations for teaching reading comprehension in adolescents (see Table 7.1), beginning with the need for "direct, explicit comprehension instruction, which is instruction in the strategies and processes that proficient readers use to understand what they read, including summarizing, keeping track of one's own understanding, and a host of other practices" (p. 4). We describe some of these practices below, but the idea that the instruction is explicit is key (Chapter 5), given that many assume that simply reading broadly and frequently will in itself improve comprehension. To this point, a study of teachers whose students achieved higher and lower reading achievement scores showed that children developed better comprehension skills when instruction was explicit (Knapp, 1995, p.8): "Students do not acquire the ability to search for deeper meaning by osmosis. Teachers must structure opportunities for children to learn how to analyze and think about what they have read."

TABLE 7.1. Recommendations for Enhancing Reading Comprehension from the Carnegie Report

1. Provide explicit instruction in the strategies and processes that support comprehension.
2. Teach comprehension in content areas.
3. Self-directed learning should motivate students to read and write.
4. Support collaborative learning around a variety of texts.
5. Provide intervention in small groups for those who struggle with reading comprehension, writing, and content areas.
6. Employ diverse texts that range in difficulty level and topics.
7. Require intensive writing in all subject areas.
8. Develop technology as an instructional tool.
9. Provide assessments of student progress and program efficacy.
10. Provide extended time for literacy. In secondary schools, 2–4 hours of literacy instruction and practice in language arts and content classes is needed each day.
11. Provide ongoing professional development in literacy.
12. Evaluate student and program outcomes.
13. Create teacher teams across content areas that meet regularly.
14. Provide leadership from teachers and principals who understand reading instruction.
15. School districts should have a comprehensive, coordinated literacy plan from preschool to high school that is interdisciplinary, interdepartmental, across grade, and coordinated with

outside resources and the community.

One other approach attempts to provide intervention for those domain-general cognitive processes (e.g., working memory) thought to support reading comprehension. These studies are not reviewed as there is currently little evidence that cognitive instruction on isolated skills transfers to reading comprehension (Melby-Lervåg et al., 2016). To illustrate, teaching phonological awareness as an isolated skill and without explicit transfer of the task to letters leads to better phonological awareness skills, but little generalization to reading (Foorman et al., 1997). Such instruction has to be embedded as part of a reading program for transfer to occur. Studies on the relation of working memory to complex cognitive tasks such as nonverbal reasoning have found that strategy use mediates these relations (Unsworth & Spillers, 2010), and training encoding strategies attenuates individual differences in working memory (Robison & Unsworth, 2017). Findings such as these suggest that the nature of the relation of working memory and reading comprehension may be more complex than is typically discussed in the literature, but generalization to reading comprehension is still questionable if the instruction is not embedded into reading instruction. Studies that attempt to integrate components of cognitive intervention into skills-specific interventions for children with difficulties in reading comprehension are underway, but not yet published.

Fortunately, there are interventions that are effective specifically in the area of reading comprehension with students who vary in the extent of impairment in word recognition, fluency, and comprehension. There is strong evidence that instruction specifically targeting reading comprehension is associated with positive outcomes regardless of the source of difficulty, even in children with decoding problems. Many interventions specifically addressing reading comprehension take place at a classroom level and generally target students generically identified with LDs. People identified with SRCDs are rarely targeted by intervention studies, with the exception of a few studies in the domain of inference.

Empirical Syntheses:Reading Comprehension Strategies

In Swanson et al.'s (1999) meta-analysis, described in Chapter 6, strategy instruction was specifically effective with students with LDs who had reading comprehension difficulties. The NRP report (NICHD, 2000) identified 203 studies that involved text comprehension. However, because many of the studies had limitations in their research designs, the final database was not adequate for empirical synthesis. The 203 studies on text comprehension instruction identified 16 different types of instruction, with eight providing a firm scientific basis indicating that they improved comprehension. These included comprehension monitoring, cooperative learning, graphic and semantic organizers, instruction in story structure, question answering, question generating, summarization, and multiple strategy teaching.

In a meta-analysis of 40 studies published from 1995 to 2006 of students with LDs, Berkeley, Scruggs, and Mastopieri (2010) evaluated the effects of three classes of instruction: (1) fundamental reading skills instruction (phonics training); (2) text enhancement interventions, including supplementing text or changing text to increase comprehension, and the use of graphic organizers, technology with hypermedia, and accompanying video-based vocabulary instruction; and (3) questioning/strategy instruction, such as teaching students strategies for self-questioning, identifying story themes, identifying type of text structure, main idea strategy instruction with a self-monitoring component, activating prior knowledge, making predictions, and summarizing. Of relevance to the focus on reading comprehension interventions in this chapter, the effect sizes for text enhancements and questioning/strategy instruction were moderate in size on more proximal criterion-based measures (0.62 and 0.75, respectively), and were somewhat smaller on norm-referenced tests (0.46 and 0.48). Interventions delivered by researchers were associated with higher effect sizes than those delivered by teachers or other adults; nonetheless, reading comprehension instruction delivered by interventionists other than the researcher was still effective. The inclusion of a peer mediation component in questioning/strategy instruction was not more effective than questioning/strategy instruction alone. Similarly, the inclusion of a self-regulation component was not statistically significant in terms of moderating the effectiveness of questioning/strategy instruction, although the effects were positive. However, because studies with peer mediation and self-regulation components increased over the time period captured by the meta-analysis, the influence of these two components on strategy instruction bears reassessment in future meta-analyses with a larger corpus of studies. In another meta-analysis, Dexter and Hughes (2011) found moderate-to-large effect sizes for a variety of graphic organizers in relation to enhanced reading comprehension, with evidence of near and far transfer.

A meta-analysis of 14 experimental studies specifically examined trials that targeted children with SRCD (Lee & Tsai, 2017). The largest effect sizes were for different forms of reciprocal teaching (Palinscar & Brown, 1985), which yielded an effect size of 0.86. Interventions that targeted metalinguistic awareness and explicitly taught higher-level oral language strategies had moderate effect sizes of about 0.50. This study demonstrates that studies targeting children with SRCD are emerging and that comprehension instruction specifically improves reading comprehension in these children. However, Lee and Tsai raised concerns about the criteria used to identify SRCD and the extent to which reading comprehension was the only reading problem experienced by the children in some studies.

Given the increasing prevalence and identification of reading comprehension disabilities in older students, a synthesis of 29 studies and partial meta-analysis of 13 reading intervention studies (published between 1994 and 2004) for struggling readers in grades 6–12 is of interest (Edmonds et al., 2009). These studies were categorized as word study, fluency, multicomponent, and comprehension (e.g., explicit strategy instruction) interventions. Similar to Berkeley et al. (2010), there was a large mean effect size of 0.89, with larger effects on researcher-developed measures than on standardized reading comprehension measures (1.19 vs. 0.47). When broken down by intervention type, there were no statistically significant effects of word study and fluency interventions on reading comprehension. Effects were found for comprehension strategy and multicomponent interventions, with the largest effects for comprehension strategy interventions (1.23 vs. 0.72, respectively). Among the struggling reader population, studies with samples that involved students with disabilities showed the largest effects, underlining the effectiveness of explicit comprehension-specific strategy interventions for students with disabilities in reading comprehension (also see Scammacca et al., 2007). Unfortunately, subsequent meta-analyses with larger and better controlled studies have shown reductions in effect sizes (Scammacca et al., 2015; Figure 6.5).

Reading interventions for older students with reading comprehension disabilities are likely to be most effective when the intervention components are aligned (differentiated) with their specific reading needs. In this respect, it is worth noting that multicomponent interventions produce moderate-sized effects; however, the specific effects of word study versus comprehension-focused intervention components on word decoding, reading fluency, and reading comprehension cannot be disentangled in such studies.

In general, effects of intensive interventions (i.e., those consisting of 75 or more sessions) on reading comprehension in older students are smaller than those for younger students (e.g., Wanzek et al., 2013 vs. Wanzek & Vaughn, 2007). The reasons for this may include differences in accurate identification in younger and older children (i.e., more false positives in younger children), along with the possibility that older children in reading intervention studies are those who are persistently low readers who have not responded to interventions in earlier grades; the nature of reading comprehension tests in the later grades where high levels of background knowledge, content-specific vocabulary knowledge, and reasoning skills are required for success; and the idea that shorter interventions (as represented by studies in the other meta-analyses) may provide an initial "boost" in performance that is not maintained (Wanzek et al., 2013). Because not many of the less extensive interventions in these previous meta-analyses measured maintenance (see Berkeley et al., 2010), it is difficult to address this latter hypothesis about the source of differences in effect sizes between meta-analyses.

Examples of Research-Based Interventions

Collaborative Strategic Reading

A review of collaborative strategic reading is an interesting example of approaches that are used at the classroom level (Vaughn, Klingner, & Bryant, 2001). In collaborative strategic reading, the teacher presents strategies to the class as a whole, using modeling, role playing, and think-alouds. Students are explicitly taught to apply strategies involving why, when, and how events occur in the text they are reading. After they develop some proficiency with the strategies, they are divided into groups on the basis of their proficiency in applying the strategies. In the groups, students perform in defined roles as they collaboratively implement the strategies in expository text. In collaborative strategic reading, four strategies are taught to students, including (1) a *preview* component, in which students attempt to activate background knowledge; (2) *comprehension monitoring* during reading by identifying difficult words and concepts in the passage and using strategies that address what to do when text does not make sense; (3) restudying the *most important idea* in the paragraph; and (4) *summarization/question asking*. The results of several studies showed that many students made significant gains in reading comprehension and academic content. However, some students showed little response, highlighting the importance of carefully monitoring the progress of students with LDs who are receiving a classroom-based intervention.

Promoting Adolescents' Comprehension of Text Program

In a series of studies, Vaughn, Wanzek, and colleagues (e.g., Vaughn et al., 2012; Wanzek et al., 2016) have implemented the Promoting Adolescents' Comprehension of Text (PACT) reading comprehension intervention in middle school and high school social studies classrooms. These interventions at the classroom level involve team-based learning and knowledge application (Michaelsen & Sweet, 2011), which is a form of collaborative learning; activation of background knowledge; teaching of essential words using a variety of explicit verbal and visual strategies and supports; critical reading during which time teachers stop students at key points to facilitate discussion and understanding; and team-based knowledge/comprehension checks that provide an opportunity for practice/retrieval, further peer discussion, and feedback. At the whole-classroom level, this intervention resulted in small-to-moderate effects on the acquisition of content knowledge and content-area comprehension. For students with reading disabilities, the effects were often larger compared to those for typical readers even though students with disabilities were being instructed in the same classrooms as their nondisabled peers (Swanson, Wanzek, Vaughn, Roberts, & Fall, 2015). Similar to the meta-analyses above, there were only small effects (often not statistically significant) on standardized tests of global reading comprehension.

These content-area interventions display several of the features of strategy instruction and collaborative learning. However, this type of instruction might also be thought of as text-based instruction meant to increase content-area knowledge. Given the importance of world knowledge (Ahmed et al., 2016) as well as content-specific academic vocabulary (Lesaux, Kieffer, Kelley, & Harris, 2014) for reading comprehension, such knowledge-building interventions in specific content areas are worth further study (see examples of such approaches in Lesaux et al., 2014, and Lawrence, Crosson, Paré-Blagoev, & Snow, 2015).

Peer-Assisted Learning Strategies

In a similar way, work on peer-assisted learning strategies (PALS) for reading in grades 2–6, in which the instructional focus is on comprehension strategies, has documented moderate effects on reading comprehension for some students with LDs, as well as for their low-, average-, and high-performing classmates, in settings where English is the dominant language (D. Fuchs, L. S. Fuchs, Mathes, & Simmons, 1997) and where Spanish is the dominant language (Saenz, Fuchs, & Fuchs, 2005; also see the What Works Clearinghouse (2012) report. PALS for reading employs several of the elements of comprehension strategy instruction described above, including peer-mediated learning where pairs of students work together, with a stronger reader and a less skilled reader who take turns being the coach and the reader; partner reading and retelling where the coach listens to and corrects the reader's errors and then asks the reader to retell what he or she has learned; paragraph shrinking, which is a main idea exercise; and prediction relay, in which the reader predicts what will happen and then verifies it with further reading. As with Collaborative Strategic Reading (Vaughn et al., 2001), an unacceptable proportion of students with LDs demonstrate insufficient response to PALS.

Learning Strategies Curriculum

A long-term program of research from the Center for Research on Learning at Kansas University (Schumaker, Deshler, & McKnight, 2002) has identified a series of strategies, or teaching routines, that impact not only the learning of students with LDs, but all students in the classroom. These teaching routines involve a variety of domains, including reading comprehension and writing, as well as a variety of organizational skills in school and out of school (e.g., homework). Largely implemented in secondary school environments and at a classroom level, these routines have been organized into the Learning Strategies Curriculum, which focuses on three major demands presented by standard curriculum: acquisition, storage, and expression of information. For acquisition, the teaching routines involve strategies that facilitate word recognition and reading comprehension (paraphrasing, visual imagery, recall of narrative text, self-questioning, and related strategic activities).

A series of research studies, many of them involving probe assessments in single-case designs, have shown that adolescents with LDs can be taught complex learning strategies and that implementation of these strategies results in improved academic performance (Schumaker et al., 2002). Effect sizes are consistently in the large range for various strategies. Studies that involve classroom-level instruction of organizational skills not only show that such instruction improves organizational skills and overall performance in students with LDs, but also reveal that students without LDs who are showed these strategies also improve with explicit instruction in this domain (Hughes, Ruhl, Schumaker, & Deshler, 2002).

Vocabulary Interventions

Given the strong longitudinal relations between vocabulary and reading comprehension and large direct and indirect effects of vocabulary knowledge on reading comprehension, researchers have focused on vocabulary as a specific intervention target. There are several hypotheses about why vocabulary and reading comprehension are related, including, but not limited to, direct unidirectional effects of vocabulary knowledge on comprehension, reciprocal causal effects between vocabulary and reading comprehension, and a third variable such as general verbal ability that affects vocabulary and reading comprehension independently (Nagy & Scott, 2000). The first and second hypotheses would lead to the prediction that vocabulary interventions should result in direct changes in reading comprehension.

The NRP (NICHD, 2000) report identified 47 studies involving vocabulary instruction. However, because many of the studies had limitations in their research designs, the final database was not adequate for empirical synthesis. It was difficult to separate and classify the many different variables and methodologies included in experimental research involving vocabulary instruction. Based on a narrative review of the best available evidence, the report made several recommendations for vocabulary instruction including: (1) explicit instruction for words required to understand specific texts; (2) repetition and opportunities for multiple exposures of words likely to be encountered across several contexts; (3) exposure to newly taught words in a variety of rich authentic contexts; (4) restructuring of vocabulary tasks to support low-achieving students and students at risk; (5) active, deep engagement in new word learning; (6) the potential usefulness of computer-aided instruction; (7) that new words can be learned incidentally, but that there are certain conditions that promote such incidental learning; and (8) the use of a variety of instructional methods. The NRP recommended that vocabulary instruction be a core component of reading comprehension instruction, which was novel at the time because researchers and practitioners did not consider vocabulary to be an interesting and important literacy topic (Nagy & Scott, 2000). Many questions were not addressable, including the relative effectiveness of different types of instructional approaches, whether instruction had an effect on text-level reading comprehension (which is the purported purpose of providing vocabulary instruction), and whether participant demographics (e.g., age) moderate effectiveness. The report could not look specifically at studies of students with LDs.

A systematic review of vocabulary interventions for students with LDs that included both group and single-subject designs (Jitendra, Edwards, Sacks, & Jacobson, 2004) found evidence of effectiveness for several types of vocabulary instruction. These included (1) *keyword/mnemonic strategies* (use of a "keyword" that sounds similar to the new word accompanied by images of the meaning of the known keyword interacting with an image of the meaning of the new word); (2) *cognitive strategy instruction* exemplified by the work of Bos and Anders (1990) using semantic feature analysis, semantic/syntactic feature analysis, and semantic mapping—in this approach, words are categorized according to overlapping and differentiating semantic features; and (3) *explicit instruction* (see Chapter 5) in which word meaning is explicitly provided and the teacher checks for understanding, promotes students' use of the new words, and fosters more independent learning. Due to limitations in the primary studies, however (e.g., variations in the type of comparison group, wide variation in dependent measures), some of which were also noted in the NRP for studies of typically developing readers, it was difficult to draw conclusions about the relative effectiveness of these various approaches for students with LDs. Furthermore, the effects were mainly for researcher-constructed measures of taught words and so it is unclear whether these interventions lead to improvements in text-level reading comprehension for students with LDs; in fact, there were very few studies that assessed the effects of these vocabulary interventions on reading comprehension.

In contrast, a meta-analysis from PreK to 12th grade (with half of the studies conducted in grades 3–5) looked explicitly at the effects of vocabulary interventions on passage-level comprehension (Elleman, Lindo, Morphy, & Compton, 2009). Effects of vocabulary instruction were small and not statistically significant on norm-referenced passage-level comprehension tests. The effect for researcher-created passage-level comprehension measures aligned with instruction, was significant and moderate (d = 0.50). In what will become a common theme across meta-analyses of specific skills instruction, the mean effect size for passage-level comprehension was much larger for students with reading disabilities (d = 1.23) than for those without reading disability (d = 0.39). Elleman et al. (2009) also looked at the effects of vocabulary instruction on the acquisition of vocabulary. Here, there were significant effects on norm-referenced and proximal researcher-created vocabulary measures (d = 0.29 and 0.79, respectively), but effect sizes on vocabulary did not differ for students with and without reading disabilities. It is worth noting that the vocabulary interventions with students with LD that produced these large effects on passage-level reading involved moderate-to-high levels of small-group discussion. Similar effects tied to discussion are obtained for students without LD using the Word Generation Program (Lawrence et al., 2015). Several studies employed semantic feature analysis, suggesting the importance of depth of processing in vocabulary instruction as well as collateral effects on increasing content-specific domain knowledge through discussion and semantic mapping. Furthermore, extended discussion of text meaning may have effects on learning by increasing motivation and engagement (Kamil et al., 2008).

Given that early vocabulary knowledge affects later reading comprehension, studies on the effects of vocabulary interventions for young children are of interest. A meta-analysis of oral vocabulary instruction for children in pre-kindergarten and kindergarten (Marulis & Neuman, 2010) found larger effects on researcher-created versus norm-referenced measures of vocabulary (g = 1.21 vs. 0.71). Explicit vocabulary instruction and the combination of explicit and implicit instructional methods were associated with stronger effects than implicit instruction alone (g = 1.11, 1.21, and 0.62, respectively). Interventions delivered by researchers and teachers had similar large effects, though those delivered by parents were also effective. The effectiveness of vocabulary interventions did not differ for children from middle-class versus low SES homes; however, vocabulary interventions were less effective for children who were low SES and who had another risk factor for learning such as low achievement or a special education designation (e.g., an IEP) than they were for middle-income peers with these same learning risk factors (g = 0.77 vs. 1.35). Such information suggests that vocabulary interventions for children with more than one risk factor for learning (e.g., socioeconomic disadvantage and early learning difficulties) may need to be more intensive and explicit than what is offered by current early vocabulary interventions.

Explicit vocabulary interventions have educationally relevant effects on vocabulary acquisition in younger and older children at risk for or with LDs and on understanding text that uses newly instructed words. It is worth keeping in mind, however, that most vocabulary interventions provide instruction on a limited set of words; when effects are large, it is for instructed words and texts containing instructed words. The acquisition of word knowledge is complex; for example, learning of new word meanings can be an incremental process taking many exposures across multiple contexts (Nagy &

Scott, 2000). Direct instruction is meant to shorten this incidental and incremental learning process. It is estimated that students without disabilities increase their reading vocabularies by thousands of words a year (Nagy, Herman, & Anderson, 1985), partly through incidental word learning during reading (see meta-analysis of contextual word learning by Swanborn & De Glopper, 1999). When one thinks about vocabulary instruction for students with comprehension difficulties, there is clearly a tension between the number of new words that students need to learn each year and the number of words that vocabulary interventions have been shown to deliver. Compton, Miller, Elleman, and Steacy (2014) argued that current comprehension interventions, including those for vocabulary, are strikingly nongenerative in their effects. In contrast, the acquisition of word knowledge in typical oral language development is highly generative; so too is the acquisition of word knowledge from text in which learning the meaning of a new word can facilitate the learning of other new words in the same text or even across different texts (reviewed in Nagy & Scott, 2000). Compton et al. (2014) suggest that computational models involving latent semantic analysis (i.e., revealing semantic connections between words and concepts from analyses of large bodies of text) and the explicit teaching of "microworld" knowledge based on such analyses may be one way to better mimic and provide explicit instructional materials for the type of contextually based vocabulary learning that occurs in natural language development.

Inference Interventions

Although inference making is included both implicitly (e.g., "why" questions) and explicitly (e.g., prediction relay activity in PALS) in many of the broader comprehension strategy instruction programs discussed above, strategy instruction does not primarily focus on the making of inferences; indeed, it is difficult to determine what proportion of strategy-based instruction involves the making of inferences that are important for maintaining semantic coherence. Compton et al. (2014) argued for a new generation of reading research that is more strongly tied to the theoretical literature on how reading develops and that more explicitly targets the sources of knowledge and skills that support typical reading development and that are deficient in children with reading comprehension difficulties. They identify inferencing as one of these skills that, according to all cognitive models of reading comprehension, is necessary for constructing both a semantically and referentially coherent text-based representation and situation model. The situation model is critical for learning from text (Kintsch, 1988). In the studies discussed earlier in this chapter, inference-making that maintains semantic coherence is a strong direct predictor of reading comprehension in adolescents, and inference-making skills are uniquely predictive of later reading comprehension in longitudinal studies.

Despite the importance of inference making for reading comprehension and its status as a core deficit in individuals with reading comprehension disabilities, there are relatively few inference-making interventions for either typically developing students or for students with reading disabilities. Most studies were conducted in the 1980s and 1990s. Hall's (2016) synthesis of intervention studies for children with reading comprehension difficulties showed some evidence of effectiveness despite wide variation in the type of inference taught (e.g., text connecting vs. knowledge-text inferences) and the instructional method for how inferences were taught (e.g., prior knowledge activation and prediction vs. key word method vs. practice answering inference questions).

A more recent meta-analysis of inference-making interventions that was not confined to students with LDs similarly found evidence for the effectiveness of such interventions (Elleman, 2017). To illustrate the wide variation in procedures used to instruct inference making, the 25 studies in the Elleman (2017) meta-analysis variously employed: (1) inference question practice and generation; (2) the clue word or key word method in which students answer inference questions and underline and sometimes discuss the clue words or phrases that support their answers; (3) activation of background knowledge with strategies for how to fill in the gaps in understanding; (4) graphic organizers; (5) identification of different kinds of inferences with practice on each type; (6) use of self-explanation and elaboration to make inferences using knowledge; and (7) use of character motives and putting oneself in the shoes of characters to make inferences about what characters do and why.

The majority of studies employed clue word and knowledge activation procedures. In contrast to many of the strategy instruction studies in the comprehension strategy meta-analyses above, 70% of these inference interventions lasted 10 hours or less, which is a relatively nonintensive intervention. Few studies instructed more than one type of inference and, as with other comprehension intervention studies, the use of proximal researcher-created measures was much more common than the use of norm-referenced reading comprehension measures.

Overall effects of inference instruction on general comprehension, inferential comprehension, and literal comprehension were d = 0.58, 0.68, and 0.28, respectively. Similar to the findings for vocabulary instruction, effects for less skilled readers were larger than those for skilled readers on measures of inferential comprehension (d = 0.80 vs. 0.55). Interestingly, there was also a large effect for less skilled readers on literal comprehension, whereas inference intervention did not improve the literal text comprehension of more skilled readers (d = 0.97 vs. 0.06). Why inference instruction generalizes to literal comprehension for less skilled readers is unknown, but is a positive, if unexpected, benefit. Regardless of the mechanism at work, making inferences is likely to be much easier when the representation of the literal meaning of the text is accurate and semantically coherent.

Although we have included inference interventions under skills-specific instruction, most of these interventions included explicit instruction in *how* to make inferences, and used procedures similar to some of the comprehension strategy instruction interventions described earlier. Furthermore, about half of these studies included a comprehension-monitoring component as part of the intervention (Elleman, 2017). Similar to what was found for vocabulary interventions and in the Hall (2016) inference intervention synthesis, the data were not sufficient to test the relative effectiveness of different approaches to inference intervention or to specify essential and nonessential components of inference-making instruction. Nor were there enough studies to distinguish between effects for children with SRCDs versus for those with more generalized difficulties in reading.

An example of an inference-specific intervention designed for secondary school students with reading disabilities comes from the work of Fritschmann, Deshler, and Schumaker (2007). INFER uses a combination of clue word identification and activation of background knowledge procedures as student work through passage reading and the answering of inference questions. INFER consists of five steps:

1. ***I**nteract with the passage and the questions* involves students in previewing the passage, reading the questions, and mentally categorizing the question as factual or "think and seek" (i.e., inferential).
2. ***N**ote what you know* involves students in further classifying the inference questions into purpose, main idea/summarization, prediction, and clarification questions at the same time as using the questions to activate knowledge and underline key words in the questions that told the students what information to look for in the passage.
3. ***F**ind the clues* involves students in finding and underlining the clues in the passage related to the key words they identified in the questions in step 2 and creating tentative mental answers to the questions.
4. ***E**xplore more details* prompts students to look for additional clues in the text that would support their tentative answers.
5. ***R**eturn to the question* requires students to go back to each question to ensure that an answer had been selected.

It is clear, that even though the focus is on inference making, children are being taught an explicit and systematic set of strategies or procedures to help them make inferences. This intervention was tested in a single-subject design with several 9th graders with LDs. Large effect sizes were found for these students.

Text Structure Interventions

Narrative text structures (character, setting, plot, solution) may be more familiar to children than knowledge of other types of text structure because of early and frequent exposure to stories in video-based, oral, and written formats. Expository text structures are typically less familiar to children, even though learning from expository text in social studies and science becomes increasingly important after the early primary grades. In the Reading Systems Framework (Perfetti et al., 2005), text structure is hypothesized to be one of the important sources of knowledge implicated in reading comprehension. Meyer (1987) identified five expository text structures: description, sequence, compare–contrast, cause–effect, and problem–solution. Text structure interventions are explicit interventions either in reading, writing, or both in which specific text structures (one or more of these five types) are instructed using a variety of methods including graphic organizers, identifying text structures in reading, producing text structures in writing, and identifying and using signal words associated with the five types of text structure (e.g., *because* in cause–effect).

A meta-analysis of 45 studies of expository text structure interventions (Hebert, Bohaty, Nelson, & Brown, 2016) tested a number of factors hypothesized to moderate the effectiveness of such interventions. Teaching of *more* text structures was associated with larger effects on expository text comprehension than was teaching of only one structure. The inclusion of a writing component in instruction was also associated with larger effect sizes. Instruction in expository text structure was effective for both elementary and secondary school students. The effect size for students with disabilities was large (0.96); however, effect sizes were smaller in studies where the comparison group received an alternate comprehension intervention. These findings suggest that explicit instruction in expository text structure is associated with improvements in expository text comprehension for students with disabilities, but whether it is more or similarly effective as other explicit comprehension interventions is not clear.

One example of a text structure program for narrative text that has been used for students with LDs is the *Theme-Identification Program,* which is presented in Figure 7.4. Williams and colleagues have completed studies that focused on middle school students with LDs (e.g., Wilder & Williams, 2001) as well as second- and third-grade students (Williams et al., 2005). The goal of the program is to help students derive themes, the overall meaning of stories abstracted from the specific plot components. In the Theme-Identification Program, two introductory sessions focus on plot components; the remaining

12 lessons address the identification of a story's theme. Each lesson is organized around a single story and includes prereading discussion of the theme concept; reading the story aloud; discussing the important story information; using organizing questions as a guide (i.e., the "theme scheme"); transfer and application of the theme to other story examples and real-life situations; review; and activity.

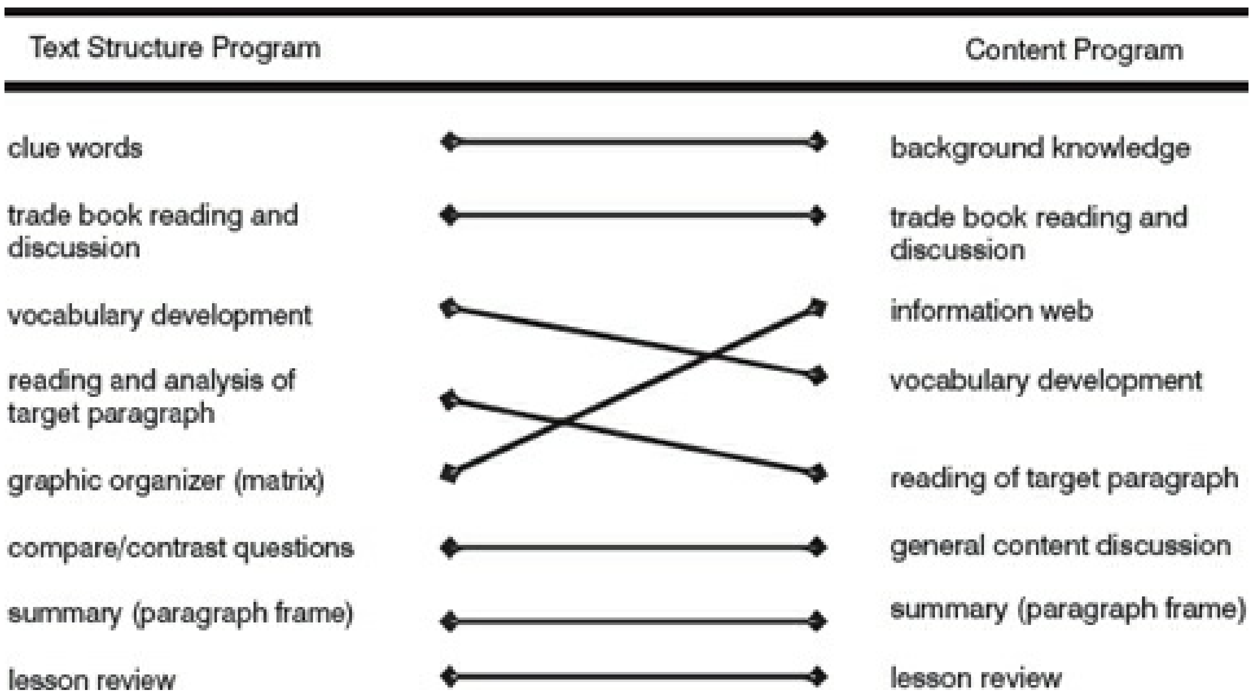

FIGURE 7.4. An overview of the Theme-Identification and content programs. Solid lines with diamond end points indicate comparable sections in each of the two instructional programs. From Williams et al. (2005, p. 541). Copyright © 2005 the American Psychological Association. Reprinted by permission.

The heart of the program is the theme scheme, which provides a set of questions that organize the important story components to help students follow the plot and derive the theme. The teacher models how to answer the eight questions leading to a theme, and students gradually assume increasing responsibility for asking the questions and identifying the theme. In addition, the students rehearse and commit to memory these questions so they can apply the theme scheme guide to untaught stories. Toward the end of instruction, transfer instruction is provided in an explicit manner, with two additional questions employed to help students generalize the theme to other relevant situations. Studies that use this program at the whole-class level (e.g., Williams et al., 2002) have shown benefits for high-, average-, and low-performing students relative to their classmates receiving more traditional programs that emphasize vocabulary and plot.

In similar studies that instruct text structures important for understanding expository text, Williams and her colleagues have shown that (1) these expository text structure interventions are effective even for children in the early elementary grades, for example, for social studies and science in second and third graders (Williams, 2005; Williams et al., 2014); (2) are consistent with the findings from the meta-analysis above, such that the teaching of more text structures results in bigger effects and the effects of teaching one text structure such as compare–contrast does not transfer to better understanding of expository text with other text structures such as cause–effect; and (3) using instructional time to include explicit teaching of expository text structure does not result in less learning of content-area knowledge compared to comparison groups that receive more traditional instruction with the same content. Thus, the additional advantage of text structure instruction is its potential for generalization to new texts in new content areas without sacrificing the learning of specific academic content in social studies or science.

Summary: Reading Comprehension Interventions

Table 7.2 summarizes essential principles for teaching reading comprehension to children with LDs. The importance of explicit, multicomponent instruction and engagement in reading are clearly apparent. The empirical syntheses show that both skills-specific and comprehension strategy interventions are effective for improving comprehension in children with reading comprehension difficulties; however, the field lacks studies that assess the relative effectiveness of various approaches. In addition, there is little data on whether these interventions are additive or at least stronger in a multicomponent reading comprehension intervention. When these research-based interventions are implemented at the classroom level, which is the case in many of the studies cited above, students with LDs gain as much as, and often more than, their typically developing peers from such instruction. This is an important point if we think that the function of interventions for students with LDs is to not only improve academic outcomes, but to also close the gap with their typically achieving peers. Connor et al. (2009) found that for teaching comprehension at the classroom level, there was little difference in teaching to the entire classroom versus forming small groups for supplemental instruction in the classroom; this was not the case for code-based instruction, where small-group instruction was more effective than whole-classroom instruction. Note that even in Connor's work, both code-based and meaning-based instruction is provided, and students receive differentiation in both areas. This parallels what is known about Tier 1 and Tier 2 instruction: the strongest effects occur when instruction is differentiated and addresses reading as a complex skill requiring multiple components of instruction, that is, decoding, fluency, and comprehension (Foorman et al., 2017). Importantly, the most effective instruction for children with reading comprehension difficulties, whether for skills-specific interventions or comprehension strategy interventions, is explicit rather than incidental or purely based on exposure, and occurs in the context of actual reading of connected text, not as isolated training independent of academic content.

TABLE 7.2. Interventions: Fundamental Principles for Teaching Reading Comprehension

1. Explicitly teach comprehension skills and strategies such as question generation and question answering (literal and inferential questions) and comprehension monitoring.
2. Model comprehension strategies using "think-alouds"; scaffold students in the use of strategies and gradually transfer strategy use to students during independent and collaborative reading; provide organizational support to facilitate self-regulation.
3. Make use of cooperative learning to encourage engagement with text through discussion and question asking-and-answering routines.
4. Explicitly teach sources of knowledge important for reading comprehension including: (a) text structure for narrative text and for different types of informational texts in social studies and science (e.g., cause–effect; compare–contrast; sequence; problem–solution); and (b) word and world knowledge necessary for understanding texts in core academic content areas.
5. Work on oral language development throughout schooling, including vocabulary and morphology.
6. Teach learning adjuncts in content, such as using graphic organizers to organize the ideas in a text and their relations to each other in history, social studies, and other areas.
7. Engage learners in reading a broad range of high-interest texts scaffolded to their reading

level.

The need to involve content teachers (e.g., those who teach history, science, language arts) in explicit reading comprehension instruction is critical, as considerable time will be needed to enhance comprehension abilities in all students. Engaging older students, whose interest in schooling in general often diminishes in the face of other interests, is also very important. The finding that comprehension interventions produce stronger effects on more proximal measures of comprehension-related outcomes than on norm-referenced measures may be partly due to limitations with norm-referenced assessments of reading comprehension (see Chapter 5), which may lack sensitivity and deep coverage in the measurement of the knowledge and skills that might be changing as a function of intervention.

Finally, although there are limitations of current interventions to produce generative learning in terms of transfer to untaught materials and maintenance of effects over time, there is considerable evidence for the effectiveness of comprehension-specific interventions for students with disabilities in reading comprehension. There are medium-to-large effects for students with LDs in some meta-analyses for strategy use, text structure interventions, vocabulary intervention, and inference intervention. However, what may be driving improved comprehension in each type of intervention are unidentified common variables such as increased engagement and deeper, more effortful processing of text. This topic clearly needs more investigation and may help account for the finding that students with LDs improve significantly in comprehension through whole-classroom instruction that is explicit and comprehensive.

CONCLUSIONS: SRCDs

It is clear that there are children whose primary problems in reading reflect difficulties in comprehension, rather than decoding and fluency, and that oral language disorders do not account for all comprehension difficulties. Many issues remain concerning the measurement of reading comprehension, especially since determinations of level and quality depend on how it is assessed (Fletcher, 2005). There simply are not enough latent variable studies of reading comprehension. In addition, more research needs to focus on children identified with SCRD, especially from neurobiological and instructional perspectives. Neurobiological studies of poor comprehenders are scant, and much of the understanding of this aspect of poor reading comprehension builds on studies of children with lower-level difficulties. Few intervention studies identify poor comprehenders as a specific subgroup.

However, studies involving intervention in reading comprehension instruction show that comprehension can be improved, even in students with WLRD. In these studies, much of the impact on gains in reading comprehension stems from strategy instruction often included as part of a comprehensive approach to reading instruction in children with word recognition and fluency difficulties. It is well known that students with LDs in a variety of domains do not spontaneously identify strategies, and, if they are taught strategies, they do not implement them in the absence of specific instruction that promotes generalization. Strategic instruction promotes self-regulation and raises the student's level of independence. Such instruction addresses the "executive function" deficiencies commonly observed in students with LDs in a variety of academic domains. More research is needed to determine how to better promote independent use and transfer of taught comprehension skills to new reading contexts, which is critical for enabling students with LDs to be able to use *reading to learn*.

CHAPTER 8

Mathematics Disabilities

Calculation and Problem Solving

In turning to LDs in mathematics, we identify two broad domains of competency from Table 4.1. One domain involves computational skills, epitomized by paper-and-pencil arithmetic; the other is problem solving, epitomized by word problems. Unlike reading, where differentiations of decoding versus listening and reading comprehension are supported by frameworks like the SVR (Hoover & Gough, 1990), there is no simple view of mathematics and no consensus about the broad domains of impairment in people with LDs in mathematics. Although there is a phenotype with adequate reading and poor mathematics, many demonstrate problems with both domains (Figure 2.2; Rourke & Finlayson, 1978; Willcutt et al., 2013). This overlap is generally considered to reflect comorbidity, but there are different views about the overlap of reading and mathematics LDs (Ashkenazi, Black, Abrams, Hoeft, & Menon, 2013). As in word- and text-reading LDs, the cognitive correlates vary with the domain of mathematical competency (calculation vs. problem solving), but also according to specific mathematical knowledge and procedures (addition and subtraction, fractions, algebra; Geary, 2005). Concepts of numerosity are also important, involving comparisons of magnitudes, the representation of quantities, and symbolic extensions. None of the cognitive correlates of mathematics LDs show a unique relation like that of phonological awareness to word reading (Chen & Li, 2014). Math development and instruction are hierarchical; like the relation of decoding to comprehension, basic skills predict more complex skills.

Several aspects of mathematical learning involve domain-general skills, such as working memory (WM) and cognitive control.

In this chapter, we discuss these different perspectives for broad competencies and for specific numerical skills in the context of LDs and typical development. We also address neurobiological correlates and instruction. One difference in this chapter relative to reading is that we more directly address curriculum issues because mathematics curricula are not as closely aligned with research.

DEFINITION AND CLASSIFICATION

Definitions of LDs in mathematics have developed more like the exclusionary definitions historically characteristic of WLRD. As noted in Chapter 2, the U.S. federal statutory and regulatory definitions of LDs refers to disabilities in mathematical calculations and problem solving (IDEA, 2004), whereas the National Joint Committee on Learning Disabilities (NJCLD, 1988) definition of LDs refers to significant difficulties in “reasoning and/or mathematical skills.” DSM-5 (American Psychiatric Association, 2013a) uses the term “specific learning disorder” (SLD) with a specifier—that is, SLD *with impairment in mathematics*. Further descriptors may include difficulties in number sense, accuracy and fluency in arithmetic facts, calculation, and mathematical reasoning. IQ–discrepancy definitions have been used, but as discussed in Chapter 3, have not shown strong validity and have diminished in use.

The International Classification of Diseases–10 (ICD-10; World Health Organization, 2013) provides research criteria for the identification of a “specific disorder of arithmetical skills” that includes developmental acalculia (also referred to as dyscalculia). ICD-10 excludes arithmetical problems due to a reading or spelling disorder and difficulties in math secondary to an intellectual disability or poor teaching:

> Involves a specific impairment in arithmetical skills that is not solely explicable on the basis of general mental retardation or of inadequate schooling. The deficit concerns mastery of basic calculation skills of addition, subtraction, multiplication, and division rather than of the more abstract mathematical skills involved in algebra, trigonometry, geometry, or calculus. (*http://apps.who.int/classifications/icd10/browse/2015/en#/F81.2*)

Because of this persistent vagueness and the parochial nature of the quality of extant definitions, no consistent standards have been established by which to judge the presence or absence of LDs in math. Adding to this dilemma is the fact that “LDs in mathematics,” “developmental arithmetic disorder,” “developmental dyscalculia,” “mathematics disabilities,” and “specific mathematics disabilities” are broad terms used for a variety of impairments in mathematics skills ranging from calculations to problem solving. As Fleishner (1994) suggested, in some cases the term “mathematics learning disability” has been used synonymously with the term “dyscalculia” to denote *specific* (as opposed to generalized) deficits in calculation or mathematical thinking. *Specific* usually implies that oral language, reading, and writing are intact (Rourke & Finlayson, 1978; World Health Organization, 2013). However, math deficits are frequently associated with other LDs (Ashkenazi et al., 2013; Fleishner, 1994; L. S. Fuchs, D. Fuchs, & Prentice, 2004; Willcutt et al., 2013) and children with LDs in calculation with and without WLRDs respond similarly to math interventions (L. S. Fuchs et al., 2008a). If reading is involved, the math (and reading) problem is more severe than if either occurs in isolation (Figure 2.2).

The term “developmental dyscalculia” has regained some prominence in the math cognition research community and is particularly prevalent in developmental science and neuroscience research. This resurgence may be related to the significant interest and increase in studies attempting to identify a “core deficit” in number sense among individuals with LDs in mathematics (see below). In these more recent conceptualizations, developmental dyscalculia has been described as not identical to arithmetic and mathematical disabilities, although heterogeneity with other LDs and co-occurring difficulties in domain-general abilities is acknowledged (Kaufmann, Wood, Rubinsten, & Henik, 2011). Because of the many definitions of dyscalculia in the literature, the DSM-5 Neurodevelopmental Work Group decided against using this term. Whether developmental dyscalculia can be reliably distinguished from mathematical disabilities more generally or other types of disabilities within the domain of mathematics is an empirical question for which there is little evidence at the present time.

The ICD-10 and DSM-5 definitions focus on calculation difficulties, and ICD-10 seems to exclude “math reasoning” difficulties, which were part of

IDEA until 2004 and were then replaced by "mathematics problem-solving skills." We believe there is evidence supporting a distinction between calculation and problem-solving skills, the latter representing word problems with complex cognitive demands for reasoning. Although there is clearly overlap in the cognitive correlates, some children have problems in either one or both domains (L. S. Fuchs et al., 2006a). At the curriculum and intervention level, these domains are clearly separable, so we will focus on calculation and problem-solving distinctions, justifying it in part by a review of component skills in mathematics and examination of cognitive correlates and intervention.

Prevalence

Depending on definition and thresholds (cut points), LDs involving mathematics have been found to be about as prevalent as disabilities involving reading. In a review, Fleishner (1994) found that studies of the prevalence of math LDs have produced similar estimates. More recent studies give estimates of 5–6% (Shalev et al., 2000) and 3.6% (Lewis et al., 1994). The latter study broke prevalence into those who had only arithmetic disability (1.3%) and those with both arithmetic and reading disabilities (2.3%). These estimates contrasted with 3.9% for specific reading disabilities.

Studies that are European in origin have cut points that tend to be more stringent (below the 5th percentile) than in North American studies. A North American study (Barbaresi, Katusic, Colligan, Weaver, & Jacobsen, 2005) utilized three definitions of LDs in math (regression-based IQ–achievement discrepancy, unadjusted discrepancy of IQ and math achievement, and low achievement in math). The prevalence of math disability in this unselected birth cohort ranged from 5.9 to 13.8%, depending on the definition. Specific LDs in math that did not involve reading occurred in about one-third to one-half of the sample of those with math LDs, depending on the definition. L. S. Fuchs et al. (2005a), in a study of 564 first-graders, compared prevalence estimates based on 17 identification methods grouped into four categories: IQ–achievement discrepancy, low achievement with average IQ, and two categories tied to intervention response and growth over time that utilized benchmarks based on low achievement or change over time. Not surprisingly, prevalence estimates varied considerably within and across categories. Definitions based on IQ–achievement discrepancy yielded a prevalence rate of 1.77%. A low achievement definition with a cut point at the 10th percentile yielded a prevalence rate of 9.75%, consistent with a dimensional perspective. Methods based on response to instruction benchmarks yielded prevalence of less than 1% if final low achievement was the benchmark and a standardized test was used, and 6–9% if the low achievement benchmark was derived from a curriculum-based assessment. Estimates based on both slope and low achievement yielded a prevalence rate of about 4%. As we discussed in Chapter 3, all of these estimates reflected decisions about cut points, the role of intervention, and the type of math skill that was assessed. The attributes of mathematics LDs are also dimensional and normally distributed.

Sex Ratio

Most studies have not found sex differences in the prevalence of LDs in math (Shaley et al., 2000), although Barbaresi et al. (2005) did find male preponderance ratios of 1.6 to 2.2:1, depending on the definition. Because the latter study depended on access to records documenting a disability, there is always the possibility of bias in terms of who was referred for an evaluation. Another approach is to investigate the sex ratio for math disability that is not accompanied by reading disability. Devine, Soltész, Nobes, Goswami, and Szücs (2013) found no sex differences for children between 7 and 10 years of age with low math and either average or above average reading using both absolute threshold criteria and relative discrepancy criteria between math and reading to identify the math LDs. There were also no sex differences between two different types of math (e.g., calculation vs. understanding shape). Spelke (2005) reviewed literature on sex differences in mathematical ability, failing to find differences between males and females in cognitive and neurobiological mechanisms at a variety of age levels. Furthermore, behavior genetic studies show no evidence for differential sex-related genetic or environmental contributions to mathematics ability/disability on either math achievement tasks or measures of magnitude processing (Kovas, Haworth, Petrill, & Plomin, 2007; Petrill, Kovas, Hart, Thompson, & Plomin, 2009; Tosto et al., 2014). As we discussed in Chapter 6, it is possible that any differences in sex ratio are due to differences in the distributions and variance of math-related skills between males and females.

Developmental Course and Outcomes

A 3-year longitudinal study in grades 4–8 (Shalev, Manor, Auerbach, & Gross-Tsur, 1998) reported that 47% of those with math disabilities in grade 5 met criteria for such disabilities (arithmetic scores below the 5th percentile) in grade 8. In a 6-year follow-up of this sample, Shalev, Manor, and Gross-Tsur (2005) found that 95% of those identified with math problems in grade 5 continued to perform in the bottom 25% of students in grade 11; 40% continued to meet the original definition of math disability (significantly below the 6th percentile). Data from the Early Childhood Longitudinal Study —Kindergarten showed that children entering and exiting kindergarten below the 10th percentile in math had a 70% chance of scoring below the 10th percentile 5 years later (Morgan, Farkas, & Wu, 2011). Thus, as in reading (Shaywitz et al., 1999), difficulties with math are persistent. Note that in Shalev et al. (2005), only 47% of students in grade 8 and 40% in grade 11 met the original definition, implying that students are moving around the thresholds. This phenomenon has also been observed in longitudinal studies of elementary school children (Gersten, Jordan, & Flojo, 2005; Mazzocco & Myers, 2003). As in reading, the small measurement error of the tests and the use of firm cut points on normally distributed dimensions likely account for some of this variability.

ACADEMIC SKILL DEFICITS

Mathematics Calculations versus Problem Solving

As the summary of definitional issues suggests, a major problem in defining mathematics LDs is the need to identify a set of key academic skill deficits that represent markers for LDs in mathematics. Ultimately, this identification would proceed from a model identifying critical mathematics competencies, much like reading can be broken into component skills involving word recognition and comprehension, or word- and text-level disorders.

Automaticity deficits, often referred to as reading or math fluency, are common in children with LDs in all domains (Chapter 10). We focus on computational versus problem-solving skills as broad domains in which LDs can occur, but this focus is not well aligned with mathematics curricula.

These curricula are organized into many strands, each of which is presumed to represent a separate component. For example, at the primary grades, the major curricular focus is whole numbers, which is conceptualized with three subdomains: understanding of numbers, calculations, and word problems. In the intermediate grades and middle school, the next major curricular focus is rational numbers. This includes common fractions, decimals, and proportions and has its own set of subdomains: part–whole understanding, measurement interpretation, calculations, and word problems. In high school, curriculum offerings include algebra, geometry, trigonometry, and calculus. The latter are not typically considered domains of LDs as exemplified by the ICD-10 definition of specific arithmetical difficulties.

Also in contrast to reading, little is understood about how curricular components relate to one other; which aspects of mathematics performance are shared or distinct; how difficulty in one corresponds to difficulty in another, both concurrently and from one grade to the next; and whether instruction in one or another area produces better learning in a third domain. Such understanding is important not only for an empirically guided framework for mathematics LD identification, but also for guiding the organization of curriculum and the design of instruction.

One approach for increasing understanding about the extent to which math calculations and problem solving represent separate components of mathematical cognition is to examine whether the cognitive abilities that support performance in these domains are shared. We discuss four large-scale investigations that examined the role of a large battery of cognitive abilities for predicting performance in calculations and word problems, looking for evidence that these two mathematical competencies are separable. Such evidence would support the validity of a hypothetical classification of math LDs that differentiated calculations and problem solving.

Swanson and Beebe-Frankenberger (2004) assessed 353 first through third graders on a set of cognitive predictors and mathematics outcomes concurrently and found that WM contributed to strong performance across domains involving calculation and word-problem skills. They also identified abilities that were uniquely associated with one or the other mathematics domain: phonological processing for calculations (also see section on language in this chapter) versus fluid intelligence and short-term memory for word problems. Swanson, Howard, and Saez (2006) then assessed these children 1 year later, and again identified different predictors of the two mathematics outcomes at the later time point: controlled attention, vocabulary knowledge, and visuospatial WM for calculations, but the executive control component of WM for word problems.

This finding of unique cognitive abilities for calculations versus word problems has been corroborated. With 312 third graders, L. S. Fuchs et al. (2006b) examined the cognitive correlates of calculations versus word problems, while controlling for the role of calculation skill in word problems. Teacher ratings of inattentive behavior correlated with both mathematics domains, but the remaining abilities differed: for calculations, phonological decoding and processing speed; for word problems, nonverbal problem solving, concept formation, and language. Similar findings, which indicate that development of calculation and word-problem skill depends on different constellations of numerical versus more general cognitive abilities, have been provided at first grade by L. S. Fuchs et al. (2005a, 2010a, 2010b).

L. S. Fuchs et al. (2008b) looked specifically at the predictors of difficulties in math calculations and problem solving. Third graders (n = 924) were first representatively sampled from 89 classrooms; assessed on calculations and word problems; classified as having difficulty with calculations or with word problems or with both domains or with neither domain; and measured on nine cognitive dimensions. Difficulties (represented by performance below the 25th percentile) occurred across both domains with the same prevalence as difficulty with a single domain, and the prevalence of specific difficulty in calculations versus word problems was similar. Figure 8.1 shows the cognitive profiles for the typically achieving group and groups with difficulties with calculation, problem solving, and both domains. Multivariate profile analysis on the cognitive dimensions in Figure 8.1 showed that specific calculation difficulty was associated with strengths in language processes and weaknesses in processing speed. Problem-solving difficulty, in contrast, was associated with deficient language processes and lower socioeconomic status (not depicted).

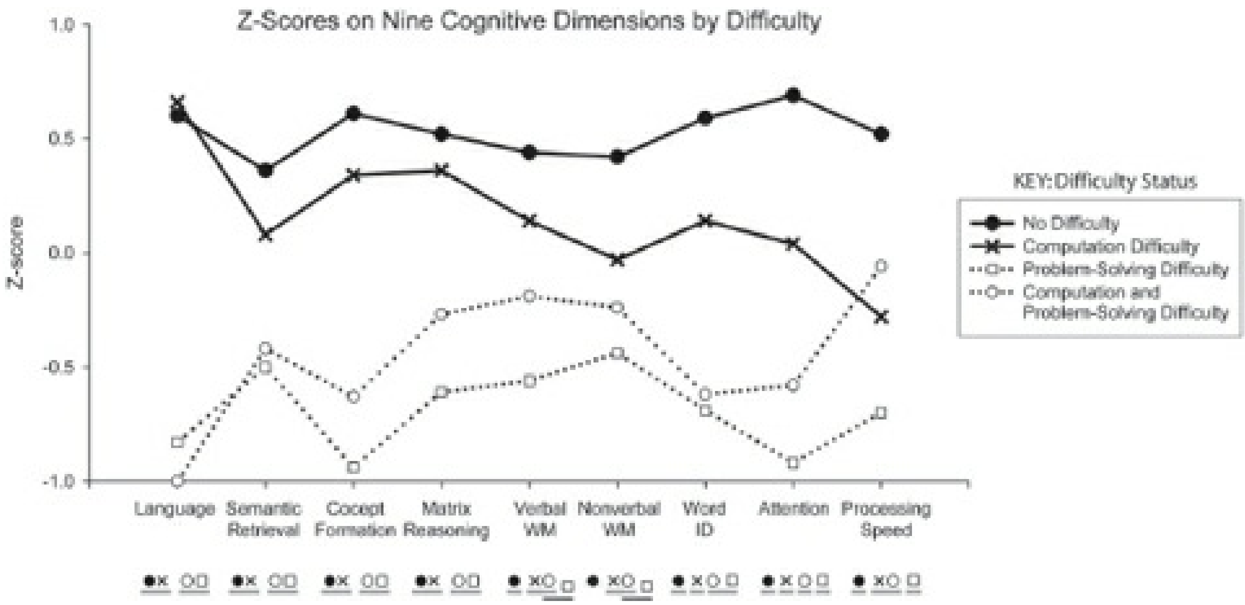

FIGURE 8.1. Cognitive profiles for groups with difficulties with calculation, problem solving, and both domains, and typically developing children. Specific computational difficulty (computational difficulty without problem-solving difficulty) was associated with deficient processing speed, whereas problem-solving difficulty, regardless of whether it occurred alone or in combination with computational difficulty, was associated with poverty and deficient language ability. From L. S. Fuchs et al. (2008b). Copyright © 2008 the American Psychological Association. Reprinted by permission.

Across studies, these results suggest that different combinations of cognitive abilities underlie the development of computational and problem-solving competencies. These differences suggest that these two domains of mathematics knowledge may be distinct; represent different types of LDs; and, as we discuss below, require different forms of intervention.

Number-Related Core Deficits in Mathematical Competencies

Other approaches to defining mathematical competencies have focused on two potential sources of numerical difficulty that may contribute to core competencies in mathematics: deficits in representations of magnitude (number sense) and difficulty in mapping symbols to magnitudes.

Representations of Magnitude

The perspective that core deficits in representations of magnitude explain LDs in mathematics rests on the assumption that an evolutionary need exists to understand magnitude and compare quantities (Butterworth, 2005). The capacity for magnitude comparisons can be observed in humans and nonhumans, and from infancy into adulthood. They have come to represent what Dehaene and Cohen (1997) characterized as "number sense." In studies of humans, infants at very early ages can discern differences in the numerosity of small sets (Starkey, Spelke, & Gelman, 1991). They can also discriminate quantity between larger arrays, first for larger ratios (Xu & Spelke, 2000) and then with greater precision for smaller ratios with increasing age. Infants also appear to infer the numerical value of manipulations on quantity, such as when adding to or subtracting from small sets (Wynn, 1998). Debate exists about whether this inherent sense of quantity depends on one or two representational systems, one for precisely representing small quantities and the other for approximately representing larger quantities (Geary, 2013). There is also controversy about whether and how symbolic representations are mapped from one or both of these systems (Lyons & Ansari, 2015).

Studies of magnitude processing in human infants were conducted to determine what basic cognitive mechanisms are present prior to significant environmental input and whether these mechanisms are specialized for number (Wynn, 1998; Xu & Spelke, 2000). More recently, researchers have been interested in understanding whether *individual differences in magnitude processing* might account for *individual differences in the acquisition of mathematical abilities,* leading to the hypothesis that a deficit in number sense measured by performance on magnitude-processing tasks might be a marker for LDs in mathematics. Research in this area has grown rapidly in the last few years, with 82% of such studies published between 2011 and 2014 and included in a 2015 meta-analysis on this topic (Schneider et al., 2017).

Three meta-analyses provide evidence for a small, consistent relation of nonsymbolic magnitude processing and performance on school-relevant math measures (Chen & Li, 2014; Schneider et al., 2017), with correlations in the range of .20–.24. Significantly larger, but still relatively modest correlations (.30), are found for magnitude comparison using symbols

(Schneider et al., 2017). These relations are consistent across age, but stronger for math outcomes measuring early number and operations skills and for mental arithmetic than for math problem solving and geometry. When math-related cognitive abilities (executive processes, language) and symbolic number processing are controlled, the relation of nonsymbolic magnitude processing to math achievement outcomes is diminished (Chen & Li, 2014; Fuhs & McNeil, 2013; Lyons, Price, Vaessen, Blomert, & Ansari, 2014).

No meta-analysis of magnitude processing exists for individuals with LDs in math. However, a review by De Smedt, Noël, Gilmore, and Ansari (2013) showed that deficits in symbolic magnitude comparison are consistently found across studies of children with LDs in mathematics, whereas the presence of deficits in nonsymbolic magnitude processing in such individuals is inconsistent. Mixed findings may be related to both child-based (e.g., age) and task-related variables. For example, in their study comparing typically developing children to those with LDs in mathematics, Bugden and Ansari (2016) found no group differences on congruent trials (where the number of dots and the area those dots take up are correlated). In contrast, differences between the groups were found on incongruent trials (where the number of dots is not correlated with area). Furthermore, visual–spatial WM was strongly related to this effect for incongruent trials. Similarly, Fuhs and McNeil (2013) reported that inhibitory control mediated the relation of performance on incongruent trials and math achievement in preschoolers. These findings suggest that task analysis of nonsymbolic magnitude-processing tasks are important for distinguishing the contributions of number-specific versus more domain-general abilities to math performance. One possibility is that problems in nonsymbolic magnitude-processing characterize a subtype of children with LD in mathematics. Although few studies have been conducted specifically on this topic, a latent profile analysis provided no evidence for a subgroup of children with low nonsymbolic magnitude processing and low math achievement (Hart et al., 2016).

Other evidence that can be helpful for determining the role of a variable such as magnitude processing on school-based mathematical abilities comes from longitudinal and experimental studies. Here, the evidence is also inconsistent. One longitudinal study showed a significant, modest correlation of magnitude processing at 6 months of age with early math knowledge at 3–4 years (Starr, Libertus, & Brannon, 2013), while another study of children from 3 to 4 years of age found that early symbolic number understanding predicted nonsymbolic magnitude processing 7 months later, rather than the other way around (Mussolin, Nys, Content, & Leybaert, 2014). Schooling significantly improves the acuity of the approximate number system, suggesting reciprocal developmental relations between magnitude-processing and formal mathematical abilities (Piazza, Pica, Izard, Spelke, & Dehaene, 2013). In a 1-year longitudinal study of 74 children moving from third to fourth grade, Vanbinst, Ansari, Ghesquière, and De Smedt (2016) found correlations of symbolic numerical processing and math calculations of about .4, which paralleled correlations of phonological awareness and timed word reading in the same sample. Phonological awareness also shows reciprocal relations with word reading during schooling.

Few experimental studies have looked at the effect of training in nonsymbolic magnitude processing on mathematical skills, with most conducted with adults (Park & Brannon, 2013) and typically developing children (e.g., Hyde, Khanum, & Spelke, 2014). However, Brankaer, Ghesquière, and De Smedt (2014) found an effect of training in nonsymbolic magnitude processing on symbolic magnitude processing in children with mild intellectual disability. In these studies, training on nonsymbolic tasks transfers to symbolic arithmetic when the nonsymbolic tasks involve *operating* on or manipulating large nonsymbolic quantities (e.g., estimating approximate nonsymbolic sums and differences when adding and subtracting from large dot displays). These findings raise the question of what aspect(s) of nonsymbolic magnitude processing account for transfer to symbolic arithmetic (Lyons & Ansari, 2015). Furthermore, because interventions designed to improve performance on nonsymbolic tasks have not been compared to other interventions that also improve symbolic arithmetic performance, comparative effectiveness cannot be determined.

Although there has been great interest in the hypothesis that individual differences in a phylogenetically ancient and ontogenetically early number sense capacity underlie LDs in mathematics, data from a variety of sources fails to provide strong evidence for this position. There is little support for the diagnostic use of nonsymbolic magnitude comparison tasks for assessments of children with or at risk for LDs in mathematics or for specific training in nonsymbolic magnitude comparison per se for these children at the present time.

Mapping Symbols to Magnitudes

The second, contrasting perspective within the domain of number-related sources of LDs in mathematics is that children with mathematics LDs have intact approximate number systems (i.e., ability to compare large nonsymbolic quantities), but are slower at mapping those quantities to the number words and Arabic numerals used in the formal mathematics system (Rousselle & Noël, 2007). As we saw from the meta-analyses described above, symbolic magnitude processing is more strongly related to math achievement than is nonsymbolic magnitude processing, and this is true across the continuum of math ability (Schneider et al., 2017) and for children with LDs in mathematics (De Smedt et al., 2013).

However, whether these findings reflect a mapping problem per se is not known. In a large cross-sectional study from 1st to 6th grades, symbolic magnitude comparison was a strong predictor of achievement in mental arithmetic in grades 1 and 2. In contrast, ordering of symbolic quantities was more predictive of mental arithmetic in the later elementary grades (Lyons et al., 2014).

Like the findings discussed above for nonsymbolic magnitude processing, it seems important for the field to better understand what these various nonsymbolic and symbolic number tasks actually assess and how they map onto formal mathematical learning and performance.

A Broader View

These perspectives about mathematical competencies reflect different theoretical positions, with direct implications for how research progresses in mathematics LDs. When mathematics LDs are viewed as core deficits in an inherent sense of number (the first perspective), the key challenge in defining and understanding mathematics LDs is to identify the competencies that make up this basic capacity of the human and nonhuman brain, which is not a straightforward and well-supported account of mathematics difficulties.

An adequate explanation of long-term mathematical development, especially competencies that involve more complex forms of mathematical reasoning—including word problems—appears to require a broader view. This includes the language system, regardless of whether language is considered to simply *facilitate* the development of mathematical skills (Gelman & Butterworth, 2005) or thought to be *causally* implicated in the development of core mathematical skills and concepts (Carey, 2004).

This broader view also encompasses a third perspective, in which mathematics comprises subdomains of knowledge, built on other general cognitive or neuropsychological systems such as the language system, the visual–spatial system, and the central executive that sustains attention and inhibits irrelevant information (Ashkenazi et al., 2013; Geary, 2013). Difficulties in mathematics may arise from these and related cognitive systems or their interactions and may lead to patterns of deficits in different mathematics competencies as well as begin to explain the overlap of math and reading LDs. In this framework, difficulty in mapping nonsymbolic to symbolic number may reflect a general learning deficit that applies across academic domains. Geary (2013) focused on poor attentional control, in combination with general cognitive abilities such as WM, inhibitory control, and language. Together, these processes mediate or account for the relation of nonsymbolic magnitude processing to math achievement.

There is some evidence for this broader framework (Ashkenazi et al., 2013; Geary, 2004, 2005), but the research base is not adequate to determine how the value of this approach compares to the more modular approach. It is apparent that children with mathematics LDs vary in component mathematics skills and in the cognitive processes related to these skills (L. S. Fuchs et al., 2005a, 2006b; Hanich, Jordan, Kaplan, & Dick, 2001). However, regardless of the model of mathematics LDs, *understanding the core processes underlying LDs in mathematics begs the question of the mathematics skill deficits that identify LDs.* Much of the research base has focused on children identified in terms of calculation difficulties. This focus is not surprising, as the early neurological literature often described adults and children with "dyscalculia," based on their inability to perform simple arithmetic calculations orally or on paper.

One problem with an exclusive or at least dominant focus on dyscalculia or calculations is that mathematics competence is multifaceted, just as proficient reading involves accurate word recognition, fluent word and text reading, and comprehension—each possibly determined by multiple core cognitive processes (L. S. Fuchs et al., 2005a, 2006b). Mathematics involves calculations, itself the product of number knowledge; retrieval of facts; and the application of procedural knowledge. Word problem solving involves calculation, language, and reasoning (and, in school, reading skill) (Geary, 1993). Any successful execution of mathematics competencies requires that the person is attentive, organized, able to switch sets, and work quickly enough to avoid overloading WM stores that retain information needed for online access of different kinds of information (Geary, 2013). Many of the tests used to measure mathematics achievement tend to confound multiple components of mathematics. For example, test items sample from potentially quite different domains of mathematics and focus primarily on calculations or word problems.

The value of the focus on basic numerical competencies is that it facilitates analysis into discrete mathematics components. If the investigator carefully specifies the component of mathematics to be assessed and then looks at the cognitive correlates of that component in a sample that varies in such competencies, the links between breakdowns in mathematics competencies and cognition may become more apparent. *Researchers and practitioners must carefully specify the component of numerical ability being evaluated and treated because the number-specific and cognitive correlates likely vary.*

CORE COGNITIVE PROCESSES

Given the difficulty in defining a set of core academic skill deficits that identify individuals with LDs in mathematics, it is not surprising that research has not advanced to a level that allows the specification of a set of core cognitive processes that underlie LDs in mathematics. Much depends not only on the type of theoretical orientation but also on the specific academic competencies used to identify the LDs. In this section, we focus on the general cognitive processes associated with different types of mathematical competencies. In keeping with the overall theme of this book, we also consider the general cognitive processes associated with math LDs when they are accompanied by other types of LDs.

Cognitive Processes and Academic Deficits

Comorbidity with Reading Disabilities

One hypothesis in the field of mathematics LD is that comorbid mathematics and reading difficulty is a subtype of LD associated with distinctive profiles of mathematics and cognitive deficits and with disproportionately poor response to mathematics intervention (Ashkenazi et al., 2013; Geary, 1993; Rourke & Finlayson, 1978). Concurrent mathematics and reading difficulty is often referred to as an example of *comorbid* LDs. Cirino et al. (2015) used lower-order skills (calculation and word reading) to identify students with mathematics difficulty (n = 105), reading difficulty (n = 65), both (n = 87), or neither (n = 403) and tested them on mathematical competencies not used to define the groups (word problems, mathematics concepts, arithmetic, procedural calculations) as well as several cognitive abilities. Results indicated that the strongest mathematical contributors to the distinction between groups with and without comorbid difficulties were word problems and math concepts, favoring the noncomorbid groups. In terms of cognitive profiles, articulation speed, phonological awareness, and verbal comprehension differentiated the comorbid and noncomorbid groups. These findings are consistent with some studies (Jordan, Hanich, & Kaplan, 2003; Willcutt et al., 2013). However, there is variability across studies, with some reporting no difference on WM (e.g., Andersson, 2010; Vukovic, 2012), but not others (Cirino et al., 2015; Willcutt et al., 2013). Set shifting has emerged as a unique marker of math difficulties regardless of the presence of a WLRD (Willcutt et al., 2013).

L. S. Fuchs, Geary, Fuchs, Compton, and Hamlett (2016a) also focused on the predictors of lower-order reading and math skill,but used path analysis to look at both indirect and direct effects of these predictors in a longitudinal framework. Children (n = 747) were assessed on cognitive and linguistic processes and on basic numerical and reading-related competencies at start of first grade; on addition retrieval at end of second grade; and on calculations and word reading at end of third grade. Path analysis revealed attentive behavior, reasoning, visuospatial memory, and rapid automatized naming (RAN) indirectly contributed to both calculation and word-reading outcomes, via retrieval. However, there was no overlap in domain-general direct effects on calculation (attentive behavior, reasoning, WM) versus word-reading (language, phonological memory, RAN) outcomes. The pattern suggests the ease of forming associative relations and the abilities engaged during the formation of these long-term memories are common to calculations and word reading and can be indexed by addition fact retrieval, but that further growth in calculation and word-reading competence is driven by a different constellation of domain-general abilities.

The dominant approach, as reflected in these studies, has been to define LDs on lower-order skills: calculations and word-level reading. L. S. Fuchs, D. Fuchs, and Compton (2013a) examined how lower- versus higher-order skill definitions affect profiles, again with third-grade children. With LDs defined on lower-order skill (calculations/word-level reading), language comprehension distinguished the two groups because it was a relative strength for noncomorbid students. In contrast, with LDs defined on higher-order skill (word problems/reading comprehension), language comprehension was a distinctive dimension because it was a relative weakness for comorbid students. Thus, language comprehension and comorbidity may have more meaning when LDs are defined on higher-order skills. This suggests the need to study comorbidity also using higher-order skill definitions of LDs, especially in relation to specific reading comprehension difficulties because there should be overlap given the impairment in listening comprehension (Chapter 7).

L. S. Fuchs, D. Fuchs, Compton, Hamlett, and Wang (2015a) focused on higher-order reading comprehension and mathematics problem solving. The hypotheses were that (1) word problem solving is a form of text comprehension that involves language comprehension processes, WM, and reasoning, but (2) word problem solving differs from other forms of text comprehension by requiring word-problem-specific language comprehension as well as general language comprehension. At the start of second grade, children (n = 206) were assessed on general language comprehension, WM, nonlinguistic reasoning, processing speed (a control variable), and foundational skill (arithmetic for word problems; word reading for text comprehension). In spring, they were assessed on word-problem-specific language comprehension, word problems, and text comprehension. Path analytic mediation analysis indicated that effects of general language comprehension on text comprehension were entirely direct, whereas effects of general language comprehension on word problems were partially mediated by word-problem-specific language. Otherwise, the effects of WM and reasoning operated in parallel ways for both outcomes. Not surprisingly, there was overlap in the pathways toward word-problem and reading comprehension competence.

To summarize, these studies suggest that comorbidity with reading problems may have more salience when considering higher-order performance in math problem solving and reading comprehension. Additional research is required to include randomized control trials with adequate statistical power so that interactions between comorbid status and intervention condition can be detected. In such research, the intervention should include a component that speaks to the hypothesized cognitive and linguistic differences between students with versus without comorbidity. There is also overlap of mathematics LDs with ADHD and oral language disorders, which is discussed below.

Calculation versus Word-Problem Difficulty

Studies have also examined whether the core cognitive processes associated with math LDs differ as a function of whether LDs are defined in terms of calculation versus word-problem difficulty. To gain insight into this issue, representative samples of children are assessed on a large battery of cognitive measures and then assessed on a later, separate occasion on calculation and word-problem measures. The value of the cognitive measures is evaluated in predicting the mathematics outcomes.

As explained in the section on component mathematics skills, evidence indicates key differences in these two aspects of math. For example, Swanson et al. (2006) found that controlled attention, vocabulary knowledge, and visuospatial WM were unique predictors of calculations, whereas the executive control component of WM was uniquely predictive for word problems. L. S. Fuchs et al. (2010a) found that facility with representing small and large quantities and numerical WM predicted the development of calculation and word-problem skill, but the set of domain-general cognitive processes associated with the two mathematics outcomes differed. For calculations, only WM (in the forms of counting span) was uniquely predictive. In contrast, regarding word problems, language comprehension, attentive behavior, reasoning ability, and WM (in the form of sentence span) were uniquely predictive. This again suggests that the core cognitive processes associated with mathematics LDs appear to differ, depending on the mathematics skills used to define LDs.

A meta-analysis of one particular cognitive process in relation to math, WM, found the largest relations of WM for calculations and word problems (.35 and .37, respectively) compared to other types of mathematics such as geometry (Peng, Namkung, Barnes, & Sun, 2015). Furthermore, the relation of WM and math was greater for children with disabilities in math in addition to other behavioral and neurocognitive disorders (e.g., ADHD, fetal alcohol syndrome) than it was for both typically developing children and those with LDs only in math (.52 vs. .34, and .25).

Other Components of the Mathematics Curriculum

A few studies have addressed components of the mathematics curriculum beyond whole-number calculations and word problems. One set of studies focuses on fractions, where it has been suggested that learning suffers at least in part because of what Ni and Zhou (2005) referred to as whole-number bias. That is, children apply their understanding of whole numbers to the solving of fraction problems, resulting in conceptual and procedural errors. For instance, whereas whole numbers have a unique predecessor and successor, such as 5 following 4 in the counting sequence, fractions do not have these unique features. Overgeneralization of this property of whole numbers contributes to children's difficulty understanding that ¼ is larger than ⅕. Also, fractions and whole numbers differ in how operations are executed and interpreted. For example, multiplying whole numbers always results in a larger number, but multiplying fractions often results in a smaller one. Given that instruction in the early school years favors whole-number skills, children must make conceptual leaps from whole-number understanding to acquire conventional understanding of fractions commensurate to their informal understanding.

The developmental sources of children's difficulty with fractions are not well understood, but several investigations lend insight. Seethaler, Fuchs, Star, and Bryant (2011) found that third-grade calculation skill as well as domain-general competencies of language, nonverbal reasoning, and the central executive component of WM uniquely predicted fifth-grade performance on rational number calculations. Hecht and Vagi (2010) found that attentive behavior and fraction concepts in fourth grade predicted procedural fraction outcomes in fifth grade. Controlling for prior mathematics achievement, Jordan et al. (2013) found that third graders' ability to order numbers on a number line predicted both their understanding of fraction concepts and procedural facility with fractions in fourth grade, whereas central executive capacity made unique contributions to fraction procedures, and language and attentive behavior made unique contributions to fraction concepts.

Two additional studies frame their investigations to generate an integrated account of factors that contribute to fraction knowledge and, by association, to the understanding of LDs involving fractions. Hecht, Close, and Santisi (2003) proposed a model whereby domain-general cognitive processes predict fraction knowledge directly as well as indirectly through numeracy skills. They tested their model for fifth graders using the domain-general predictors of central executive capacity and attentive classroom behavior and the domain-specific mediators of fraction concepts and simple whole-number calculations. Attentive behavior and central executive capacity were both related to simple whole-number calculations, but only attentive behavior was related to fraction concepts. Also, skill with fraction concepts was significantly related to fraction calculations, word-problems, and estimation skill. By contrast, simple whole-number calculation skill was related only to fraction calculations.

In a second study that attempted to provide an integrated account of fraction learning, Vukovic et al. (2014) used a longitudinal design to investigate the effects of domain-general and domain-specific competencies on individual differences in children's understanding of fractions. Students were assessed at 6 years of age on reasoning, language, attentive behavior, central executive, and visual–spatial sketchpad and on number knowledge competencies. At 7 years, they were tested on whole-number concepts and procedures that reflect formal school learning (number line estimation and arithmetic calculations), and at 10 years, they were assessed on fraction concepts. Relations between domain-general competencies and children's later understanding of fractions concepts were all indirect—mediated by whole-number skills. The effect of language ability was through number line estimation and arithmetic calculations; the effect of the visual–spatial sketchpad, through number line estimation; and the effect of attentive behavior, via arithmetic calculations.

Taken together, these findings suggest that domain-general competencies at school entry provide building blocks for later learning, but are not directly involved in how children understand fraction concepts 4 years later. The findings indicate that domain-general competencies are related to understanding early fraction concepts through their influence on the *learning* of precursor whole-number skills.

The second main finding was that all of the whole-number variables—first-grade number knowledge, second-grade number line estimation, and second-grade arithmetic calculations—were uniquely related to fraction concepts. First-grade number knowledge had a direct effect on fraction concepts. Second-grade number line estimation and whole-number operations were also significantly related to fourth-grade fraction concepts. These findings suggested that early whole-number knowledge is foundational for the acquisition of fraction concepts, as predicted by the National Mathematics Advisory Panel (Geary et al., 2008a). It is also noteworthy that number knowledge was the only first-grade variable (not the first-grade domain-general cognitive variables) to have a *direct* effect on later fraction concepts.

Even fewer studies are available to inform understanding of how cognitive processes support students' development of *algebraic knowledge* and how they explain mathematics LDs. The Peng et al. (2015) meta-analysis reported a significant, but small, relation of .27 for the relation of WM to algebra. Working with 10-year-olds, Lee, Ng, Ng, and Lim (2004) found that central executive processes, Performance IQ, and literacy predicted concurrent prealgebraic word-problem skill, although relations were small. Lee, Ng, Bull, Pe, and Ho (2011) found that pattern recognition and calculations completely mediated the effects of updating (a form of WM). Tolar, Lederberg, and Fletcher (2009) showed that fluency with calculations had stronger effects on algebra than WM or spatial ability, which nevertheless did have moderate effects. L. S. Fuchs et al. (2016c) assessed 279 children on seven domain-general cognitive resources as well as arithmetic calculations and word problems at start of 2nd grade and on calculations, word problems, and prealgebraic knowledge at end of 3rd grade. Arithmetic calculations and word problems were foundational to prealgebraic knowledge, but results also revealed direct contributions of reasoning and oral language to prealgebraic knowledge, beyond indirect effects that were mediated via whole-number calculations and word problems. In contrast, attentive behavior, phonological processing, and processing speed contributed to prealgebraic knowledge only indirectly via whole-number calculations and word problems. WM was not a significant predictor.

Summary: Cognitive Processes and Academic Deficits

As studies on fractions and algebraic thinking illustrate, understanding the development of component mathematics skills that first become central components of the mathematics curriculum in the intermediate grades and secondary levels is more complex than studying curricular skills that are relevant at the start of school. The effects of cognitive processes on later mathematics learning often occur indirectly, via their effects on learning the simpler primary-grade math skills. The Vukovic et al. (2014) study suggested this may be entirely the case for fractions; the L. S. Fuchs et al. (2012) study suggested that some cognitive processes, especially higher-order language and reasoning, exert direct and indirect effects, whereas the effects of lower-order cognitive processes such as processing speed are entirely mediated through early targets in the mathematics curriculum.

Examining the cognitive processes involved in comorbid forms of mathematics LDs and in the development of whole-number calculation and word-problem skills, fractions, and algebra indicates that WM, language, and a handful of other cognitive processes are differentially correlated with one or more components of mathematics skill. In this next section, we consider how the cognitive processes most commonly related to math (WM/executive processes and language) may exert an effect in the development of mathematics competencies.

Cognitive Processes and Mathematics Learning

In the next section, we move from associations of different ways of defining mathematics skill components and cognition to more general cognitive processes implicated in broadly defined mathematics learning in calculation and problem-solving abilities. These processes include WM, attention, and language.

Working Memory

WM and mathematics skills are correlated (average correlation = .35; Peng et al., 2018), particularly for calculations and word-problem solving. Whether this role is supportive or causal is debatable (see Menon, 2016). Most of the research on children with and without LDs in math is correlational in nature rather than experimental. To our knowledge, few studies assess requirements for WM resources during the learning of new concepts and procedures and subsequent performance in these areas of math. For example, dual task procedures that were used in the cognitive literature to establish a causal relation between WM and math performance in adults (LeFevre, DeStefano, Coleman, & Shanahan, 2005) have not been widely applied to investigate the nature of the role of WM in children's math learning and performance. Experimental studies that measure the effects of WM interventions on math are in their infancy, although early findings for children *without* LDs in math have not been robust when WM is trained as an isolated skill (Jacob & Parkinson, 2015; Melby-Lervåg et al., 2016). Longitudinal studies, although correlational, involve time precedence that can inform us about the direction of the relation of two skills over development. In a study of children with spina bifida, who have high rates of math LD, Barnes et al. (2014) found that WM abilities at 36 months of age partially mediated achievement in math calculations and math problem solving, but not math fact fluency, 5 to 6 years later. These findings suggest that early WM abilities play a role in the learning and/or performance on more complex math tasks including multidigit calculation and problem solving.

Despite the lack of causal studies, it is likely that WM is implicated in both mathematical learning and performance given what is known about the role of WM for engaging in cognitive tasks that are evolutionarily novel (Geary, 2013; Menon, 2016) and that involve higher-order reasoning processes. Menon (2016) reviewed studies showing that visuospatial WM was particularly weak in children with math LDs. In addition (see below), neuroimaging studies of math identify a neural network that is involved in mathematical processing, parts of which are activated when cognitive resources such as WM and attention are needed for task performance (Ashkenazi et al., 2013).

What are the mechanisms by which WM may affect mathematical learning and/or performance on math tasks? In very young children, WM is necessary for forming mental models that help the child to solve concrete mathematical problems, such as manipulating nonsymbolic quantities (Rasmussen & Bisanz, 2005), and also for early counting (Noël, 2009). As verbal memory develops, children begin to rely more on verbal codes that reflect automaticity to accomplish mathematical tasks, but may also rely on visual–spatial codes during math learning and performance. Geary (1993) proposed that WM is used to learn math facts during the transition from counting strategies to direct retrieval such as being able to hold two addends together in memory at the same time when computing the sum. In multidigit calculation, as well as mental arithmetic, WM may be needed to hold intermediate sums, products, and so forth. During word-problem solving, WM may be required to integrate ideas in the linguistic input as well as map and coordinate the linguistic and mathematical components of the task. The type of code used to accomplish various math tasks may depend on (1) the type of strategy the child brings to bear to solve the math problem (Menon, 2016); (2) the type of math task (LeFevre et al., 2010); and (3) the age of the child and and his or her stage of skill learning. For example, some researchers have suggested that visual–spatial resources may come into play when first learning new complex math skills, regardless of age (Geary, 2011; Raghubar, Barnes, & Hecht, 2010).

These studies show that the relations of WM and math LDs are complicated, and additional research is necessary to determine how WM might affect mathematical learning in children with math disabilities, with and without WLRDs (Geary, Hoard, Byrd-Craven, & DeSoto, 2004; Menon, 2016). Children who have both reading and math LDs tend to have more severe problems in WM than those who have problems in only reading or math (Figure 2.2). Similarly, relations between WM and math LDs are likely to be complex because different aspects of WM may be related to different mathematical skills.

Attention

The association of attention difficulties and mathematics achievement is significant and large even among children with LDs who do not meet diagnostic criteria for ADHD. Several studies have highlighted the role of behavioral indices of attention and concentration as a robust predictor of math skills. For example, L. S. Fuchs et al. (2005a) found that teacher ratings of attention predicted arithmetic skills even when controlling for several other cognitive abilities. Fuchs et al. also found that teacher ratings of attention (i.e., distractibility) were a unique predictor of calculation, paralleling findings reported in studies of children identified with ADHD (Ackerman, Anhalt, & Dykman, 1986). Severity of the LD in math is related to ratings of inattention, which predicts difficulties in both conceptual and procedural aspects of math (Cirino, Fletcher, Ewing-Cobbs, Barnes, & Fuchs, 2007; Raghubar et al., 2009; Willcutt et al., 2010a). The question is exactly what teachers evaluate when they complete such scales; it may be that teachers are simply rating children according to academic competencies.

There is evidence that relations of attention to math extend beyond teacher ratings. Mathematics and inattention are phenotypically and genetically correlated and these associations are not fully accounted for by other academic or cognitive variables (Greven, Kovas, Willcutt, Petrill, & Plomin, 2014). Attention switching measured at the child level was the only unique predictor of difficulties in math in a multivariate study of the neuropsychological correlates of LDs in math with and without comorbid

LDs in reading (Willcutt et al., 2013).

What is the mechanism by which attention is thought to be important for mathematical competence? In an analysis of individual items from teacher ratings to predict outcomes in math interventions, Cirino et al. (2007) found that an item involving attention to details was most predictive. Hyperactivity-impulsivity items are generally not related to math competence (L. S. Fuchs et al., 2005a), so disinhibitory control is not likely a key element. More generally, attention is a “hub” cognitive domain because it performs a gatekeeping function for the acquisition of skills across many cognitive and academic areas, including the acquisition of executive skills such as WM and the set of skills subsumed under self-regulation. The ability to maintain vigilance or sustain attention across tasks that have multiple components and that unfold across time would seem to apply both to new learning of math concepts and procedures and to performance on many different math tasks.

Another mechanism involves executive control of attention, which is the ability to flexibly switch between rules during task performance by focusing on task-relevant information and inhibiting task-irrelevant information. This is very important for attending to details, such as mathematics signs and operands, algorithmic steps, and specific clues in math word problems. How vigilance and executive attention operate during different types of math learning and performance has not been studied experimentally, though training in vigilance and executive attention have not been found to improve math learning in preschool children over and above the effects of a math-specific intervention (Barnes et al., 2016a).

Language

Carey (2004) proposed that language is important in enabling formal math learning, as well as for development in areas of math like geometry that are often considered to be the least verbal (Spelke & Tsivkin, 2001). Language provides a set of symbols, such as the counting words that have no inherent meaning but which set the stage for the mappings between previously distinct representational systems, such as quantitative and language systems. The resulting integrated representations are more powerful and result in new mental structures. There is also longitudinal evidence that language/verbal abilities are drivers of developmental change in other math-related abilities such as executive processes and WM (e.g., Fuhs & Day, 2011).

Some evidence for the importance of language in the development of early math skills comes from studies of toddlers and preschoolers in which even the development of calculation using small numbers varies according to linguistic quantifiers that differ across languages (Hodent, Bryant, & Houde, 2005). In the Pathways to Mathematics model (LeFevre et al., 2010), vocabulary and phonological awareness at 4 years of age accounted for unique significant variance on all math achievement outcomes, including number line tasks and geometry, at 7 years of age. Although the role of language in mathematical competence has been explicated for typical development, aberrant development of the language system could be expected to result in deficits in certain aspects of mathematical function even from a very young age.

From a different perspective, other studies suggest that children who are impaired in both reading and math typically show more severe and pervasive disturbances of oral language than children who are impaired only in word recognition (e.g., Peterson et al., 2017). This pattern seems to hold for all comorbid associations of LDs and ADHD. The difficulties of children with comorbid LDs involving mathematics reflect problems in learning, retaining, and retrieving math facts, which are essential to precise calculation. These lead to pervasive difficulties with math. Thus, Jordan and Hanich (2000) found that children with both reading and math difficulties showed problems in multiple domains of mathematical thinking.

Language impairments clearly lead to difficulties in the acquisition of math skills, although the precise mechanisms of action are not well explored. For example, measures of the quality of lexical representations such as phonological awareness are better predictors of counting knowledge and counting procedures than other aspects of language such as oral language vocabulary knowledge (Barnes et al., 2011), but little is known about math-specific vocabulary. It is likely that language is both *directly* related to many aspects of math and also *indirectly* related to math through its developmental influences on other math-related correlates such as executive skills.

Summary: Cognitive Processes and Mathematical

Learning

In reading, there is agreement about the important domains of competency and consistency in the cognitive correlates. Controversies in children with dyslexia, for example, are often pitted as alternatives to the dominant phonological-processing hypothesis, usually as alternative explanations in a single deficit model or as a needed expansion because of a multiple deficit or subtype model. In contrast, there is no consensus about core mathematical competencies, and the associations with cognitive skills vary depending on the identified competency. We have focused on calculation and problem solving as two broad domains of mathematics LDs, partly because much of the research can be organized under these two broad rubrics, but also because (1) this subdivision has been formally tested from a classification framework and (2) aligns with instructional approaches. As we turn to neurobiological processes, we will see that most of the research focuses on children and adults defined on the basis of a computational deficit.

NEUROBIOLOGICAL FACTORS

When the first edition of this book was published (Fletcher et al., 2007), there were very few studies of either brain structure or brain function in children with and without LDs in math. The past few years have seen an increase in such studies. Also rapidly emerging are studies of the familial segregation and heritability of math disability.

Math and the Brain

Figure 8.2 shows areas of the brain implicated in mathematics, based largely on studies of arithmetic (Ashkenazi et al., 2013; Peters & De Smedt, 2017). The involvement of these circuits varies depending on the task, but does represent a network involving posterior parietal areas strongly implicated in number processing; occipitotemporal regions involved in the representation of numbers; prefrontal regions involved in cognitive control and automaticity; and the hippocampus, essential for retrieval of number facts.

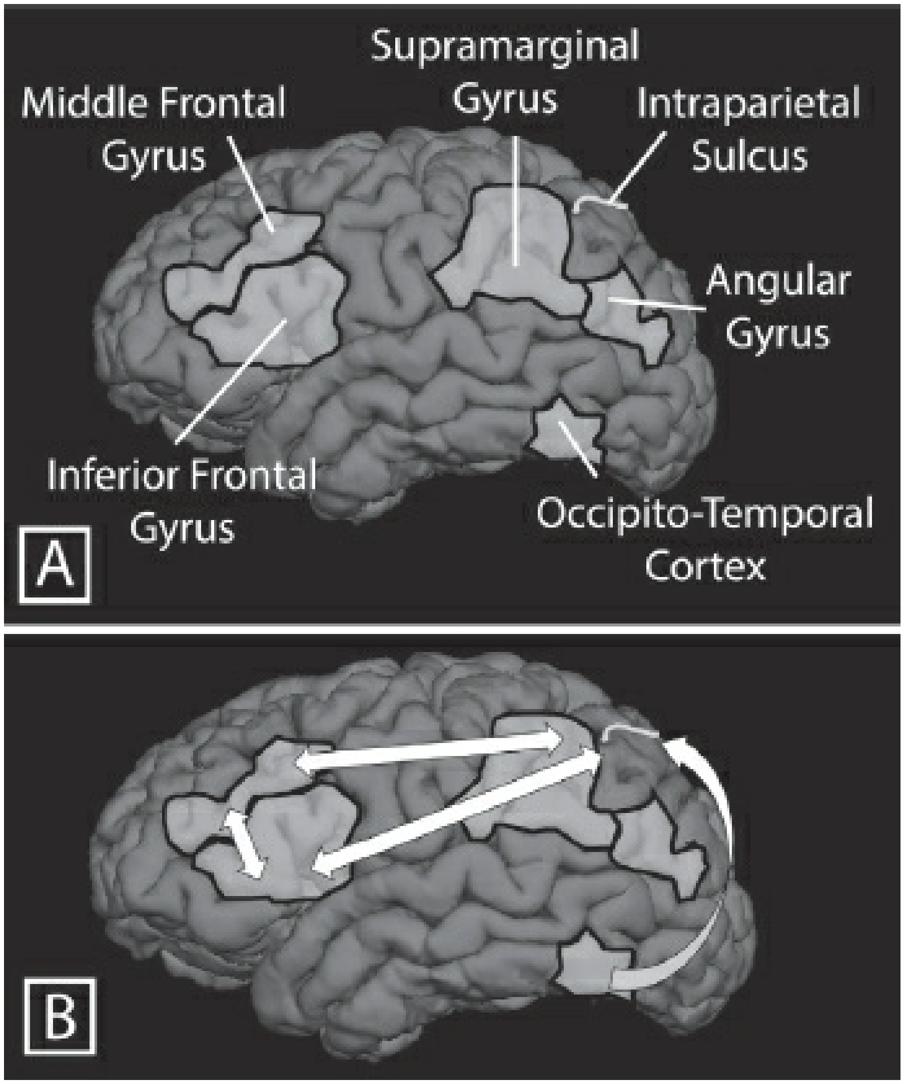

FIGURE 8.2. Neural networks involved in arithmetical skills. The posterior parietal circuits include the posterior superior parietal lobe and the intraparietal sulcus, implicated in numerical processing; the supramarginal, angular, and fusiform gyri; frontal areas including the dorsolateral and ventrolateral cortices. Not depicted on this sagittal slice is the hippocampus, important for fact retrieval. Courtesy Victoria Williams.

Functional neuroimaging studies of math in children as well as structural studies of the brain, including those on the connectivity of different brain regions, have largely addressed the neural correlates of simple and more complex arithmetic as well as magnitude processing. Investigations of the neural correlates of other aspects of math reviewed in this chapter have not been completed. Furthermore, imaging studies of typically developing children outnumber those involving children with LDs in math.

Brain Function:Typical Development

In functional imaging studies, the regions of brain that are associated with mathematical function differ somewhat depending on the type of math task that is being performed. Meta-analyses of the neural correlates of mathematics processing reveal that parietal regions, namely, the inferior and superior parietal lobules, are associated with magnitude comparison (Kaufmann et al., 2011). Earlier in this chapter we discussed uncertainty about the relation of nonsymbolic and symbolic number processing. This meta-analysis also showed that neural activation on nonsymbolic and symbolic magnitude comparison tasks in children involves largely nonoverlapping areas both within and extending beyond the parietal lobes (Kaufmann et al., 2011).

In contrast to magnitude processing, calculation was associated with a more distributed frontoparietal network (Kaufmann et al., 2011) and represents a parietal–frontal shift as magnitude-processing skills become automatized (Matejko & Ansari, 2015). However, simple arithmetic seems to be more associated with the posterior parietal than with the frontal regions, with more complex computations engaging the frontal regions. The greater involvement of frontal regions in more complex math tasks may be related to the recruitment of domain-general cognitive processes such as attention, WM, and the like needed for aspects of math that are not automatized (Menon, 2016). A shift with age/experience/training on tasks such as calculation, from more frontal to more parietal involvement (Kaufmann et al., 2011; Rivera, Reiss, Eckert, & Menon, 2005) is consistent with dual process theories of cognition: more novel tasks requiring higher-order reasoning processes require greater WM resources compared to tasks that can be performed using more automatized processes (Geary, 2013).

In another meta-analysis, Arsalidou and Taylor (2011) found strong evidence aligning simple number tasks with the inferior and superior parietal lobes. Different kinds of calculation tasks activated these regions as well as the middle and superior frontal gyri. There were also differences in activation for different calculation tasks, especially in the distribution across the frontal regions and cingulate gyrus. More complex tasks require more cognitive resources, especially those involved in WM and other aspects of cognitive control mediated by anterior brain regions. Interestingly, activation of the occipitotemporal regions was also consistent, showing a role for the ventral system in the processing of visual representations of numbers. There was also more activation of the left fusiform gyrus than the right, again implicating this region for highly specialized orthographic feature extraction. However, there has also been inconsistent evidence for a “number form area” in the left *inferior* temporal gyrus (Yeo, Wilkey, & Price, 2017).

Evans, Flowers, Luetje, Napoliello, and Eden (2016) found overlap in areas related to addition, subtraction, and word reading in 7–29 year-old-individuals with no evidence of LDs. Subtraction activated bilateral intraparietal sulci and supramarginal gyri, right insula, inferior frontal gyrus, and cingulate, reflecting its greater complexity. Simple addition and word reading shared activity in the right middle temporal gyrus and left superior temporal gyrus. Word reading was associated with the left fusiform gyrus. Age and experience were associated with reduced left frontal activity and greater right temporoparietal shift in activity on all tasks. The issue of overlap in regions associated with math, reading, and attention would help illuminate the comorbidity issue.

Brain Structure: Typical Development

Some studies have found relations of brain volumes of cortical and subcortical areas and different math outcomes. Evans et al. (2015) found relations of gray matter volume at age 8 in the posterior parietal cortex (see Figure 8.2) and the prefrontal cortex. These volumes predicted the development of mathematics abilities, but not reading abilities. The strength of functional coupling among these regions also predicted gains in numerical abilities, supporting the idea of a network of brain regions that works in concert to promote numerical skill acquisition. In other studies, Supekar et al. (2013) found strong connectivity of frontal and parietal regions, and the basal ganglia, with the hippocampus (Plate 7). These connections increased significantly in an 8-week tutoring program.

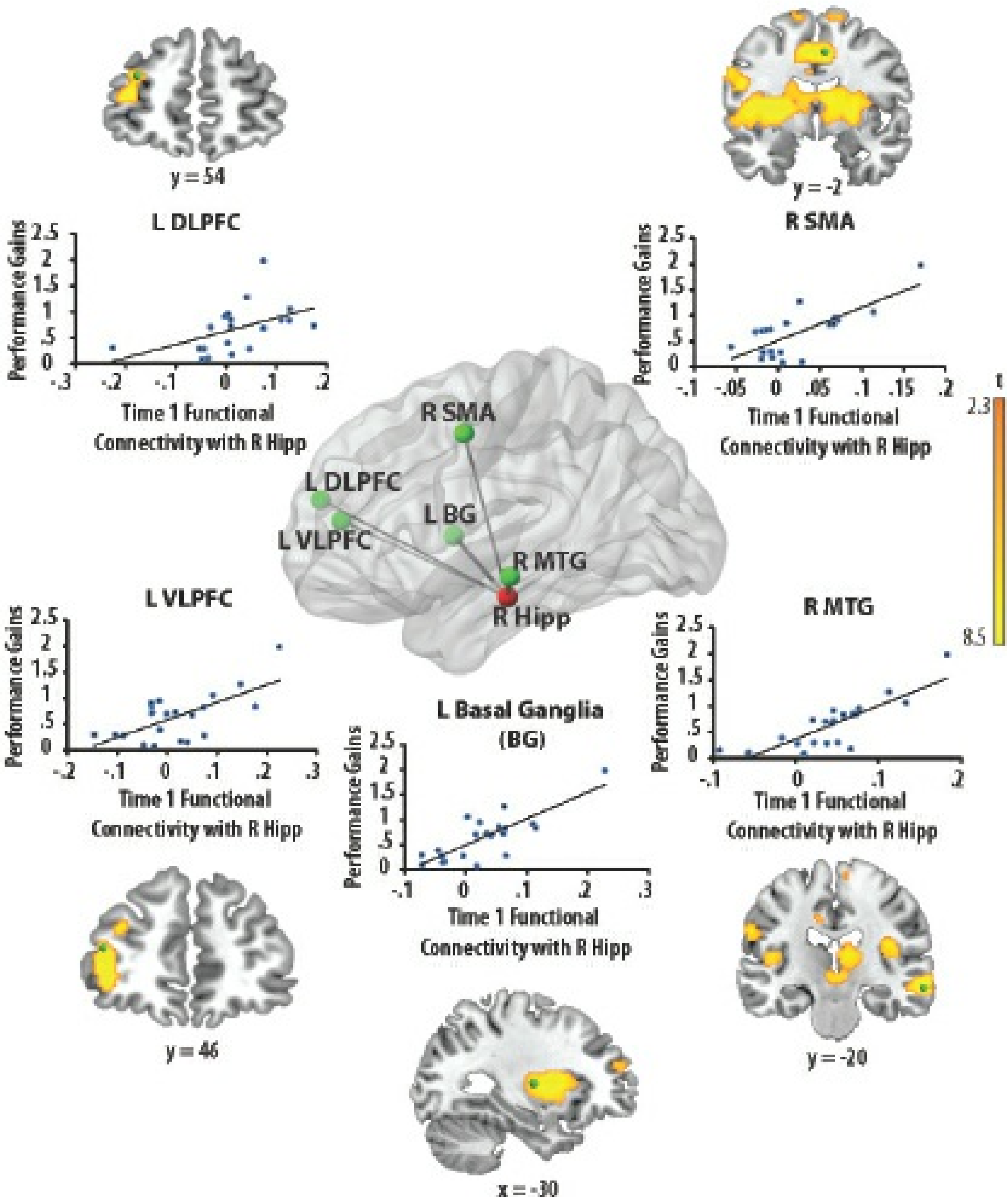

PLATE 7. Functional connectivity of the right hippocampus (R Hipp) before math tutoring. Composite 3D view of connectivity network is shown in the central panel with the right hippocampus seed ROI showing peak connectivity with the hippocampus. Surrounding panels show brain areas correlated with performance gains with tutoring. From Supekar et al. (2013, p. 8223). Reprinted with permission.

A systematic review of diffusion tensor imaging (DTI) studies that measure connectivity between regions of brain implicated in math also found evidence for the role of a left frontal–parietal white matter tract for calculation in children (Matejko & Ansari, 2015). In a study that looked at regions involved in different arithmetic operations in children, addition and multiplication was related to the arcuate fasciculus (an area associated with language and reading), but this relation was no longer significant when phonological processing was covaried (Van Beek, Ghesquière, Lagae, & De Smedt, 2014). This finding is of interest given the overlap of LDs in math and reading, and also given findings from longitudinal studies showing relations of early phonological processing to later arithmetic and reading abilities (Barnes et al., 2014).

Children with Mathematics LDs

Although studies of children with LDs in mathematics and of children with neurodevelopmental disorders associated with particular deficits in mathematics are few in comparison with WLRD, the evidence suggests that difficulties in math are associated with atypical function, brain structure, and connectivity in inferior parietal or temporoparietal white matter (Ashkenazi et al., 2013; Matejko & Ansari, 2015). In a summary of fMRI studies of children with math LDs, defined largely by calculation deficits, Ashkenazi et al. found reduced activation for number and quantity tasks in the right intraparietal sulcus and less activation in the ventral occipitotemporal stream bilaterally. In general, they highlighted the bilateral intraparietal sulci (see Figure 8.2), the ventral occipitotemporal region, and then recruitment of prefrontal areas and the cingulate gyrus for more complex arithmetical tasks (see Plate 7).

In aMRI studies, Kucian et al. (2014) used DTI to compare white matter integrity in children with LDs in math and age-matched controls. The main findings were that the brains of children with LDs in math showed deficits in fiber projections between parietal, temporal, and frontal brain regions. A region of brain implicated in the integration and control of distributed brain processes, the superior longitudinal fasciculus, showed reduced functional anisotropy. Such findings suggest that children with LDs in math might show underengagement in this neural network during the performance of math tasks. Yet, in an fMRI study of simple addition and subtraction, children with LDs in math showed *greater* activity in parietal, temporal, and frontal regions than their typically developing peers (Rosenberg-Lee et al., 2015). Furthermore, this pattern of hyperactivation was found for both the frontoparietal network involved in problem solving and requiring WM and attention and the default mode network, a brain network that is typically deactivated during cognitive-demanding tasks. Jolles et al. (2016) used a resting state (task-free) analysis of functional brain activity and found evidence of *hyperconnectivity* of the intraparietal sulcus with other components of the frontoparietal network bilaterally in a group with mathematics calculation LDs. In contrast, there were more frequent spontaneous low-frequency fluctuations in multiple frontal and parietal regions, supporting the view that there is a disruption of the entire network in children with math LDs.

In the only study involving magnetoencephalography and clearly defined groups with LDs in math and both math and reading, Simos et al. (2008) found generally suppressed neural activity in children with both reading and math LDs, possible related to more general difficulties with symbolic tasks. In contrast, for children with only math calculation difficulties, there was *increased* activity in right hemisphere inferior and superior parietal regions, with reduced reliance on left hemisphere parietal areas in comparison with typically developing children. Most notable were differences in the *timing* of neurophysiological activity, where children with only math calculation difficulties showed much earlier activation of prefrontal areas associated with numerical processing. In general, the timing differences were more pronounced than differences in which cortical regions were engaged during the math-processing tasks, suggesting a more poorly organized neural circuit.

Regardless of whether these mathematically important brain networks are under- or overactivated in children with LDs in math, these findings suggest differences in parietal organization and connectivity of its functional circuits in children with LDs in math rather than simply delayed maturation of parietal circuits. Functional neuroimaging studies addressing numerical and magnitude processing show a network of parietal circuits (inferior, superior, and intraparietal) extending to prefrontal areas. The involvement of different parts of this network varies depending on the numerical tasks as well as requirements for language processing (Ansari & Lyons, 2016). Activation is reduced with age and experience, with less reliance on frontal circuits (Rivera et al., 2005). In children with mathematics LDs, defined primarily on the basis of computational skills, the network is characterized predominantly by *increased* activity that is different from typically developing children in the degree and timing of activation and generally more disorganized (Iuculano et al., 2015; Simos et al., 2008).

Intervention Imaging Studies

Ansari and Lyons (2016) argued that studies of neural changes in relation to intervention would help elucidate the nature of the neural networks involved in mathematics. Although the neural networks show effects of training for academic skills such as reading (see Chapter 6), little research on the effects of intervention on brain has been conducted in math. Kucian et al. (2011) trained 16 children defined with developmental dyscalculia and 16 age-matched controls 8–10 years of age on a computer game that trained number line skills. The children completed the intervention 15 minutes per day, 5 days per week, for 5 weeks. In general, both groups improved in spatial representation of numbers and correctly solved math problems. In baseline fMRI assessments to a number line task, the children with developmental dyscalculia showed reduced bilateral parietal lobe activation; after intervention, there were reductions in parietal lobe activation in both groups, suggesting greater automaticity of calculation ability. However, neither calculation ability nor brain activation were normalized by this intervention, which is not surprising given its short duration.

In reading intervention studies, the focus is on teaching reading skills. For math, one prediction for such studies would be that intervention in arithmetic ought to primarily affect the frontoparietal network that emerges from the studies discussed above. However, an intervention study with 24 third-grade children revealed that different neural circuits might be involved in arithmetic fact *learning*. In this study (Supekar et al., 2013), the hippocampus, a structure involved in the laying down of long-term representations, was important for predicting response to an intervention that involved conceptual instruction and the development of fluency in simple arithmetic problem solving. Preintervention hippocampal volume and intrinsic functional connectivity of the hippocampus with prefrontal cortex and basal ganglia predicted improvements in arithmetic performance.

In a subsequent study, Iuculano et al. (2015) investigated the effects of an 8-week tutoring program (about 15–20 hours of intervention) focused on number knowledge, operations, and retrieval processes delivered individually using computer-delivered methods. The 15 children with math LDs (involving calculation, some also with problem-solving difficulties and none with word-reading problems) were compared with 15 typically developing children who also received the intervention. Before tutoring, children with mathematics LD showed increased activity in prefrontal, parietal, and ventral occipitotemporal brain regions commonly associated with math processing. After intervention, there was clear improvement in the accuracy of intervention-targeted areas and little difference in strength or pattern of neural activity relative to the typically developing children (see Plate 8). The results supported the normalization of brain function within mathematically relevant regions, paralleling reading intervention studies reviewed in Chapter 6.

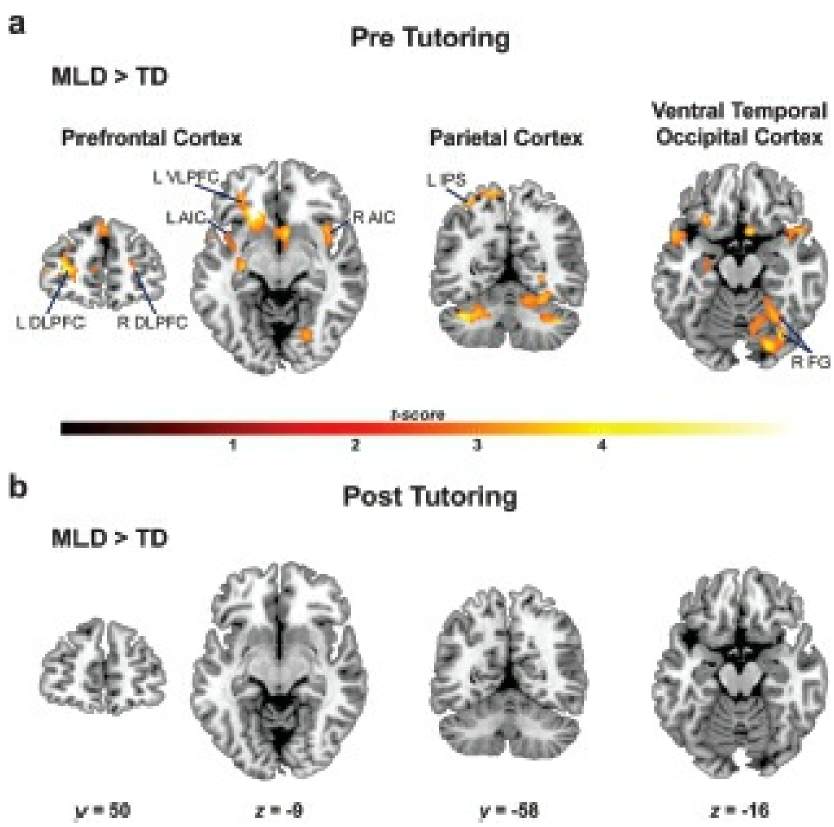

PLATE 8. Normalization of brain activity after 8 weeks of math tutoring. (a) Before tutoring, children

with math LDs showed significant differences in brain activation levels compared with TD children in the bilateral dorsolateral prefrontal cortices (DLPFC); the left ventrolateral prefrontal cortex (VLPFC); the bilateral anterior insular cortices (AIC); the left intraparietal sulcus (IPS); and in the ventral occipitotemporal cortex, including the right fusiform gyrus (FG). (b) After 8 weeks of tutoring, functional brain responses in children with math LDs normalized to the levels of TD children. From Iuculano et al. (2015, p. 4). Reprinted under Creative Commons License Deed 4.0.

Genetic Factors

Family Patterns

As in reading disabilities, an emerging research base demonstrates heritable factors in math disabilities. Math disabilities run in the family. Gross-Tsur, Manor, and Shalev (1996) found that 10% of children with a specific math disability had at least one other family member who complained of difficulties with math. Another 45% had another type of LD. Those with a family history of math disabilities were more likely to have persistent difficulties in math. Shalev et al. (2001) found that prevalence of math disabilities was quite high in mothers (66%), fathers (40%), and siblings (53%) of probands with math disabilities. Shalev et al. concluded that the prevalence of math disabilities was about 10 times higher in those with family members who had math disabilities than in the general population.

Twin Studies

Although genetically sensitive studies of math are still relatively fewer in number than studies of reading and of ADHD, activity in this domain has increased in the past 5 years. There are large-scale longitudinal twin studies conducted in the United States and the United Kingdom as well as smaller twin studies with more detailed assessment of academic and cognitive abilities. There are also studies that investigate horizontal relations or covariance between disabilities. For example, to what extent do genetic and environmental influences overlap for math and other disabilities such as those in reading decoding and reading comprehension. Other studies investigate the genetic and environmental contributions to different types of math abilities, including calculations, word problems, and magnitude processing.

Early genetic studies in mathematics yielded large variations in univariate heritability estimates, likely due to between-study variability in sample size, sample age, and the type of math ability measured. A systematic review of the few genetic studies in math up to 2008 reported moderate heritability (Willcutt et al., 2010b). Heritability estimates from more recent studies typically range from moderate to large (.42–.68), but the size of these effects as well as those of effects for shared and nonshared environment seem to depend on age and type of math ability measured, as discussed below.

According to the *continuity hypothesis*, estimates of heritability are assumed to be similar across levels of the academic skill in question. Data from behavior genetic studies of mathematics are consistent with this hypothesis using samples selected for math disability (Kovas et al., 2007), high math achievement (Petrill et al., 2009), and in unselected samples (Kovas et al., 2007). The genetic studies of math provide strong evidence for the continuity hypothesis, which is consistent with a dimensional view of LDs in which the attributes represent academic performance at the low end of a normal distribution.

In the past 5–10 years, behavior genetic studies of mathematics have become more multifaceted both within the study of mathematics ability itself and in the relation of math to other learning and behavioral disorders. As discussed in Chapter 6, *generalist* genes involved in learning have been proposed to underlie ability and disability in both reading and math, and in ADHD (Plomin & Kovas, 2005). Genomewide association studies suggest that about 50% of the correlation between reading and math is due to shared genetic effects and the contributions to this association are polygenic in nature (Donnelly, Plomin, & Spencer, 2014). This strong genetic overlap of reading and math is in keeping with epidemiological studies showing the high rates of comorbidity for LDs in reading and math (Gross-Tsur et al., 1996; Knopik & DeFries, 1999). Of interest in recent studies are analyses that more closely investigate the nature of the relation of math and reading. For example, when the relations of math and reading are disaggregated for reading decoding versus reading comprehension, mathematics and reading comprehension are found to share the greatest phenotypic and genetic overlap (Harlaar, Kovas, Dale, Petrill, & Plomin, 2012), and may, therefore, be more likely to share common cognitive correlates than other aspects of reading.

LDs in math also have phenotypic and genetic overlap with attention disorders (Gross-Tsur et al., 1996). Of interest to the discussion of cognitive sources of abilities and disabilities in math, the genetic relations of math and attention have recently been investigated for different types of attention symptoms. Greven et al. (2014) found a genetic relation for math and attention that was greater for inattention than for hyperactivity/impulsivity. These genetic findings converge with those from behavioral studies in which only symptoms of inattention are strongly related to math ability and disability (L. S. Fuchs et al., 2005a; Raghubar et al., 2009).

Why are these genetic studies of math that disaggregate for types of reading and ADHD important? Such findings help to refine behavioral and neurobiological models of LDs in math and reading. For example, the possibility that some genetic markers for inattention and math ability or for reading comprehension and math may be shared helps to narrow or constrain the search for genes involved in math ability/disability. Such findings may also be relevant for designing research on math interventions that are sensitive to more precise conceptions of the nature of heterogeneity or comorbidity.

In keeping with more precision in how reading and attention abilities are conceived and measured in genetically sensitive studies of math, some twin studies of math have looked at whether genetic and environmental sources of variance are similar for different types of math abilities including calculations, word-problem solving, and magnitude comparison. Hart, Petrill, Thompson, and Plomin (2009) found somewhat different genetic and environmental influences for math calculation, math problem solving, and math fluency. Estimates of genetic and environmental variance also varied with age, with heritability estimates generally becoming greater across time in keeping with genetic findings in other academic and cognitive domains.

For math problem solving, there was considerable genetic overlap with reading (for both decoding and comprehension) and general cognitive ability (a composite of verbal, nonverbal, and verbal WM tests). The strongest genetic influences within the domain of math and across time were on math fluency, and math fluency shared genetic overlap with general cognitive ability and RAN (but not reading), suggesting unique genetic influences that may be related to speed of processing. Interestingly, shared environmental influences were greater for calculations than for other math abilities. The authors suggested that this greater environmental influence may be seen for skills that are more directly instructed, particularly in countries in which there is more variability in this instruction (i.e., in the United States compared to the United Kingdom and East Asian countries in which there are national math curricula). Genetically sensitive studies of different types of math abilities in relation to math-related cognitive processes such as language and WM are beginning to emerge (e.g., Lukowski et al., 2014).

In terms of heritability of the types of “number sense” abilities that were discussed earlier, Tosto et al. (2014) found that performance on nonsymbolic magnitude tasks was modestly heritable (32%), a lower heritability estimate than those found in most other studies of math. In contrast to many other math skills studied in behavior genetic studies, nonshared environmental factors explained most of the variance (68%). One interpretation of these findings is that large approximate magnitude comparison is an ability that is evolutionarily important, and in keeping with other such abilities may not show much variability in heritability.

Summary: Neurobiological Studies

Neuroimaging studies of math processing in children have increased over the past few years and are relatively consistent in showing the importance of a left frontoparietal network for simple and more complex calculations. Neuroimaging studies of children with LDs in math, longitudinal studies of math that incorporate neuroimaging, and studies that investigate response of brain to intervention are burgeoning. The math network is characterized by predominantly increased activity that is different from typically developing children in the degree and timing of activity. This disorganized network becomes more focally organized with intervention and the few intervention-imaging studies of math intervention show primarily normalizing, not compensatory, activation patterns. Most studies of the heritability of math skills are twin studies that parse different cognitive skills into those that are heritable from influences that represent shared and nonshared environmental components. About 50% of the variability in mathematics skills can be explained by heritable factors. The idea of generalist genes that explain comorbid associations of reading, math, and ADHD is as prominent in this literature as in the reading literature (Chapter 6). Estimates of the influence of heritable and shared environmental influences vary across domains of mathematical competency. There is little research on candidate genes specific to mathematical competencies; indeed, in contrast to reading, where the examination of candidate genes for dyslexia is common (Chapter 6), there is little evidence of genes specific to mathematics LDs, with even more of a focus on dimensional concepts and individual differences.

INTERVENTIONS FOR MATHEMATICS DISABILITIES

The mathematics curriculum is complex and has multiple components. It is unclear how different curricular components relate to one other and whether strengthening performance in one domain can be expected to transfer to other domains. Such understanding is important not only for an empirically guided framework for LD identification, but also for guiding the organization of curriculum and the design of instruction. Failure to produce strong performance across curricular components creates challenges for addressing the needs of students with mathematics LDs. For example, remediating whole-number calculation skill cannot be expected to improve whole-number word-problem skill (e.g., L. S. Fuchs et al., 2009). Moreover, once performance on whole-number calculations and word problems is strengthened, it is likely that intervention on fraction calculations and on fraction word problems is still required.

At the same time, as we discussed in Chapter 5, general instructional principles effectively apply across curricular components even as the specific task analyses, explanations, materials, and instructional activities differ. We begin with core classroom instruction and then focus on interventions designed to supplement classroom instruction for these students. We highlight interventions designed to address three major stumbling blocks for these students: simple addition and subtraction, word problems, and fractions concepts and operations.

Core Classroom Instruction

Principles for effective classroom instruction for the students with mathematics LDs are similar to the instructional practices required to address other LDs. Classroom instruction should be explicit, academically focused, and foster high levels of engagement and frequent opportunities for student response and feedback. Some basal or developmental programs used for students with LDs have these characteristics. For example, Connecting Math Concepts (Engelmann, Carnine, Engelmann, & Kelly, 1991) is based on a behavioral/task-analytic model frequently used for primary- and elementary-age students with LDs. It contains highly structured lessons involving frequent teacher questions and student answers. Studies have demonstrated the efficacy of this program for students with LDs (Adams & Carnine, 2003). Direct Instruction programs in math yielded an effect size of d = .55 (95% confidence interval .46–.65), in the medium range (Stockard et al., in press). The effects were robust across multiple methodological and sample variations.

Mathematics PALS (e.g., L. S. Fuchs et al., 1997) incorporates similar instructional principles, but takes a different approach. As a classwide peer-tutoring program, it is designed as a supplement to the classroom teacher's basal program whose goal is to differentiate instruction and provide more intensive practice. All students in the class are paired to work in highly structured ways, based on thorough task analyses of component mathematics skills. Each pair works on a skill on which the lower-performing student in the pair requires support. PALS instruction is explicit, targets procedural skills as well as conceptual knowledge, and provides students with carefully structured peer-guided practice, feedback, and cumulative review. When PALS is implemented in general education classrooms as a supplement to the teacher's instructional program, it improves outcomes for students with LDs as well as for their low-, average-, and high-performing classmates at kindergarten through grade 6 (e.g., L. S. Fuchs et al., 1997; L. S. Fuchs, D. Fuchs, & Karns, 2001a; L. S. Fuchs, D. Fuchs, Karns, Yazdian, & Powell, 2001b).

A number of classroom teaching techniques have been shown to be generally useful in helping students with LDs develop mathematics skills. Rivera and Smith (1988), who summarized research on the value of modeling in teaching calculation skills, found that, just as in reading comprehension (Chapter 7), teacher demonstrations of calculation algorithms and higher-level procedural steps were effective in increasing both calculation and problem-solving skills. Lloyd (1980) tested the value of strategy training with students who are deficient in mathematics. In this type of intervention, a task analysis of the relevant cognitive operation is modeled and explained to students. When students have mastered the component skills, strategies are provided that help the students integrate the steps and apply them in different problem-solving contexts. Finally, cognitive-behavioral models of intervention have given rise to the development of self-instructional strategy techniques to help guide students with LDs through a variety of problem-solving contexts (Hallahan et al., 1996). A key component in this technique is to teach a student first to verbalize the steps that should be used in solving a particular math problem. Once the student has mastered the application of the problem-solving algorithm, the student is taught to self-instruct, but using subvocal directions. This technique has been shown to be useful with both elementary-age students (Lovitt & Curtiss, 1968) and adolescents (Seabaugh & Schumaker, 1993).

Supplemental Intervention

Even with well-designed classroom instruction, many students with mathematics LDs experience continued difficulty and require supplemental intervention. Such intervention occurs primarily in the form of small-group tutoring, which is usually designed as a Tier 2 *supplemental* intervention (see Chapter 5). In this section, we illustrate the application of the design principles in Chapter 5 in validated tutoring programs for addressing three major stumbling blocks for students with mathematics LDs at the elementary grades. We begin with word problems, a domain with which many children with math LDs struggle. Then we address arithmetic skills, a domain that has received more attention than most components of the mathematics curriculum, and fractions, an area that has received relatively little attention. We conclude with a discussion of the limitations of intervention research in mathematics and identify areas for future research.

Word-Problem Intervention

Pirate Math (see L. S. Fuchs et al., 2009) is an example of a well-researched intervention for word-problem difficulties, while building the foundational skills required for word problems: arithmetic, procedural calculations, and prealgebraic knowledge. The program incorporates a pirate theme because within this schema-broadening instructional program, students are taught to represent the underlying structure of word-problem types using equations that include *x* to represent the unknown quantity. "They find *x*, just like Pirates find *x* on treasure maps."

Pirate Math comprises four units: an introductory unit that addresses foundational skills as well as three word-problem units each focused on a different type of word problem. By teaching word problems in terms of "type of word problems," students learn to search novel word problems (ones they have not seen before) in terms of the types of word problems for which they have learned solution strategies. A large percentage of simple word problems can be categorized as three problem types: combine, compare (referred to instructionally as difference), and change problems. Teaching word problems in terms of categories illustrates the instructional principle of minimizing the learning challenge: instead of viewing every novel word problem as unfamiliar, students learn three word-problem types, each with a predictable solution strategy.

Every tutoring lesson in Pirate Math is manualized as a series of scripts in a defined scope and sequence. Scripts are studied, not read or memorized, so that teachers' instructional style is authentic. Teachers can also differentially emphasize different components of the scope and sequence, particularly if they are teaching in homogeneous small groups composed of children with similar needs. Pirate Math runs for 16 weeks, with 48 sessions (three per week). Each session lasts 20–30 minutes. Instruction, as outlined below, is *explicit and systematic*; it is designed with care to *minimize the learning challenge*; it is rich in *concepts*; it incorporates *systematic practice* as well as *cumulative review*; and it relies on *systematic reinforcement* to encourage good attention, hard work, and accurate performance.

The introductory unit addresses mathematics skills foundational to word problems. This includes *counting strategies* for deriving answers to arithmetic problems, algorithms for two-digit addition and subtraction procedural calculations, methods to find *x* in any position in simple equations (e.g., $a + b = c$; $d - e = f$), and strategies for checking work in word problems.

Each of the three word-problem units focuses on one word-problem type and, after the first word-problem-type unit, subsequent units provide systematic, mixed cumulative review that includes previously taught problem types. Each word-problem unit lesson comprises six activities. The first is the *counting strategies review and flash card warm-up* already described. With *word-problem warm-up*, the next activity, students explain how they solved a word problem from the previous day's paper-and-pencil review.

Conceptual and strategic instruction is the third activity. It is the heart of the lesson. Tutors provide explicit instruction in the underlying structure of and steps in solving the three types of word problems, along with instruction on identifying and integrating transfer features to broaden students' understanding for each problem type. This approach allows them to identify the word-problem types when novel problems include problem features that are not essential to the problem type, such as unfamiliar vocabulary or irrelevant information or tables that present relevant information. The tutor relies on role playing, manipulatives, instructional posters, modeling, and guided practice. In this component of the lesson, students solve three word problems, with decreasing amounts of support from the tutor.

Total is the first problem type addressed. In the total unit, tutors teach students to RUN through a problem: a three-step strategy prompting students to Read the problem, Underline the question, and Name the problem type. Students used the RUN strategy across all three problem types. Next, for each problem type, students are taught a "meta-equation" to represent the underlying structure of that problem type, and they use this meta-equation (e.g., for *total* problems, the meta-equation is Part 1 + Part 2 = Total, or P1 + P2 = T) to structure their solution strategy. The strategy for difference problems and change problems follows similar steps but uses meta-equations specific to those problem types.

For each problem type, explicit transfer instruction occurs in multiple ways. Students are taught that because not all numerical values in word problems are relevant for finding solutions, they should identify and cross out irrelevant information as they RUN through the problem. They learn to recognize and solve word problems with the missing information not only in the easiest, third slot of the meta-equation, but when the missing information occurs in the first or second position of the meta-equation. They learn to apply the problem-solving strategies to word problems with more complex calculations or with money.

Sorting word-problems is the fourth activity. Tutors read aloud flash cards, each displaying a word problem. The student identifies the problem type, placing the card on a mat with four boxes labeled “Total,” “Difference,” “Change,” or “?”

Paper-and-pencil review is the final activity, in which students have 2 minutes to complete number sentences asking the student to find *x* and 2 minutes to complete a word problem. Tutors provide corrective feedback and note the number of correct problems on the paper.

Pirate Math also includes a systematic *incentive* program. Throughout each session, tutors use a timer, which is set to ring at unpredictable intervals. If all students in the group are “on task” when the timer rings, each earns a gold coin. If one or more students are not on task, no one earns a coin. Students can also earn gold coins for completing Bonus Problems correctly. At the end of the lesson, each gold coin is placed on the student’s individual “treasure map,” which leads to a treasure box; when reached, the student chooses a small prize.

Pirate Math is validated for improving word-problem skills among students with math LDs (e.g., L. S. Fuchs et al., 2008a, 2009, 2011). For example, L. S. Fuchs et al. (2009) identified third-grade students with substantial difficulty in calculations and word problems. These children were randomly assigned to three conditions: control (the school program without any research-based mathematics tutoring); 13 weeks (three times per week) of Pirate Math intervention; or 13 weeks (three times weekly) of Math Flash intervention, a validated program focused entirely on calculation with no word problem instruction. Students were pre- and posttested on calculation and word-problem measures. On calculations, Pirate Math and Math Flash students improved comparably and significantly more than students in the control group. The effect size comparing Math Flash to the control group was large (0.85). The effect size comparing Pirate Math to the control group was similar (0.72). Yet, given that Pirate Math allocated only 5 minutes of every session to calculations (whereas Math Flash spent 20–30 minutes per session on calculations), Pirate Math's effects on calculations are noteworthy. At the same time, however, effects on word problems clearly favored Pirate Math. The effect size comparing Pirate Math to the control group was large (0.89), and there was no significant difference between Math Flash and the control group. The effect size comparing Pirate Math to Math Flash was 0.72. These findings indicate that Pirate Math, a comprehensive intervention focusing on word-problem solving that includes practice on foundational arithmetic skills, provides a general benefit for math-problem solving and calculations, which is not seen with programs restricted to instruction in only foundational skills (see Chapter 5).

Arithmetic Intervention

Students with mathematics LDs show consistent delays in the adoption of the efficient counting procedures associated with arithmetic success, make more counting errors when executing counting strategies to solve arithmetic problems, and fail to make the shift from counting strategies toward memory-based retrieval, as typically developing children do (e.g., Geary, Hoard, & Bailey, 2012a; Goldman, Pellegrino, & Mertz, l988). Most of these students with LDs eventually catch up to peers in skilled use of counting procedures, but difficulty with retrieval tends to persist (Geary et al., 2012a; Jordan et al., 2003). Students with LDs retrieve fewer answers from memory and when they do retrieve answers, they commit more errors (Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007).

Some researchers consider simple arithmetic fluency to be a signature deficit of mathematics LDs (e.g., Geary et al., 2012a; Goldman et al., 1988; Jordan et al., 2003). Moreover, remediating arithmetic deficits in older students can be difficult (L. S. Fuchs et al., 2010b). For these reasons, there is a pressing need for early intervention in arithmetic. Few randomized control trials have been conducted with early arithmetic intervention as their intervention focus.

We identified five randomized control trials assessing the efficacy of arithmetic intervention for first-grade students with mathematics difficulties. L. S. Fuchs et al. (2006b) conducted a randomized control trial to assess the efficacy of practice with simple arithmetic problems. An addition or subtraction problem with its answer briefly flashed on a computer screen; then students generated the problem and answer from short-term memory. This is based on the assumption that practice strengthens retrieval when problems and answers are simultaneously active in WM (Geary, 1993). Compared to an analogous computer-assisted spelling practice condition, arithmetic practice (10 minutes, twice weekly for 18 weeks) produced significantly better performance for addition but not subtraction; effect sizes were 0.95 and −0.01.

Two randomized control trials combined practice with number knowledge tutoring. In L. S. Fuchs et al. (2005a), tutoring occurred three times per week for 16 weeks. Each session included 30 minutes of tutor-led instruction designed to build number knowledge plus 10 minutes of computerized arithmetic practice, as just described. Results favored tutoring over a no-tutoring control group on other mathematics measures (for concepts and applications, effect size = 0.67; for procedural calculations, effect size = 0.40–0.57; for word problems, effect size = 0.48), but effects were not reliable on simple arithmetic (effect sizes = 0.15–0.40). Bryant et al. (2011) also integrated practice with number knowledge (four times per week for 19 weeks), this time with greater success on arithmetic outcomes. In each session, 20 minutes were devoted to number knowledge and 4 minutes to practice. Effects were significantly stronger for tutoring compared to a no-tutoring control group on simple arithmetic (effect size = 0.55).

Only one randomized control trial focused exclusively on number knowledge. Smith, Cobb, Farran, Cordray, and Munter (2013) evaluated Math Recovery, in which tutors adapt lessons to meet student needs as reflected on embedded assessments. Tutors introduced tasks and had students explain their reasoning, but practice was not provided. Tutoring was to occur four to five times per week, 30 minutes per session, across 12 weeks, but the median number of sessions was 32. Results favored intervention over the control group on arithmetic, but the effect size of 0.15 was substantially lower than in studies that combined number knowledge tutoring with practice and took a more explicit approach to instruction.

These studies suggest that first-grade intervention is effective for enhancing some forms of mathematics learning among at-risk first graders—when intervention combines number knowledge tutoring with practice.

However, these studies contrasted the intervention condition against a no-tutoring control group, without including two competing intervention conditions. So these studies do not provide the basis for understanding whether the effects of intervention are simply due to more instruction. They also fail to inform practitioners about what components of intervention contribute to positive effects.

L. S. Fuchs et al. (2013c) therefore contrasted two intervention conditions as well as a comparison group that received typical instruction. The major emphasis in intervention was developing interconnected knowledge of numbers, but a small portion of each session was devoted to practice. In one condition, practice was designed to reinforce the relations and principles that serve as the basis of reasoning strategies and that support fact retrieval. The other form of practice was more rote: it was designed to promote quick responding and use of efficient counting procedures to generate many correct responses and thereby form long-term representations to support retrieval. Both practice conditions occurred on the same content, encouraged strategic behavior, and provided immediate corrective feedback.

There were two major distinctions between the intervention conditions. First, one intervention condition encouraged a variety of number-principle strategies (i.e., relying on number lists, arithmetic principles such as cardinality, the commutative principle, subtraction as the inverse of addition, and efficient counting procedures). The other intervention condition only encouraged efficient counting strategies. The second distinction involved practice. In the condition that encouraged a variety of strategic behaviors, practice did not involve speeded execution of the student's chosen strategy; rather, the focus was on executing strategies thoughtfully to emphasize number knowledge. In contrast, in the condition that relied exclusively on counting strategies, practice was speeded. Below, we use the terms *nonspeeded practice* and *speeded practice* to refer to these conditions.

To understand the efficacy of number knowledge tutoring when combined with speeded versus nonspeeded practice, each tutoring condition was compared against an at-risk no-tutoring control group that received the same classroom instruction as the tutored groups. To understand how the type of practice affects learning and whether the effects of intervention are attributable to more than simply providing extra instructional time, we contrasted the two tutoring conditions against each other. To provide insight into whether different forms of intervention help narrow the achievement gap, we included a group of low-risk classmates, who received the same classroom instruction as the at-risk tutored and control groups. The study included more than 900 children in 40 schools in 233 classrooms.

Tutoring occurred for 16 weeks, three times per week, 30 minutes per session. In both intervention conditions, 25 minutes of each 30-minute session were the same, designed to foster number knowledge. The last 5 minutes, which involved practice, differed. To foster engagement, the program uses a space theme. Children are encouraged to "blast off into the math galaxy" by improving their mathematics knowledge; some manipulatives are shaped as space rockets.

The results showed that on arithmetic, number knowledge intervention with nonspeeded practice produced significantly better learning compared to at-risk control students who did not receive tutoring. The effect size was 0.38. This lends support to studies indicating the important role number knowledge plays in developing competence with simple arithmetic (De Smedt, Verschaffel, & Ghesquière, 2009; Duncan et al., 2007; Rousselle & Noël, 2007). At the same time, nonspeeded practice did not help at-risk students narrow the achievement gap (effect size = 0.07). By contrast, incorporating speeded practice in number knowledge intervention produced superior arithmetic improvement compared to low-risk classmates, with an effect size of 0.39, thereby narrowing the achievement gap. Speeded practice was also substantially more effective than number knowledge intervention with nonspeeded practice (effect size = 0.51). This extends earlier randomized control trials by isolating the effects of speeded practice, delivered in the context of intervention to build number knowledge. Results indicate a substantial role for speeded practice in promoting arithmetic learning. We also found no evidence that speeded practice inhibits development of number knowledge or word-problem skill, despite the fact that rote responding was involved in speeded practice. Both number knowledge intervention conditions produced comparable number knowledge and word-problem learning, which was superior to at-risk control students.

Fractions Interventions

Half of middle and high school students in the United States are still not proficient with the ideas and procedures taught about fractions in the elementary grades (Hiebert & Wearne, 1985; National Council of Teachers of Mathematics, 2007; Geary et al., 2008a). Yet, competence with fractions is considered foundational for learning algebra, for success with more advanced mathematics, and for competing successfully in the American workforce (NMAP, 2008; Geary, Hoard, Nugent, & Bailey, 2012b). For these reasons, the NMAP recommended that high priority be assigned to improving performance on fractions, a theme reflected in the Common Core State Standards (CCSS; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Therefore, getting an early start with intervention on fractions for students with mathematics LDs is important. Some research has, however, focused on older students with LDs.

The most successful intervention studies have adopted an explicit instructional approach; some focus entirely on fraction procedures while others attend to fraction concepts. Using a multiple-baseline design, Joseph and Hunter (2001) demonstrated experimental control for a cue-cards strategy across three eighth-grade students with mathematics LDs. A teacher initially taught students to use the cue card, which supported a three-pronged strategy for adding or multiplying fractions and reducing answers. After students showed competence in applying the strategy, they used the cue card while solving problems. In a maintenance phase, the cue card was removed. All three students showed substantial improvement with introduction of the cue card strategy, and maintenance was strong. The study focus was, however, entirely procedural in terms of instruction and outcome.

Kelly, Gersten, and Carnine (1990) also took an explicit approach to fraction instruction, but focused on procedures and concepts. They randomly assigned 28 high school students with mathematics LDs from three classes to 10 sessions of teacher-mediated videodisc instruction or conventional textbook instruction. Direct instruction was employed in both conditions, but only videodisc instruction provided mixed problem-type instruction, separated highly confusable concepts and terminology in early instructional stages, and provided a broader range of examples to avoid misconceptions. Both groups improved substantially from pretest (40% on a 12-item test) to posttest (96% vs. 82%), with the videodisc group improving significantly more. Yet, despite the instructional focus on concepts and procedures, the fraction measure was largely procedural. A few items required students to name fractions from pictures or distinguish numerators from denominators; the remaining items were procedural.

In two other studies, intervention focused primarily on understanding of fractions and assessed outcomes on concepts as well as procedures. Butler, Miller, Crehan, Babbitt, and Pierce (2003) contrasted two explicit instruction conditions with 50 middle school students with mathematics LDs. Both conditions carefully transitioned students from a conceptual emphasis, largely based on part–whole understanding, to algorithmic rules for handling fractions, and from visual to symbolic representations. Only one condition included concrete manipulatives. Both groups significantly improved across 10 sessions. On one measure, in which students circled fractional parts of sets, those who received 3 days of manipulatives improved significantly more; on the other four measures, the difference between conditions was not significant, providing minimal evidence regarding the importance of concrete representations. Without random assignment or a control group, however, conclusions are tentative.

These studies provide the basis for only tentatively concluding that explicit instruction, based on part–whole understanding of fractions, enhances fraction learning among middle and high school students with mathematics LDs. Yet, none of these studies focused on younger students, when a strong focus on fraction concepts and procedures begins in the curriculum. To extend the focus to younger students, L. S. Fuchs et al. (2013c) designed and tested the efficacy of intervention at fourth grade. This intervention focused primarily on conceptual understanding, which is important for learning and maintaining accurate procedures with fractions (Hecht et al., 2003; Mazzocco & Devlin, 2008; Ni & Zhou, 2005). Two types of conceptual knowledge (Kieren, 1993) were addressed. The first was part–whole understanding, with which a fraction is understood as a part of one entire object or a subset of a group of objects. This type of understanding is typically represented using an area model, in which a region of a shape is shaded or a subset of objects is distinguished from the remaining objects. The second type of conceptual knowledge, the measurement interpretation of fractions, reflects cardinal size (Hecht et al., 2003) and is often represented with number lines (e.g., Siegler, Thompson, & Schneider, 2011). In American schools, fractions are taught primarily via area models that underpin part–whole understanding. Measurement understanding is assigned a subordinate role (addressed later and with less emphasis).

This study was innovative because the major emphasis was the measurement interpretation of fractions—although we also incorporated a smaller amount of time on part–whole interpretations to build on students' incoming understanding of fractions, as addressed in their classrooms. In emphasizing the measurement interpretation, we sought to avoid understanding of fractions exclusively as part–whole relationships, which may create difficulty for conceptualizing improper and negative fractions (NMAP, 2008). A focus on the measurement model is also in line with the fourth-grade CCSS's focus on understanding of fraction equivalence and ordering (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010).

This intervention program, *Fraction Face-Off!* (L. S. Fuchs, Schumacher, Malone, & Fuchs, 2015c), included 36 lessons taught over a 12-week period (three 30-minute lessons per week). In line with the measurement interpretation of fractions, the content focuses primarily on representing, comparing, ordering, and placing fractions on a 0 to 1 number line. This focus is supplemented by attention to part–whole understanding (e.g., showing objects with shaded regions) and fair shares representations to build on classroom instruction. In this way, number lines, fraction tiles, and fraction circles are used throughout the lessons, with stronger emphasis on part–whole representations in early lessons. In Lesson 22 (of 36), after fraction concepts are well established, the program also introduces fraction calculation procedures. Throughout the program, the focus is on proper fractions and fractions equal to 1; improper fractions greater than 1 are introduced with addition and subtraction of fractions. To reduce calculation demands, denominators did not exceed 12 and excluded 7, 9, and 11. This tutoring content mirrored classroom instruction with the following exceptions. Fraction Challenge's focus is narrower, with greater emphasis on measurement understanding, whereas classroom instruction emphasized part–whole understanding more than intervention; Fraction Challenge focuses on calculations substantially less than classroom instruction; and Fraction Challenge uses a more limited pool of denominators.

As with Pirate Math (word-problem intervention described above) and Number Rockets (L. S. Fuchs, D. Fuchs, & Bryant, 2005b) (arithmetic intervention described above), the fraction program also encourages students to regulate their attention and behavior and to work hard. Tutors teach students that *on-task behavior* means listening carefully, working hard, and following directions and that on-task behavior is important for learning. Tutors set a timer to beep at three unpredictable times during each lesson. If all students are on task when the timer beeps, all students receive a checkmark. To increase the likelihood of consistent on-task behavior, students cannot anticipate time intervals. Also, on each practice sheet, two of 16 problems are bonus problems. As the tutors score the practice sheet, they reveal which problems are bonus items. Students receive a checkmark for each correctly answered bonus problem. At the end of the lesson, tutors tally checkmarks for each student and award them with a "half dollar" per checkmark. At the end of each week, students shop at the "fractions store" to spend money earned during tutoring. All items in the store are listed in whole-dollar amounts at three price points so students must exchange half dollars for whole dollars and determine what they can afford. In this way, to use the fraction store, students must rely on their fraction knowledge, while exercising judgment about buying a less expensive item versus saving for a more expensive one. In Lesson 19, we replace half dollars with quarter dollars.

In the study of this intervention, when randomly assigning students to intervention versus control conditions L. S. Fuchs et al. (2013c) stratified participants by the student's level of incoming mathematics deficits on whole numbers (severe vs. less severe) while also stratifying by classroom. Participants were 259 fourth graders with mathematics difficulty, from 53 classrooms in 13 schools: 129 intervention students (60 more severe and 69 less severe), and 130 control students (66 more severe and 64 less severe). Another 292 students were low-risk classmates. With this sample, two measures were administered to isolate the type of the measurement interpretation of fractions. On comparing fractions (in which students place a greater than, less than, or equal sign between two fractions), the effect size favoring intervention over control children was 1.82, and the achievement gap between at-risk tutored students and their low-risk classmates narrowed, while the gap for at-risk control students increased. On fraction number line (in which students place a fraction on a 0–1 number line), the effect size was 1.14. Fraction number line data was not collected on low-risk classmates, but the posttest performance of intervention students was at the 75th percentile for a normative sample.

Because the alignment between fraction instruction and assessments for comparing fractions and fraction number line was greater for intervention than for classroom instruction, effects were also considered on released items from the National Assessment of Educational Progress (NAEP). NAEP was not aligned with intervention; it focused with comparable emphasis on measurement and part–whole understanding. Here, effects were also significant and strong. The effect size favoring at-risk intervention students over control was 0.94, and the achievement gap favoring low-risk classmates over control students remained large, while the achievement gap for intervention students decreased substantially or was eliminated.

Moreover, although classroom instruction focused on calculations more than in the intervention group, effects again favored intervention students over control. Here the effect size for fraction calculations favoring intervention students was 2.51; the achievement gap between at-risk intervention students and their low-risk classmates narrowed, even as the gap for at-risk control students increased; and intervention students' posttest performance actually exceeded that of low-risk classmates. Given that classroom instruction allocated substantially more time to calculations, this suggests that understanding fractions, perhaps specifically the measurement understanding of fractions, transfers to procedural skill, at least with respect to adding and subtracting fractions (Hecht et al., 2003; Mazzocco & Devlin, 2008; Ni & Zhou, 2005; Siegler et al., 2011).

This study extends fraction intervention research by focusing primarily on the measurement and interpretation of fractions, rather than on part–whole understanding as in earlier work, and by targeting fourth-grade students (rather than middle or high school students). Another interesting extension to the literature concerns the focus on risk severity. Response to intervention was comparable for students with mathematics LDs (more severe risk) versus students with low mathematics achievement (less severe risk)—when risk was defined in terms of whole-number deficits. That is, there were no significant interactions between risk severity and study condition, and effect sizes were similar for more versus less severe student groups. This intervention study illustrates that intervention designed to foster understanding of fraction magnitude for fourth graders with mathematics LDs is effective. The effects are strong, with the achievement gap for at-risk learners substantially narrowed.

The Problem of Transfer

The interventions reviewed for different components of mathematics raise important questions about transfer across components of the mathematics curriculum. Should mathematics skills be taught as isolated skills or components, or should more comprehensive interventions be sought? The intervention literature indicates that, although transfer may occur across some domains, it is decidedly limited across others. Across a series of randomized control trials, L. S. Fuchs and colleagues investigated the efficacy of calculations or word-problem tutoring on calculations versus word-problem outcomes. The first study investigated the effects of first-grade tutoring on multiple components of the curriculum (L. S. Fuchs et al., 2005a). In the second study, a major focus was the efficacy of tutoring specifically to enhance fluent and accurate performance on math facts among first graders at risk for mathematics LDs (L. S. Fuchs et al., 2013b). The third focused again on math facts tutoring, this time with third-grade students (L. S. Fuchs et al., 2009). The fourth study investigated the separate and combined effects of primary prevention and supplemental tutoring on word problems (L. S. Fuchs et al., 2008a) with at-risk third graders.

Across these studies, findings indicated the following: Transfer may occur from math facts tutoring to procedural calculations, indicating a connection between these two types of calculation competence. However, results indicated that math facts tutoring did not transfer to word-problem performance, even when those word problems required students to answer math facts to derive solutions.

A more recent study (L. S. Fuchs et al., 2014a) deliberately focused on the issue of transfer, not only between calculations and word problems but also from calculations to prealgebraic knowledge and from word problems to prealgebraic knowledge. Participants were 1,102 children in 127 2nd-grade classrooms in 25 schools. Teachers were randomly assigned to three conditions: calculations intervention, word-problem intervention, and business-as-usual control. Intervention, which lasted 17 weeks, was designed to provide research-based linkages between calculations or word problems (depending on condition) to prealgebraic knowledge. Multilevel modeling suggested calculation intervention improved calculation but not word-problem outcomes; word-problem intervention enhanced word-problem but not calculation outcomes. Again, these findings indicated that calculations and word problems represent distinct forms of mathematical cognition, but transfer is limited, and more generally shows that teaching lower-level skills like calculations and facts can be done in the context of more complex, higher-level interventions. Thus, a major need for mathematics intervention is the development of more comprehensive programs (see discussion of Pirate Math above).

Transfer issues are also relevant for higher-level mathematics. In general, mathematics, more than reading, is potentially complicated by the fact that the elementary school curriculum comprises multiple components within and across the grades. This problem becomes more complicated in high school, where the components of the mathematics curriculum (e.g., geometry, trigonometry, calculus, and algebra) diverge more dramatically than in the earlier grades. To illustrate, algebra involves symbolizing and operating on numerical relationships and mathematical structures. Algebraic expressions can be treated procedurally, by substituting numerical values to yield numerical results (Kieran, 1993). This suggests that understanding of arithmetic principles, as in calculations or word problems, involves generalizations that are algebraic in nature (NMAP, 2008).

The arguments on both sides, however, are guided largely by rational analyses of the content of the two domains, rather than by empirical studies that test for transfer between calculation, word problems, and algebra. Only a handful of relevant studies, all correlational, have investigated the connection between arithmetic and algebra. Lee et al. (2011) and Tolar et al. (2009) found that arithmetic calculations serve as a platform for algebra. But Lee et al. used word problems as a proxy for the prealgebraic knowledge outcome; Tolar et al. focused on college students; and both studies limited their focus on arithmetic to calculations (they did not consider word problems as a predictor of algebra outcome). L. S. Fuchs et al. (2012) simultaneously considered calculations and word problems and found both uniquely predicted third graders' understanding of the equal sign and variables, thereby providing support for a connection. L. S. Fuchs et al. (2014b), discussed in the preceding paragraphs, extended this correlational research by examining transfer within an experimental framework. Results indicated that the positive effects of calculation intervention on calculation outcomes failed to transfer to prealgebraic knowledge. By contrast, the positive effects of word-problem intervention on word-problem outcomes *did* transfer to prealgebraic knowledge. Again, teaching foundational skills in the context of more complex interventions may be more powerful.

Summary: Intervention for Mathematics Disabilities

Mathematics intervention studies focused on LDs, some of which were described in this chapter, provide the basis for thoughts about the power as well as the limitations of mathematics intervention for students with LDs. We focus first on the power. The literature indicates it is possible to design intervention programs, using the design principles outlined in Chapter 5, to enhance the learning of these students. Table 8.1 summarizes important principles for teaching mathematics to children with LDs. Interventions that incorporate explicit instruction that minimizes the learning challenge (i.e., supports success); provide students with a strong conceptual foundation and efficient procedural strategies; embeds regular, strategic, and cumulative practice; and incorporates methods to promote self-regulated learning strategies are generally efficacious. Students who receive such interventions experience substantially greater success than if left in the general education program without such intervention (as represented in the control conditions in these studies). When students with LDs do not receive these intervention services, the gap between their level of math performance and that of low-risk classmates grows, making it increasingly difficult for these children to profit from classroom instruction.

TABLE 8.1. Intervention: Fundamental Principles for Teaching Mathematics LDs

1. Teach different components explicitly (e.g., fact retrieval, procedures, problem solving) for whole numbers and rational numbers.
2. Teach arithmetic (math facts) in the context of number knowledge principles to support understanding and strategic counting.
3. Provide speeded strategic practice, with immediate corrective feedback on errors, to build long-term associations in memory and encourage automatic retrieval of answers.
4. For procedural computations, explicitly teach the most efficient algorithms as rules, while providing the conceptual basis for why those procedures work. Begin with worked examples, gradually transferring responsibility to the learners, while providing practice with corrective feedback and cumulative review across problem types.
5. For word-problem solving, teach problem types (e.g., combine problems, compare problems), introducing one problem type at a time, but systematically providing cumulative review across all taught problem types and practice in sorting problems into problem types. For each problem type, begin by providing the conceptual basis for the problem type; then explicitly teach the most efficient solution strategy, with worked examples and gradual transfer of work to the learner.
6. Explicitly teach for transfer by explaining the ways in which problems may look novel but still represent the taught problem types.
7. Promote self-regulation and independence to promote generalization.

Even so, some important limitations exist for dramatically reducing the need for ongoing and intensive services for some of these students with LDs. First, for many components of the mathematics curriculum, interventions have not been designed or systematically tested. Second, despite substantial effects for some interventions, as illustrated in this chapter, not all students respond. Lack of universal response is well documented in the literature, not only for mathematics but also for reading, and estimates of inadequate responsive range from 25 to 40% of students with LDs (O'Connor & Fuchs, 2013). This is the case, at least, when 12–20 weeks of small-group intervention are provided. It is also important to note that the rate of inadequate response in efficacy studies, which control the quality of implementation, probably underestimates the actual percentage when intervention is practiced in schools. In actual practice, fidelity of implementation is likely to be lower, with reduced effects. In addition, as students continue in school, the effects of tutoring may diminish; without additional support, some responders will reemerge with math difficulties in other areas. Even so, if it were possible to provide a longer duration intervention or to deliver interventions individually, it may be possible to reduce the rate of inadequate response. Clearly, research to further intensify interventions and to examine effects when validated practices are delivered under typical school resources is required.

We have focused on classroom and supplemental instruction (Tier 2). In Chapter 5, we provided an example of tiered math intervention that showed that enhanced classroom and supplemental instruction was far more effective than either component alone for children with math LDs. But some students did not respond and in an MTSS implementation would proceed to a Tier 3 intervention. This should be in even smaller groups with more time and differentiation of instruction. However, the level of intensive intervention that focuses on students inadequately responsive to two tiers of intervention is not available and is needed. Powell and Fuchs (2015) outlined characteristic of a potential Tier 3 math intervention. They suggested that it needed to be highly individualized and provided strategies for intensifying beyond Tier 1. The need for small steps, student explanations, teacher modeling, manipulatives, worked examples, repeated practice and error correction, and fluency practices were emphasized. Specific examples of Tier 3 interventions were provided.

CONCLUSIONS: MATHEMATICS DISABILITIES

Geary’s (1993) seminal paper on math disabilities integrated findings from cognitive, developmental, intervention, and neurobiological studies and provided a framework for subsequent research in the field. In the past 10 years, there has been a considerable increase in activity across these areas of research on math and better integration of theory, core concepts, and measures across fields. In keeping with some of the themes of this book, considerable strides in our understanding of mathematics LDs have been achieved. This includes the confluence of intervention and correlational research examining children with and without other co-occurring reading and behavioral disorders. These studies, in combination with behavior genetic studies that look at the genetic and environmental sources of overlapping and unique variability in reading, math, and behavior, provide information that is important for understanding associations between various learning and behavior disorders. Such knowledge informs how we think about risk and may be important for thinking about interventions designed to address heterogeneity.

Another trend in the field has been the increasing differentiation of mathematical domains (e.g., calculation, word problems, fractions, and algebra). The number of high-quality intervention studies has increased in less-studied areas such as fractions and algebra, which are major determinants of postsecondary success and employment opportunities. The correlational research in this area shows that some of the cognitive correlates of different math skills overlap, but that important distinctions also exist between subdomains of mathematical cognition. Intervention research that looks at transfer between these domains of math provides a more complex and nuanced picture about what it means to have an LD in mathematics. Studies that investigate the genetic and environmental sources of variability in calculation and word problems also show some differences for these two types of math. This correlational, experimental, and genetic research demonstrates how the field is using knowledge from multiple sources to inform theory and practice. Importantly, this research gets to the heart of questions about definitions of LDs in math and is beginning to provide the needed empirical basis for specifying the critical academic skill deficits that represent markers for LDs in math. The key is to consistently use measures of both calculation and problem solving to define LDs and proceed from these designations to studies of number concepts, cognition, and neurobiology. Without these anchors, samples are so diverse that synthesis is difficult.

Of importance for early identification of risk is the study of potential early developmental markers of later LDs in math. To this end, there has been an upsurge in the number of behavioral and neurobiological studies and meta-analyses devoted to investigating the role of "number sense" as an important source of individual differences in math. The relation of math achievement to this evolutionary ability to discriminate large nonsymbolic quantities is not as large and direct as has been proposed. As well, heritability is considerably lower for this ability than it is for most other aspects of math. However, key prospective longitudinal studies are needed. Such studies may be important for understanding whether and how number-sense abilities very early in life and affect the development of both early and later mathematical skills.

In addition to research on Tier 3 intervention in mathematics, additional research on other components of the mathematics curriculum, at early and later stages of mathematics development, is required to provide greater clarity about whether strong core instruction creates protection against further risk and whether, even with intervention, we can expect new forms of risk to emerge. New risk may emerge due to lack of transfer from earlier intervention. Alternatively, new topics in the mathematics curriculum may create risk for students whose prior mathematics performance has been adequate. All this creates the need not only for additional intervention work, with a focus on long-term outcomes, but also for additional research on screening for risk on topics at the intermediate grades and at the middle and high school levels. In all areas involving LDs, there must be recognition of the principals in an MTSS framework: a focus on prevention, clear anchoring on intervention in strong core instruction, and a continuum of interventions based on progress monitoring and instructional response.

CHAPTER 9

Written Expression Disabilities

Disorders involving the writing process have been studied since Ogle (1867) used the term "agraphia" to distinguish an acquired writing disorder from aphasia, an acquired language disorder. This suggested the two disorders were dissociable. In the first half of the 20th century, Goldstein (1948) and others applied clinical observation and case study methodology to explore the association and dissociation between written and oral expression in brain-injured adults, but generally concluded that writing depended on speech and must therefore have similar neural correlates. As we discuss below, this hypothesis has not held up over time.

Jones, Abbott, and Berninger (2014) indicated that there is a substantial research base on written expression. Research on the writing process tends to be multidisciplinary and cross-cultural, with little cross-communication (Reich & Grigorenko, 2012), but it has been conducted for over a century. This research base does not always penetrate LDs involving written expression, which reflects disruptions of development that have effects on subsequently emerging skills.

Patterns of impairment are not as distinct in developmental as in acquired writing disorders, and they tend not to overlap (Romani, Olson, & Di Betta, 2005). LDs in written expression involve multiple domains, including handwriting, typing, and punctuation; spelling; and composition. In a meta-analysis of studies examining writing skills in children broadly identified with LDs, Graham, Collins, and Rigby-Wills (2017) identified 53 studies and 138 effect sizes. Comparisons with typically achieving children showed the largest differences in spelling, grammar, and handwriting (1.14), writing quality (1.06), organization (1.04), vocabulary quality in text (0.89), output (0.87), genre elements (0.82), and sentence-writing fluency (0.81). Thus, writing is a common problem in students identified with LDs, with multiple manifestations. These LDs are the least well studied in terms of definition, cognitive correlates, and neurobiological factors, which is ironic because there is a strong evidence base on effective interventions for these academic domains, including children with LDs.

As this meta-analysis demonstrates, a major issue is that many children with LDs also have some form of written expression difficulty, which extends to ADHD. Written expression requires many of the same cognitive processes that are involved in other LDs, epitomized by the need for phonological and orthographic mapping in spelling, problems that affect most with WLRD. For higher-level composition processes, self-regulation and monitoring, organizational, and oral language skills are required, which are common sources of difficulty for children with specific reading comprehension difficulties (Chapter 7) and children with ADHD (Denckla et al., 2013). The degree of overlap remains unclear, although there is better understanding of comorbidity since the first edition of the book. There is more research on handwriting and spelling, and the importance of these basic skills for higher-order composition is clearly established. This research is expanding to the neurobiological area, where there is emerging neuroimaging research, especially on spelling, and more genetics research. But the area of greatest development is intervention, where the accumulation of research permits identification of evidence-based approaches in all these components of written expression.

DEFINITION AND CLASSIFICATION

Most definitions of LDs include a category for written expression. In DSM-5, written expression can be specified as a form of LD and described as impairment of written expression with possible difficulties involving "spelling accuracy, grammar and punctuation accuracy, and/or clarity or organization of written expression." All other DSM-5 criteria for LDs apply, including persistence, age of onset, adequacy of instruction, and other exclusions. Other definitions, such as in the ICD-10 or IDEA, simply specify written expression as a category in which a person may have a LD. Again, similar criteria apply to all forms of LDs.

Components of Written Expression

In addition to the general issues with defining LDs (Chapter 4), two specific issues make definitions of LDs involving written expression difficult. The first is defining exactly what academic impairments constitute a LD in written expression. The example from DSM-5 exemplifies this problem and shows that many efforts to identify academic skill deficits that would permit definition address very specific skills or simply use a test with the term "written expression." As we discuss in the section below on academic skills deficits, our preferred approach, which was proposed by Berninger (2004), separates skills involved in the mechanical act of writing (transcription) from the processes involved in composing text (generation). But this distinction begs the question of what is represented within these broad categories of written expression (Kim, 2016). The problems are not unlike those involved in unpacking the construct of "listening comprehension" (Chapter 7), but are better understood than in mathematics LDs (Chapter 8).

Comorbidity

The second problem is comorbidity (Grigorenko, 2007). In developmental disorders of reading and writing, problems with spelling often accompany WLRD, but problems with handwriting and composition can occur in the absence of WLRD. Spelling and word reading are linked by problems involving phonological and orthographic processing (Berninger, 2004). Oral language skills are involved in composition, including vocabulary and morphosyntactic knowledge. There is substantial comorbidity of oral language and written expression disorders (Bishop & Clarkson, 2003). Not surprisingly, many children with poor reading comprehension have problems with composition (Carretti et al., 2014). Oral language, reading comprehension, and composition may be influenced by similar metacognitive processes, including planning, self-monitoring, self-evaluation, and self-modification.

Other childhood disorders are associated with disorders of written expression in the absence of reading and spelling problems, including nonverbal LDs (Rourke, 1989), oral language disorders (Bishop & Clarkson, 2003), and especially ADHD (Barkley, 2015). In ADHD, problems with writing are striking and originate out of the motor system, often called *dysgraphia* because of problems with voluntary motor control and planning that are expressed in the writing process. Difficulties with motor planning and execution create even more difficulty engaging top-down cognitive control processes (self-regulation, organization, monitoring) involved in composing, which are often problems in children with ADHD who do not have dysgraphia and overlap with processes involved in reading comprehension (Denckla et al., 2013). Mayes, Calhoun, and Crowell (2000) found in a clinic sample that 70% of children identified with ADHD were identified with LDs. These were predominantly in written expression, which were twice as common (65%) as reading or math LDs. In a follow-up of this sample, Yoshimasu et al. (2011) reported that the rate of written language LDs was over five times higher for children identified with ADHD than those not identified with ADHD.

In a meta-analysis, Graham, Fishman, Reid, and Hebert (2016) synthesized studies comparing the writing characteristics of children with ADHD. They found 45 studies and 87 effect sizes, with the largest differences relative to typically achieving controls in spelling (0.80), writing quality (0.78), vocabulary (0.76), number of genre elements (0.69), output (0.64), and handwriting (0.62). Unfortunately, it was not possible to address the comorbidity of reading problems, which may explain the associations with spelling and vocabulary. Because executive function and attention deficits are generally more severe in people with comorbid WLRD and ADHD (see Chapter 2), we hypothesize that writing difficulties in those with ADHD and no WLRD would be more apparent on transcription measures.

In this vein, studies of written expression have not followed the lead of research on reading and math disabilities in distinguishing between children with specific writing disabilities versus comorbidity with other LDs and

ADHD. This issue has hampered definitional efforts, so that the classification of disorders of written expression has lagged behind that of reading and math LDs. There are still no clear operational definitions that address all components of the written language domain (Berninger & Chanquoy, 2012). Although the view that deficits in written expression invariably co-occur with WLRD has not held up, emergent research on written language indicates that many children with LDs have problems with at least one academic skill in writing, whether it is handwriting, spelling, or written discourse (Hooper et al., 1994). In an epidemiological study, Katusic, Colligan, Weaver, and Barbaresi (2009) found that 25% of those with written language LDs did not have a problem with reading, but this estimate does not take into account comorbidity with math LDs or ADHD.

It is not clear whether written language disorders are simple expressions of common underlying processes, as in the relations of phonological processing, word reading, and spelling in children with WLRD, or represent additional, independent disorders. In particular, is there a prototype for an isolated written expression disorder or for academic skill deficits involving handwriting (likely), spelling (infrequently, but see Wimmer & Mayringer, 2002, who observe reading problems in the absence of spelling difficulties in German children; Chapter 10), and composition (not known)? In adults, writing difficulties often reflect an inability to spell that, even when remediated, is closely associated with difficulties in word recognition (Rourke, 1993). When children with LD in math have difficulty with handwriting, it is because of impairments in their motor development. Their spelling errors, interestingly, are typically phonetically constrained, in contrast to children who have word recognition difficulties (Rourke, 1993). Once these two difficulties (spelling and motor skills) are taken into account, is there a subgroup of children whose difficulties are restricted to composing? The classification research that is necessary to evaluate this hypothesis has not been completed.

Prevalence

These definitional issues make studies of the epidemiology of LDs involving written expression difficult. In a birth cohort of over 5,000 children born from 1976 to 1982 in Mayo County, Minnesota, Katusic et al. (2009) found prevalence rates of 6.9–14.7% for written expression difficulties based on a search of school records, with the variability depending on the formulae, which were for low achievement (25th percentile) as well as discrepancy definitions. There were no clearly defining characteristics of these LDs. In earlier studies, Berninger and Hart (1992) found incidence rates of 1.3–2.7% for handwriting, about 4% for spelling, and 1–3% for written expression in a sample of 300 elementary school children. Hooper et al. (1993) evaluated the prevalence of composition problems in an epidemiological sample of 1,274 middle schoolers, finding rates of 6–22% with scores one standard deviation below average (about the 15th percentile) on the narrative subtest of the Test of Written Language; the variability reflected different sociodemographic factors, with higher rates in boys and minorities. Given the rates of developmental language disorders and WLRD in the general population, one could predict that written language disorders affect at least 10% of the school-age population, depending, as always, on the criteria used to define the LDs and comorbidity.

Sex Ratio

In Katusic et al. (2009), boys were two to three times more likely than girls to have LDs in written expression, depending on the definition. But this study relied on ascertainment through identification by a school or clinic, which always yields higher rates for boys than girls. Berninger and Fuller (1992) and Hooper et al. (1993) reported that more boys than girls (about 1.5:1) displayed written language deficits when level of achievement was used as the comparison variable. In contrast, Berninger and Hart (1992) found no differences in sex ratio when IQ–achievement discrepancy criteria were used. Berninger, Nielsen, Abbott, Wijsman, and Raskind (2008) evaluated reading and writing skills in a sample of children and adults identified because of family risk for dyslexia. The incidence of writing problems showed male predominance rates of about 1.5:1, and boys and men were more severely affected. The largest differences were apparent in handwriting and spelling. If men had writing impairments, spelling was more affected. The differences were not due to motor impairments, but reflected greater difficulty with orthographic processing. Kim, Park, and Park (2015) also found more severe writing problems in boys than girls in grades 2 and 3. Clearly, both the amount and accuracy of epidemiological data are lacking, particularly in comparison to studies of oral language and reading. In a twin study (Olson et al., 2013), girls outperformed boys on a writing composite. This finding is interesting because in other studies of this Colorado sample, sex differences in incidence or severity of reading and spelling problems have not been apparent.

Developmental Course and Outcomes

There are few studies of long-term outcomes in children identified specifically with disorders of written expression. Many studies focus on children identified with language and reading problems, often reporting that writing problems are persistent (Bruck, 1987). Hamstra-Bletz and Blöte (1993) found that Dutch students with dysgraphic handwriting (n = 121), who were tracked from grades 2 to 6, had lower fine-motor ability in early grades and less preference for personal style in later grades, but did not differ in writing speed from children without dysgraphic handwriting. Connelly, Campbell, MacLean, and Barnes (2006) found that college students identified with dyslexia had writing difficulties because of problems with handwriting speed and spelling in context. Addressing written expression in an unselected sample, Berninger et al. (2006a) found that individual differences in writing ability were stable through grades 1–5. Costa, Edwards, and Hooper (2016) evaluated 137 grade 1 children identified as typically developing (n = 83), writing-disabled (n = 38), and writing and reading disabled (n = 16). The rate of isolated writing disabilities was twice as high in grade 1 as writing with reading disabilities; by grade 4, the rate had increased from about 30 to 47%. It has long been known that oral language disorders are associated with significant long-term problems with written expression even when the oral language problems seem to have resolved or significantly improved. This likely reflected the later acquisition of written language skills (Bishop & Snowling, 2004). Bishop and Clarkson (2003) followed a large sample of twins in which one or both twins had an oral language disorder. Most of the twins could not spell well enough to attempt narrative production. Even twins of affected probands who had no evidence of oral language impairment on standardized tests showed more difficulties on written language narratives than age-matched controls. It is likely that written language problems are persistent across different populations and certainly in children defined with LDs in reading and oral language.

ACADEMIC SKILL DEFICITS

Transcription and Composition

Despite difficulties with definition and the question of prototypes with isolated difficulties, progress has been made in understanding the academic skill deficits associated with written expression. Writing difficulties can involve problems with handwriting, spelling, and/or composition—the expression of ideas at the level of text. As already noted, Berninger (2004) differentiated the "transcription" component of writing from its "generational" component. The transcription component involves the production of letters and spelling, which are necessary to translate ideas into a written product. The generational component involves translation of ideas into language representations that must be organized, stored, and then retrieved from memory. Transcription represents basic mechanical processes, spelling, and punctuation, whereas generation represents composition. Struggling writers focus more on the product than the process of writing, compared to typically developing and novice writers (Lin, Monroe, & Troia, 2007).

In research, there has been greater focus on transcription than generation. This progress reflects in part the fact that transcription is specific to the writing process, whereas the generational component is applicable to many aspects of language and thought. Nonetheless, the transcription and generational components are closely linked. Just as word recognition problems constrain reading comprehension, problems with handwriting and spelling constrain composing because of problems with automaticity. When children first begin to write, their capacity to generate text is constrained by their need to develop skills for writing alphabetic and logographic symbols that represent speech. Their capacity for telling stories far exceeds their ability to write what they think, consistent with a capacity model of writing (Berninger, 2004). As children become proficient with transcription, they can write more rapidly to a point where handwriting speed is less of a constraint on their ability to compose. This is a major source of individual differences in the ability to compose essays and stories in children and adults (Medwell & Wray, 2014) and even in note taking (Peverly, Garner, & Vekaria, 2014). Interestingly, handwriting fluency is an effective predictor of composition, note taking, and other written language tasks in adults (Peverly, 2006).

In a broad meta-analysis of the relation of component skills and the writing process, Kent and Wanzek (2016) included an examination of handwriting fluency and spelling to quality of writing and production. Handwriting fluency involved 17 effect sizes with writing quality and seven with production. For both domains, the average effect size was moderate (about 0.50). Spelling had similar moderate relations with writing quality, but weak relations with productivity (0.25). These relations did not vary significantly with grade level.

There are methods for assessing handwriting, spelling, and composition, although specific tests are not as well developed as many experts would hope. For handwriting, qualitative assessments of legibility of the writing sample are often employed. Spelling tests that involve the dictation of single words are common, but frequent methodological limitations involve how items are organized in terms of different orthographic conventions as well as the number of items at different levels of ability. It is also possible to score spelling errors in context. Composition is usually evaluated through coding systems that require judgments about specific components of the written narrative. The Test of Written Language (Hamill & Larsen, 2009), which uses a spontaneous writing sample in response to complex pictures, represents a formal, published test; many approaches that involve the scoring of narratives are used in research, often based on scoring rubrics for narrative discourse. Curriculum-based assessments in relation to transcription usually involve writing letters of words as quickly as possible for transcription; for generation, curriculum-based measurement involves assessments in response to prompts, which yield scores for production and quality. These types of measures have been shown to be sensitive to growth with adequate reliability and validity in elementary school children (Ritchey & Coker, 2013).

Relation of Transcription and Composition

Since the seminal work by Flower and Hayes (1980), which began efforts to view writing as composed of a set of cognitive and linguistic processes, researchers have tried to unpack the components of transcription and composition. Learning to write in order to compose has a protracted developmental course. Berninger, Nagy, and Beers (2011) asked students in grades 1–4 to complete two tasks involving sentence construction. One task evaluated sentence integrity, asking the children to write one complete sentence in response to a prompt. The second task asked them to integrate two sentences into one and maintain the meaning of the sentence. Most children could only write one complete sentence, with no relation of the sentence integrity and combining tasks until grade 4. In the earlier grades, morphological knowledge of prefixes and word endings (grade 1) and spelling (grades 2–4) accounted for how well children combined sentences. The authors concluded that students possess morphosyntactic knowledge, but cannot use it for composing sentences until grade 4 due to the constraints of transcription.

In beginning writers, Puranik and Al Otaiba (2012) gave 242 kindergarten children measures of reading, writing, cognitive, and oral language skills. Handwriting and spelling were related to the ability to express ideas in writing, but there was no relation of oral language and reading skills. In the same sample, structural equation modeling also showed no relation with early reading and unique relations of spelling and letter writing fluency, with weaker relations of oral language (Kim et al., 2011).

The relation of transcription and generation has also been expressed through frameworks similar to the SVR (Gough & Tunmer, 1986; Chapter 6). Juel, Griffith, and Gough (1986) proposed a simple view of writing in which written expression was the product of spelling and ideation. Berninger and Winn (2006) expanded the generation component into a “not-so-simple” view of written expression, adding planning, reviewing, and revising, which required WM, and cognitive control of attention (Figure 9.1). Component models like the simple and not-so-simple views of writing have been evaluated in large studies that use multiple measures of transcription and composition to reliably measure the underlying (latent) constructs.

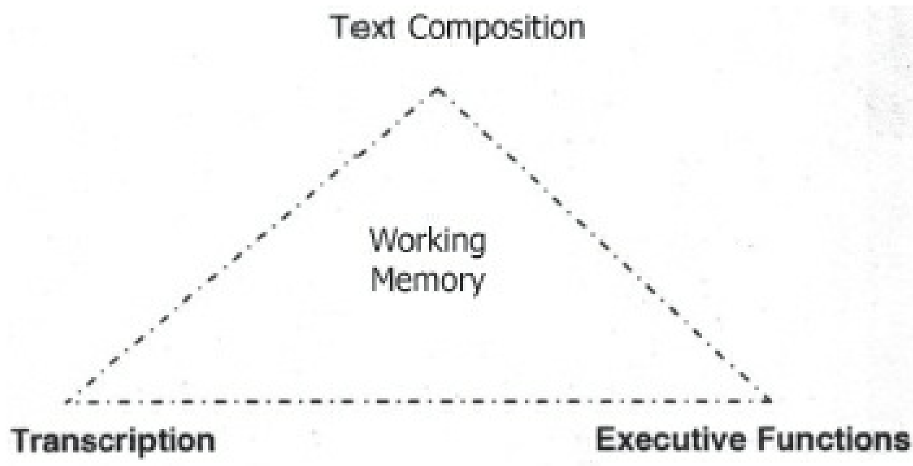

FIGURE 9.1. The not-so-simple view of written expression. Courtesy Whitney Roper.

In a series of studies of elementary and middle school children, Berninger and colleagues (Berninger, 2004; Graham, Berninger, Abbott, Abbott, & Whitaker, 1997) found that a test involving printing of the lower-case letters of the alphabet as fast as possible for 15 seconds predicts a variety of written expression outcomes. Graham et al. (1997) conducted a structural modeling study of different measures of handwriting fluency, spelling, and composition in a sample of 600 children in grades 1–6. Handwriting fluency predicted compositional fluency and quality in primary and intermediate grades; handwriting fluency and spelling predicted compositional fluency in the primary grades. Across the age range, these latent variables accounted for 41–66% of the variance in compositional fluency and 25–42% of the variance in compositional quality. The researchers concluded that the transcription component of writing constrains the amount and quality of the composition.

In a longitudinal study of children in grades 1–7, Abbott, Berninger, and Fayol (2010) examined relations according to word-, sentence-, and text-level processing. In grades 1–7, there were significant relations of spelling to spelling and spelling to composition. Wagner et al. (2011) examined the underlying correlates of handwriting fluency and written composition in grades 1 and 4 based on an analysis of writing samples. For both age groups, they identified five constructs. Three involved written composition: overall text organization, productivity, and complexity; the other two were spelling and handwriting fluency. Productivity and handwriting fluency accounted for differences in grades 1 and 4, with smaller contributions of complexity and organization; spelling and punctuation were weakly related to developmental differences. In a study that included evaluation of the dimensionality of writing in 494 children in grades 2 and 3, three dimensions emerged: writing quality, productivity (number of words and ideas), and correct and incorrect word sequence. The latter two measures were highly correlated. As such, many studies focus on the production of correct text as opposed to overall productivity, especially in older students.

Kim and Schatschneider (2017) tested a model that expanded the simple and not-so-simple views of writing, which they termed the "direct and indirect effects model of writing" (DIEW). This model explicitly attempts to operationalize the compositional component of the model into separate domains: construction of situation models and discourse-level language found in models of reading comprehension (see Chapter 7). It also includes processes like inference making and WM as components. WM is seen as a foundational resource that influences components of transcription (spelling, sentence writing) as well as discourse-level oral language, the latter involving grammar, inferencing, theory of mind (perspective taking), and vocabulary, an idea that builds on Kellogg's (1996) model of WM in writing and other models (see Olive, 2014). Kim et al. tested four variations of the model in 193 grade 1 children. They found that handwriting fluency, spelling, and discourse-level oral language were directly related to higher-order cognitive skills, foundational language, and WM as paths to written expression. There were also indirect effects of language and cognitive skills to discourse-level oral language, paralleling findings in the reading comprehension area (Chapter 7). Writing outcomes were most strongly related to discourse-level skills (path coefficient = .46), WM (.43), and spelling (.37), with smaller contributions by vocabulary (.19), handwriting (.17), theory of mind (.12), inferencing (.10), and grammar (.10).

In cross-linguistic studies, Limpo and Alves (2013) modeled the relation of transcription, revision, planning, and the student's perception of self-efficacy to the quality of writing in elementary and middle school in Portuguese children. In the younger children, the model including these variables accounted for 76% of the variability in written expression. There was a strong, direct influence of transcription n writing quality. In older children, the five factors explained similar amounts of variability, but handwriting and spelling were not directly related to writing quality. Transcription did contribute indirectly through planning and self-efficacy. Thus, in a more transparent orthography, transcription still constrained writing quality in the younger age group, but greater automaticity with development reduced these constraints in the adolescent group. In a study looking specifically at spelling in Italian beginning writers, Arfé, Dockrell, and De Bernadi (2016) found that spelling skills contributed to text accuracy and quality, but less than oral grammar skills. They concluded that spelling skills were important, but perhaps less so for a transparent orthography.

These studies clearly demonstrate that across different languages, transcription and generation can be differentiated and are closely related. They also support a capacity model in which automaticity of transcription skills is essential to composition. As we see in the next section, the components of written expression vary depending on the academic skill deficits used to define the writing problem.

CORE COGNITIVE PROCESSES

Handwriting

Berninger and her associates (Berninger, 1994, 2004) and Graham and colleagues (Graham, Harris, & Fink, 2000; Graham, Weintraub, & Berninger, 2001) reported that automaticity in the retrieval and production of alphabet letters, rapid coding of orthographic information, and speed of sequential finger movements were the best predictors of handwriting skills. Automaticity of handwriting predicted compositional fluency and quality. A deficit in fine motor skills also constrained handwriting, especially in the beginning stages of writing, which may be why sequential motor movement is related to letter production and legibility (Berninger, 2004). Peverly et al. (2014) found that handwriting speed and selective attention were more related to note taking in adults than WM, language comprehension, and speed of lexical access; handwriting speed itself was a product of fine motor fluency and speed of lexical access. In more transparent languages like Turkish, handwriting speed is also the best predictor of writing fluency (Babayiĝit & Stainthorp, 2010). Handwriting is more than a motor act, so that knowledge of orthography and planning ability also contribute to handwriting (and spelling) proficiency.

Spelling

Linguistic skills involving both phonological and orthographic processes seem important at even the earliest development of spelling abilities. Apel, Wolter, and Masterson (2006) examined the impact of phonological and orthographic processes on learning to spell. They found that young children quickly mapped orthographic information on letter patterns with minimal exposure to novel words. Letter patterns in the novel words that occurred more frequently were learned more easily, just as phonological information that occurred more frequently was mapped more rapidly. In concluding that both phonological and orthographic processes were important for spelling, Apel et al. (2006) countered other explanations for the relation of these processes in spelling, which suggested that orthographic representations were simply mapped onto phonological representations (Treiman & Kessler, 2005).

Spelling abilities are also predicted by language skills involving phonological and orthographic mappings and motor skills, especially visual–motor integration (Berninger, 2004). Because writing involves a mechanical act, it is not surprising that assessments of the motor system predict spelling abilities. There is controversy about the degree to which phonological and orthographic processes are independent and whether orthographic processing can be reliably measured as a separate process (Vellutino, Scanlon, & Tanzman, 1994). Romani et al. (2005) argued that spelling development reflects two processes, one involving phonological processing at a sublexical level and the other representing a problem with storing adequate orthographic relations as a lexical pattern that leads to significant difficulties in the accurate spelling of irregular words. Romani et al. (2005) questioned whether the lexical problem was due to problems with visual processing or difficulties in creating lexical representations, which reflect problems at a phonological and/or orthographic level growing out of the language system.

The need for phonological representations of words for spelling is obvious. Even in English, the phonological system is more predictive of word spellings than is commonly understood, especially if the historical origins of words are considered (Moats, 2005). About 50% of English spellings are regular, and many others are only slightly irregular (Joshi et al., 2008–2009). In elementary school children, knowledge of the morphemic structure of words is clearly important. Arndt and Foorman (2010) reported that in second graders, morphological errors were the most common error type, followed by orthographic errors. A similar pattern was observed in the bottom 25% of spellers on their dictated, researcher-developed measure. Bourassa and Treiman (2008) also found that children with dyslexia also made more frequent morphological errors, with a pattern resembling that of younger, typically developing children.

These error patterns are also apparent in more transparent languages, such as Greek, where the differences in errors is quantitative, not qualitative, and morphological errors are most common (Protopapas, Fakou, Drakopoulou, Skaloumbakas, & Mouzaki, 2013). At the same time, despite these error types, phonemic awareness is a robust longitudinal predictor of spelling (and reading; Nikolopoulous, Goulandris, Hulme, & Snowling, 2006). In Turkish, also a more transparent language, Babayiĝit and Stainthorp (2010) found that phonological and morphological errors were related to spelling, but level of spelling ability was the most significant factor.

Writing in logographic languages like Chinese highlights the importance of orthographic and syntactic processes in spelling. Tan, Spinks, Eden, Perfetti, and Siok (2005) showed that learning to read in Chinese was strongly related to the ability to write in Chinese. The relation of phonological awareness to reading and writing was weaker in Chinese than in an alphabetic language. Tan et al. also found that writing Chinese characters depends on two interacting mechanisms: one involving orthographic awareness that links visual, phonological, and semantic systems, the other involving motor programs that allow for the storage and retention of the characters. Writing (and reading) in Chinese children with dyslexia is predicted by naming speed, orthographic knowledge, and phonological memory (Chan, Ho, Tsang, Lee, & Chung, 2006). Packard et al. (2006) found that teaching about the morphological and orthographic structure of Chinese words increased the ability to write Chinese characters. Across languages, those with more transparent relations of phonology and orthography seem to produce less severe difficulties with word-reading accuracy, but the spelling and fluency problems are more marked, which suggests that the phonological, morphological, and orthographic components of spelling (and word reading) are dissociable (Caravolas et al., 2012; Protopapas, Simos, Sideridis, & Mouzaki, 2012).

Composition

Composition is the capacity to generate ideas in writing. In addition to transcription, it requires the formulation of ideas, organizational skills, planning, and specific writing processes, such as editing and revising. Johnson and Myklebust (1967) presented a developmental model of language learning, which posited that the ability to write is dependent on adequate development in listening, speaking, and reading. This highlights the link between different language skills and composition. In their meta-analysis Kent and Wanzek (2016) found a small-to-moderate relation of oral language skills and writing quality (ES = 0.33). However, Kim and Schatschneider (2017), a study with stronger operationalization of discourse-level language skills, found much stronger relations.

Another domain that has received some attention is executive functions, a construct that is not simple to define. In the not-so-simple view of writing, Berninger et al. (2006a) highlight the importance of supervisory attentional and other top-down cognitive control processes for revising, editing, and composing. Hooper, Swartz, Wakely, de Kruif, and Montgomery (2002) documented a role of executive functions in disorders of written expression. Controlling for level of decoding ability, comparisons of good and poor writers (identified on the basis of evaluations of narrative text) showed that poor writers had particular difficulties on measures involving initiating responses and shifting response set. De La Paz, Swanson, and Graham (1998) found that the difficulties experienced in revising written text by older (eighth-grade) students with writing problems were due in part to executive control issues. However, mechanical difficulties also contributed to these problems with revision.

In examining different handwriting modes in third- and fifth-grade children, Berninger et al. (2006b) found that inhibition and set switching were effective predictors. Obviously, the transcription and generation components must interact for an individual to produce high-quality written text. A role for executive functions in terms of planning and organizing written expression at the level of handwriting and composition is apparent and has had significant influence on the development of interventions in the written expression area (see below).

Attempting to link the language and executive function domains, Hooper et al. (1994) conceptualized writing as a complex problem-solving process reflecting the writer's declarative knowledge, procedural knowledge, and conditional knowledge, all of which are subserved by a network of neuropsychological factors, personality factors, and other conditions (including teacher–student relationships, amount of writing instruction, and the teacher's knowledge of the writing process). In this context, "declarative knowledge" refers to the specific writing and spelling subskills that the learner has acquired, whereas "procedural knowledge" refers to the learner's competence in using such knowledge while writing for meaning. Similarly, Berninger (2004) suggested that neuropsychological, linguistic, and higher cognitive constraints may be recursive throughout the development of the writing process, but that each of these constraints may exert more influence at different points in the developmental process (Hooper et al., 1994).

WM is an executive function task that has been pivotal for conceptions of the development of composition skills from the earliest process-oriented models (Hayes & Flowers, 1980). Although most researchers agree that WM is involved in writing, especially composition, it is not clear which components of written expression are most related to WM (Hayes & Chenoweth, 2007). Berninger et al. (2008) examined word- and sentence-level reading and writing in relation to WM grades 2, 4, and 6. Word-level WM contributed unique variance to reading and writing outcomes except for grade 6 writing; text-level WM was most strongly related to reading comprehension in grades 4 and 6. Kim and Schatschneider (2017) hypothesized that WM was a fundamental skill for language and cognitive processes involved in reading, as well as in transcription. These skills would represent components of writing. Higher-level language skills were expected to relate to text generation. In grade 1 students, they found that listening comprehension, vocabulary and grammar, and WM fully mediated writing skills. The strongest relations were for listening comprehension and WM, with weaker contributions of vocabulary and handwriting. In a meta-analysis of component skills of reading and writing, Ahmed, Wagner, and Lopez (2014) did not find strong relations of WM to writing quality when oral language skills were in the model. Altogether, the roles of executive functions and self-regulation for text generation are established, but the precise role of WM remains unclear, possibly because WM is involved in most executive function and higher-level comprehension tasks.

Relations of Reading and Writing

The simple and not-so-simple views of writing introduce possible dimensionality and relations of reading and writing and their reciprocity in development. Figure 9.2 highlights some of these relations and the different models that can be tested. Kent and Wanzek (2016) reported a meta-analytic correlation of reading and writing quality of .48. Mehta, Foorman, Branum-Martin, and Taylor (2005) examined the dimensionality of language, literacy, and writing in a large sample of children in grades 1–4. They reported that there was a unitary literacy factor for reading and spelling, with phonological awareness the best indicator of both domains, but declining in importance from grades 1 to 4. Writing was only weakly related to literacy at the individual level, but highly impacted by individual differences in quality of teaching. Language was dissociable from literacy and language at the student level, but perfectly correlated with the level of literacy at the classroom level. These findings show that these domains are dissociable at the student level, but are impacted by classroom-level differences. Berninger and Abbot (2010) evaluated oral expression, listening comprehension, reading compression, and written expression in students ranging from grade 1 to grade 7, finding these domains were dissociable, but with shared components. In further exploration of these data, Abbott et al. (2010) reported paths from composition to comprehension in grades 3–5 and comprehension to composition in grades 2–6. Comprehension influenced word reading in grades 1–6.

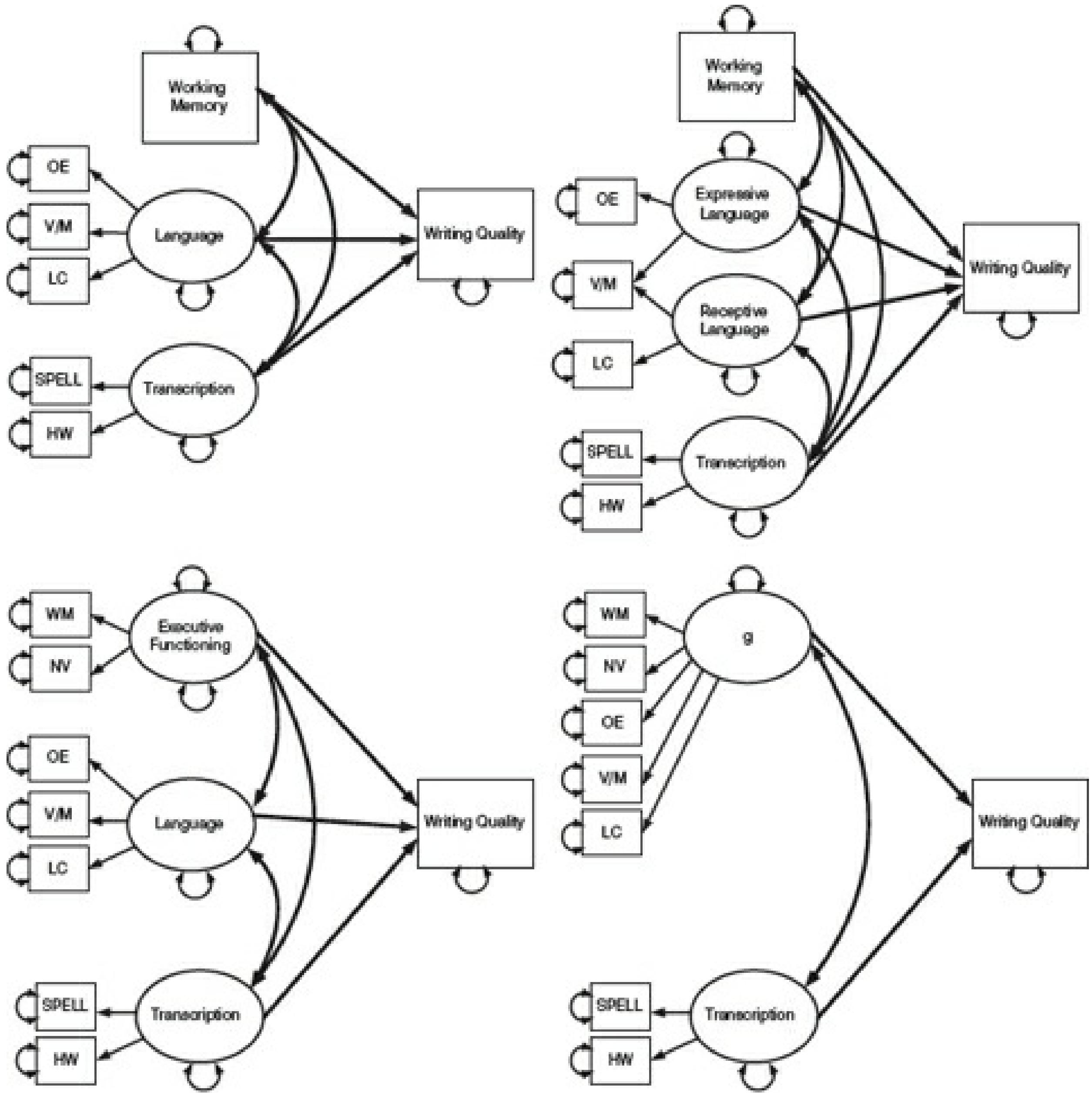

FIGURE 9.2. Different component models for assessing the relation of reading and writing. HW, handwriting; LC, listening comprehension; NV, nonverbal reasoning: OE, oral expression; SPELL, spelling; V/M, vocabulary/morphology; WM, working memory. Courtesy Yusra Ahmad.

Berninger et al. (2006a) compared the interrelations of listening comprehension, oral expression, reading comprehension, and written expression in a large sample of elementary school children in grades 1, 3, and 5. They found only moderate correlations among the four domains and also found that different neuropsychological tasks differentially predicted each domain, again suggesting that they are dissociable, but also supporting a bidirectional model. Shanahan and Lomax (1986, 1988) compared unidirectional and bidirectional models in second-grade students. They found that the bidirectional model was the best fit, followed by the reading-to-writing model and the writing-to-reading model. Reading and writing shared about 50% overlapping variance. Ahmed et al. (2014) evaluated relations of reading and writing at word, sentence, and passage levels in a longitudinal study from grade 1 to grade 4. Using latent change models to assess growth and reciprocity over time, they found that a unidirectional reading-to-writing model best fit the data at word and text levels, but that a bidirectional model was the best fit at the level of the sentence. Ahmed et al. (2014) found that oral language skills were related to reading, but not writing; transcription was predictive of writing quality, as was reading comprehension. These results generally support a model where reading has more influence on writing than writing has on reading, with some evidence of reciprocity. These findings tend to show weakest effects for writing-to-reading models. By contrast, Graham and Hebert (2011) conducted a meta-analysis examining the effects of writing instruction on reading. For students in grades 2–12, they reported an average effect size of 0.37 on norm-referenced tests of reading comprehension, in which students were required to write about material they had read. For struggling students, the effect size was moderate (0.64). Increasing the amount of writing had an effect size of 0.35. These effects, in the small-to-moderate range, also indicate that writing has an effect on reading, although across all the available studies, the effects of reading on writing seem stronger.

Summary: Cognitive Processes in Written Expression

Different academic skill deficits in written expression have different cognitive correlates: Handwriting is correlated with fine-motor, motor-planning, and WM skills; spelling with phonological analysis, knowledge of orthographic conventions specific to a child’s language of instruction, and visual–motor skills; and composition with executive functions and a variety of oral language skills (Berninger, 2004; Hooper, Wakely, de Kruif, & Swartz, 2006; Kim & Schatschneider, 2017). Motor and phonological/orthographic difficulties correlated with handwriting and spelling especially constrain the transcription component, whereas difficulties with executive functions and discourse-level language constrain the generational component. The motor and executive function problems help explain why so many children with ADHD have disorders of written expression (Barkley, 2015). Subtypes may exist, but the key classification question is whether written language difficulties occur in the absence of other LDs, oral language problems, or ADHD.

NEUROBIOLOGICAL FACTORS

There are fMRI studies of writing, spelling, and composition, some of which include analyses of brain volumes. These studies generally involve both adults and children and have small samples, with little consistency across studies. There are few anatomical studies of children with isolated LDs in written expression.

Brain Function

Writing

A meta-analysis (Planton, Jucla, Roux, & Démonet, 2013) of 18 handwriting fMRI studies identified 12 cortical and subcortical regions. Writing-specific regions included the left frontal superior and middle frontal gyrus, the left intraparietal and superior parietal areas, and the right cerebellum. There was also activation of areas involved in motor functions (e.g., primary motor and sensorimotor cortex, supplementary motor cortex), and language functions (temporal lobe regions). In a follow-up, Planton, Longcamp, Péran, Démonet, and Jucla (2017) attempted to define the role of five areas identified in the meta-analysis as central to writing. All five areas were activated by writing and drawing, suggesting lack of specificity for writing. Only the superior premotor cortex showed left lateralization during writing. Other regions included the superior parietal cortex, ventral premotor cortex, ventral occipitotemporal cortex, and right cerebellum. Although these regions seem to involve a network activated during writing for transcription and for processing of orthographic forms, they do not appear to be specific to the writing process.

In children, who lack the automaticity of adults, Berninger (2004) summarized findings from functional neuroimaging studies that addressed components of the writing process conducted by the University of Washington group. She reported that components involved in fine motor control can be related to areas of the frontal lobes and the cerebellum. These areas are well known to be involved in support of core processes that underlie writing, including motor control and planning. In subsequent studies, Richards et al. (2009a) compared good and poor fifth-grade writers on tasks involving sequential finger movements. They found that poorer writing was associated with underactivation of a broad range of areas across the brain and cerebellum, suggesting that finger movements involved a widely distributed set of brain regions.

In a study of five- and six-year-olds, Gimenez et al. (2014) found that stronger early handwriting was associated with reduced activation and increased volume of the pars triangulus of the right inferior frontal gyrus, which they related to increased neural efficiency of this region. Since the task used to activate the brain was a phonological processing task, activation of the inferior frontal gyrus is not surprising. James and Engelhardt (2012) compared typing and handwriting in prereading 5-year-olds. These children printed, traced, or typed letters and were then scanned. When shown what they had drawn, only the handwritten circuits activated the neural network associated with reading, implying that handwriting was an important contributor to early reading and not replaced by other motor activities, such as tracing shapes or typing.

These studies show a pattern of activations involving frontal and parietal regions as well as the cerebellum during writing. There does appear to be specificity of the network to writing. With the exception of studies by the Berninger and Richards's group, few evaluate children who are poor writers.

Spelling

Several studies have compared good and poor spellers. Richards, Berninger, and Fayol (2009b) compared small groups of good and poor 11-year-old spellers. When comparing judgments about letters in unfamiliar orthographic representations, poor spellers showed less activation in the left posterior cingulate and precuneous. If the task involved item pairs pronounced the same but spelled differently and correctly, poor spellers had reduced activation in the left precentral, postcentral, and inferior frontal gyri, but more activation of motor areas and other frontal lobe regions. Gebauer et al. (2012a) evaluated 31 good and poor spellers, including typically developing controls, poor spellers with no reading impairments, and poor spellers with reading impairment. In making judgments about the accuracy of orthographic representations, poor spellers had greater right hemisphere activation than the other two groups. They also activated bilateral and inferior frontal gyri when processing correct and incorrect spellings of words, with the other two groups showing bilateral differences only in the misspelled condition, so that the additional right hemisphere activation might indicate that the tasks demanded more effort. There were no differences in the integrity of frontal white matter between poor spellers and controls, but those poor in reading and spelling had reduced integrity of frontal white matter.

In intervention-imaging studies, Richards et al. (2005, 2006) evaluated brain activation in response to two different spelling interventions in children in grades 4–6 identified with dyslexia. One intervention involved an orthographically based intervention that taught specific strategies for letter patterns. The other intervention focused on morphological components of spelling, teaching children to synthesize word parts to make words and to break down words into constituent elements that supported the meaning of the words. These interventions were conducted in 14-hour-long sessions over 3 weeks with before and after fMRI based on four word-reading tasks that manipulated phonological mapping, orthographic mapping, and morphological mapping with and without phonological shifts. The investigators found unique patterns of activation for each of the four tasks at baseline in controls, with common activation across tasks of structures often associated with reading: the left inferior frontal gyrus, bilateral lingual gyrus, bilateral fusiform gyrus, and left inferior temporal gyrus. A variety of cortical and cerebellum structures were uniquely activated. The patterns were different in the children with dyslexia, always involving underactivation, and were most apparent on tasks requiring phonological mapping. After intervention involving orthographic mapping, the right inferior frontal gyrus and the right posterior parietal gyrus showed significantly greater activation in the group with dyslexia, with little change in the control group. Morphological treatment did not lead to significant changes in activation. These changes were considered normalizing. Gebauer et al. (2012b) investigated changes in the integrity of the of white matter tracts using DTI in poor spellers who underwent a training program focused on understanding and use of morphemes. At baseline, DTI revealed reduced integrity of right hemisphere tracts involving the superior corona radiate, posterior internal capsule, and superior longitudinal fasciculus. After the 5-week training, there was increased integrity of these right hemisphere regions, especially the corona radiata.

Unfortunately, these studies are generally conducted with small samples, and there is no consistent pattern in results except for differences in the activation and integrity of right hemisphere regions. This may be related to effort and task demands.

Composition

In the only study we identified of text generation in good and poor writers, Berninger, Abbott, Augsburger, and Garcia (2009a) asked 10-year-old children to generate ideas on a topic they would be asked to write about after scanning. Good writers showed more activation than poor writers in frontal and temporal regions associated with language and executive function. Poor writers showed more activation in the right prefrontal regions, which the authors attributed to WM demands.

Genetic Factors

There are a few studies of heritability of spelling and writing abilities. Raskind, Hsu, Berninger, Thomson, and Wijsman (2000) found that spelling disorders, but not handwriting disorders, aggregate in families. Other studies have also found that spelling difficulties aggregate in families (Schulte-Körne, Deimel, Müller, Gutenbrunner, & Remschmidt, 1996). These findings are consistent with twin studies, which have found strong heritability of spelling abilities, which exceed that found for reading abilities (Stevenson, Graham, Fredman, & McLoughlin, 1987). Bates et al. (2004) evaluated genetic and environmental influences on reading and spelling of real words, pseudowords, and irregular words. They reported heritabilities of .61 for real words, .71 for pseudowords, and .73 for irregular words; spelling yielded estimates of .76 for real and irregular words, and .52 for pseudowords. Evaluations of the environmental contributions were significant, representing variance due to unique environmental influences and not differences in families. In their study of adult twins reared apart, Johnson, Bouchard, Segal, and Samuels (2005) found heritabilities around .75 for different measures of word reading, .51 for reading comprehension, and .76 for spelling.

In the only study with a broad assessment of writing, Olson et al. (2013) examined 540 sets of identical and fraternal twins. The assessments included a measure of composition at the sentence level, writing fluency, and a handwriting legibility test. Measures of reading and component processes were also used. Olson et al. found that sentence composition (.72) and handwriting (.79) had substantial genetic influences, with fluency at .36. These estimates were comparable to the heritability coefficients for rapid naming and phonological awareness (.79), rapid naming (.70), word reading (.81), spelling (.87), and comprehension (.83). Only a vocabulary measure showed a shared environmental influence; unique (nonshared) environmental influences were significant for all measures. For the three writing measures, there was significant shared genetic influence, especially for composition and fluency, indicating common genetic influence, but also unique genetic influences.

In linkage studies, Schulte-Körne (2001) found evidence linking spelling to a region of chromosome 15. Similarly, Nöthen et al. (1999) reported a locus for spelling (and reading) on chromosome 15, which has also been reported for dyslexia (Grigorenko, 2005). Rubenstein, Matsushita, Berninger, Raskind, and Wijsman (2011) evaluated spelling as a dyslexia phenotype. The found four loci that explained variation in spelling that corresponded with sites also identified for reading, with some evidence of mediation by Verbal IQ. Given that reading and spelling abilities are highly correlated and represent a common factor that shares heritability (Marlow et al., 2001), it remains to be seen how these findings differ from those reported above for word reading.

Summary

Neuroimaging studies have not really isolated components specific to LDs in written language, which most studies focused on good and poor spelling. It is not possible to define a neural network specific to writing and there appears to be considerable overlap with networks associated with other complex linguistic and motor tasks. Spelling shares considerable heritability with phonological decoding. Writing shows heritability similar to reading and reading-related skills, although there are also unique genetic influences (Olson et al., 2013).

INTERVENTIONS FOR WRITTEN LANGUAGE DISABILITIES

Interventions for LDs to improve written language have been studied less extensively than instructional practices for reading disabilities or even mathematics disabilities. Interventions for transcription difficulties have been more frequently studied than for written composition. As suggested above, written expression is a complex domain that involves multiple cognitive and linguistic processes.

In the past two decades, the research base on interventions addressing different components of written composition has nevertheless grown, and Steve Graham and colleagues have synthesized this research in a series of meta-analyses. Graham, McKeown, Kiuhara, and Harris (2012) focused on the elementary grades; Graham and Perrin (2007) focused on adolescents; and Graham and Harris (2003) focused on the most thoroughly researched and effective practice for students with LDs: self-regulated strategy development intervention (SRSD). These syntheses included true experiments as well as quasi-experiments, and focused on writing quality as the outcome of interest. The first two meta-analyses included students with and without LDs and considered a range of practices. By contrast, Graham and Harris only considered SRSD and focused specifically on struggling writers. There are also additional meta-analyses on handwriting instruction (Santangelo & Graham, 2016), spelling (Graham & Santangelo, 2014; K. J. Williams, Walker, Vaughn, & Wanzek, 2017), and word processing (Morphy

& Graham, 2012).

We rely on these syntheses for identifying effective or promising practices in the area of written expression and estimate their effects; then we illustrate the research base by describing some representative studies. We first focus on intervention designed to improve foundational skills: transcription, handwriting, spelling, or keyboarding. Then we discuss interventions designed to improve the quality of written composition. Within written composition, we first address explicit instructional methods and finish by providing an overview of other instructional techniques that are sometimes used to supplement explicit instruction and sometimes used instead of explicit instruction.

Teaching Transcription Skills (Handwriting, Spelling, and Keyboarding)

Transcription is a foundational component of written expression. It includes handwriting, spelling, and keyboarding—skills that are needed to "write" or communicate the message. As we already discussed, difficulties in both printing and cursive writing stem from a number of factors that include motor deficits, visual–motor coordination problems, visual memory deficits, and orthographic processing. The term "dysgraphia" has been used historically to refer to a developmental difficulty in transducing visual information to the motor system, with motor planning and execution problems (Johnson & Myklebust, 1967), which manifests itself in an inability to copy and write legibly. In a machine-learning study designed to separate dysgraphic and nondysgraphic writing samples of elementary school Hebrew writers, Rosenblum and Dror (2017) found that the best discriminators were different indices of production time. Children with dysgraphia also put more pressure on the paper and have poor quality of letter formation.

In terms of spelling, as described in Chapter 6, English spelling (orthography) is an alphabetic system in which phonemic units (speech sounds) are represented by graphemes (letters or letter combinations). For both students and primary-grade teachers, this fundamental relation between spoken and written language is the most important aspect underlying literacy development. Spelling is often a stumbling block for individuals with WLRD:

appropriate intervention tends to improve their decoding skills, even as their spelling skills continue to suffer (Bruck, 1987). This finding, combined with the number of people who read well but spell poorly (Frith, 1980), suggests that reading and spelling are to some extent dissociated and that theoretical models of one domain will not necessarily explain the other (Moats, 2005).

In any case, serious deleterious consequences for written expression of poor handwriting or spelling are manifested in several ways. They result in misinterpretations of the author's meaning (Graham, Harris, & Chorzempa, 2002; Graham et al., 2000). They create negative perceptions about the writer, which taint overall impressions about the quality of an essay (Hughes, Keeling, & Tuck, 1983) and interfere with the execution of composing processes because cognitive resources are unduly allocated to the mechanical aspects of the process (Berninger, 2004). Moreover, they lead students to avoiding writing, which further constrains writing development.

Meta-Analyses

Across these three transcription areas, Graham et al. (2012) identified five handwriting studies, three spelling studies, and one keyboarding study that were true or quasi-experiments, that included writing quality outcomes, and that focused on the elementary grades (in this case, grades 1–3). Results indicated that explicit teaching of transcription skills improved writing quality, with a moderate effect size of 0.55. In a meta-analysis of handwriting instruction, Santangelo and Graham (2016) identified 80 studies addressing different components of handwriting instruction. All involved explicit attempts to teach handwriting; there are no controlled studies on learning handwriting as an incidental process. When compared to control conditions in which handwriting was not taught, intervention resulted in greater legibility (effect size = 0.59) and fluency (0.63). These estimates are based on multiple studies and are likely reliable. The next sets of effects are based on small samples of studies, but are meaningful. Individualizing handwriting instruction was effective (0.69), as was the use of technology for teaching handwriting (0.85), such as the use of a digitized tablet. Handwriting instruction also was associated with improved quality (0.84), length (1.33), and fluency of compositions (0.48). Ineffective practices included motor training with no explicit graphemic component in relation to no motor training (legibility, 0.10; fluency, −.07); or in relation to handwriting training (legibility, 0.18; fluency, −.06). The null result for motor training may indicate that generalization does not occur because transcription is strongly linked to composition and is heavily influenced by the cognitive and linguistic components of writing. Other ineffective practices include a program, Handwriting without Tears (Owens, 2004) (legibility = 0.13; fluency = 0.18); multisensory tracing of letters (legibility = 0.02); and copying letters from memory (legibility = 0.26).

Some other methods are ineffective. This includes the idea that spelling skills can be learned without instruction or through reading and writing (Graham et al., 2012). The most popular approaches to writing instruction, often referred to as process approaches and epitomized by the writers' workshop (Calkins, 2003), do not have strong evidence of efficacy. Process approaches are difficult to define, but typically involve planning, composing, generating ideas, and writing for real purposes and audiences, with little emphasis on foundational skills. In a meta-analysis of process-writing approaches, Sandmel and Graham (2011) identified 29 studies that yielded a statistically significant, modest effect size of 0.34 on writing quality.

Graham and Santangelo (2014) synthesized 53 studies of spelling instruction for over 5,000 students in kindergarten–grade 2. Spelling instruction improved spelling outcomes compared to no spelling instruction (effect size = 0.54) or incidental spelling instruction (0.43). Increasing time in formal spelling instruction was effective (.70). The gains generalized to contextual spelling when writing (0.94) and were maintained over time (0.53). There were also improvements in reading skills (0.44) or phonological awareness (0.51), and effects were similar across grades and levels of spelling and reading ability. In another synthesis that largely involved case studies, K. J. Williams et al. (2017) found that explicit spelling instruction was effective for teaching spelling and reading in children identified with LDs in grades 4–12.

Specific Intervention Practices

To understand effective instructional practices included in these meta-analyses, a description of some approaches and primary studies is helpful. An intervention focus on transcription skills has a long history. In 1967, Johnson and Myklebust conducted a substantial amount of clinical research, from which they developed a comprehensive task-analytic model for the treatment of handwriting difficulties. In fact, many teachers still use an even older method for the remediation of written language deficits in students with LDs, based on Gillingham and Stillman (1965). With this method, (1) the teacher models a large letter on the blackboard, writing and saying the name; (2) the student traces the letter while saying the name (this tracing stage continues until the student is secure with both the letter formation and the name); (3) the student copies the letter while saying the name; and (4) the student writes the letter from memory while saying the name. In addition to these types of multisensory intervention methods, some older studies have assessed the utility of improving handwriting by teaching students to verbally guide themselves through the process (Hayes & Flower, 1980). Graham and Santangelo (2014) and Santangelo and Graham (2016) did not find strong evidence for the efficacy of these older approaches for transcription.

More recent work has also focused on early intervention for handwriting difficulties. Berninger et al. (1997) randomly assigned first-graders with poor legibility and automaticity in handwriting to five intervention conditions: conventional repeated copying of letters; conventional imitating the motor components of letter formations; provision of visual cues for letter formations; writing letters from memory with increasing delays; and combination of the visual cues/memory component. After 24 lessons over a 4-month period, the combined treatment was more effective than control or other conditions in improving handwriting. These findings were replicated by Graham et al. (2000) and Jones and Christensen (1999).

Berninger et al. (2006b), who completed three studies evaluating different levels of intervention for first- and second-grade students, found that intervention providing practice in motor activities with no letter component or providing practice in letters with no motor component led to some improvement in letter formation and legibility. However, only explicit instruction in handwriting that combined motor and orthographic components with verbal mediation and visual cuing improved automaticity of writing and generalized to improved word recognition skills. In the second study, motor training or orthographic training in isolation did not add to outcomes produced by explicit instruction in automatic letter writing and composition. The third study showed that the addition of explicit instruction in handwriting to instruction in reading improved handwriting, but did not add to reading outcomes. This line of research highlights the importance of integrative approaches that include considerable emphasis on actually producing letters and words.

Graham et al. (2000) also conducted an experimental intervention study with first-grade students who were experiencing handwriting and writing difficulties. Thirty-eight students were assigned randomly to handwriting intervention or to a contrast condition involving phonological awareness instruction. Handwriting instruction involved 27 15-minute lessons divided into nine units. In each unit, three lower-case letters, which shared common formational characteristics, were introduced and practiced. Each lesson incorporated four activities. The first, Alphabet Warm-Up, focused on learning to name each letter, matching the name with its letter, and knowing the sequence of the letters in the alphabet. The second activity, Alphabet Practice, provided tracing and writing individual letters. The third activity, Alphabet Rockets, was designed to increase students' handwriting fluency, and the fourth activity, Alphabet Fun, allowed students to play with the letters in a creative manner. Across these four components, instruction was explicit, relying on a task analysis of the letters to focus the child's attention on the critical features and demands of the task and to provide adequate support for the child to enjoy success until independent mastery was demonstrated.

Results showed that, at posttest, students in the handwriting condition made greater handwriting gains than children in the control condition (effect size = 1.39). This advantage was still evident 6 months later (0.87). Effects were also revealed on posttest compositional fluency (1.46), but dropped to a nonsignificant effect size of 0.45 at the 6-month maintenance assessment. This pattern held for low-performing students with and without identified LDs. Interestingly, however, at posttest, students in the handwriting condition did not produce qualitatively better stories than peers in the control condition. This is in contrast to the Graham et al. (2012) meta-analysis, and it suggests that additional research on handwriting intervention, perhaps conducted in conjunction with composition tasks, is needed.

Examples of spelling intervention studies include Graham et al. (2002). These researchers focused on poor spellers in second grade, randomly assigning them to spelling versus math (control) supplementary instruction. Forty-eight sessions, each 20 minutes in duration, were conducted. The lessons were divided into eight units, each focusing on two or more related spelling patterns. The first lesson was a word-sorting activity, in which students categorized words by the spelling pattern featured in that unit. The teacher engaged the students in thinking about similarities and differences between the words—modeling the thinking process by which words might be sorted into their appropriate categories. Gradually, students assumed responsibility for sorting while articulating the features by which they categorized. Once all the words were sorted, the teacher provided the rule for the patterns emphasized in the word sort. After that, students generated words of their own. Then the deck of words was shuffled and students sorted again, trying to beat their previous times. During Lesson 2, the teacher gave each student eight study words that had occurred frequently in the student's writing and that the child had missed on the pretest. In Lessons 2–5, students employed two study procedures to learn these eight words: self-study using a set of steps and dyadic practice using games. Also as part of Lessons 2–5, teachers provided explicit instruction and practice in identifying the sound patterns associated with the unit's content, and the students worked in pairs to build words that corresponded to the spelling pattern emphasized in that unit. In Lesson 6, students took a test to determine their mastery of the eight words. They then scored the test; plotted the score on a graph; and set a goal for how many words they would spell correctly on the next unit test, which would be added to the graph. Students also completed a test assessing their spelling of nine words that contained the rimes emphasized during word sorting. Cumulative review was conducted systematically, beginning with the second unit.

Results demonstrated the value of this systematic and explicit approach to spelling instruction. As compared with peers in the math control condition, students who received the spelling intervention made greater improvements on norm-referenced spelling tests (effect size = 0.64–1.05), a writing fluency test (0.78), as well as a reading word-attack measure (0.82). Six months later, students in the spelling treatment maintained their advantage in spelling

(0.70–1.07) but not on measures of the writing fluency (0.57) or reading word attack (0.47). Spelling instruction did, however, have a positive effect at maintenance on the reading word recognition skills of students whose pretest scores were lowest.

Berninger et al. (2002) assigned third-graders to interventions that involved training only in spelling, training only in essay composition, and a combined spelling/essay composition training, along with a control condition involving keyboarding. The spelling component emphasized orthographic patterns in words, particularly at the morphological level. Both interventions that included spelling instruction produced more spelling gains than the essay condition that did not involve explicit spelling intervention. Together, these studies illustrate that many students with LDs respond to spelling intervention, and that such gains are maximized with explicit focus on letter patterns (orthography) and opportunities to practice in writing.

Keyboarding is often considered an alternative for students who present with dysgraphia or otherwise struggle with reading. The effects of keyboarding need to be considered in a developmental context because handwriting, including printing and cursive, and written spelling, seem linked with the development of writing and reading. Learning to type in order to write via computers is beneficial, especially for older students. Morphy and Graham (2012) synthesized 27 studies of weaker writers in grades 1–12 that yielded 77 independent effects. Word processing was associated with improved writing quality (0.52), greater length of stories (0.48), improved text organization (0.66), correctness of spelling and punctuation (0.61), and motivation to write (1.42). People preferred word processing over handwriting (0.64). A limited number of studies showed strong effects for word-processing programs that provided feedback on the quality of text or suggestions for revising (1.46). There was no impact on vocabulary or grammar.

These studies were predominantly of older students in middle and secondary environments and did not isolate the effects of keyboarding, the third transcription skill. In elementary students, keyboarding per se may not facilitate written language skills as much as in older students who access word processors, and may interfere with successful development of the neural networks needed for writing. Berninger and Amtmann (2003) reviewed evidence for the efficacy of a variety of compensatory tools supporting handwriting and spelling, including keyboarding, which did improve transcription if typing was slow. Students who had difficulty with automatic production of letters in a paper-and-pencil format also had difficulty with keyboard components. Alves et al. (2016) randomly assigned grade 2 students to 10-week intervention groups (four weekly 30-minute sessions) promoting handwriting, spelling, and keyboarding. Students in the handwriting condition showed greater writing fluency and wrote stories that were longer and of higher quality than students in the keyboarding group. Berninger et al. (2009b) compared grade 4 students identified with handwriting and spelling difficulties on three writing tasks by hand and by typing involving letters, sentences, and essays. Students with LDs and typically achieving students took longer to complete the sentence and essay tasks when typing than by hand, and also wrote longer essays by hand. Longcampe, Zerbata-Poudou, and Velay (2005) trained two groups of students 3–5 years of age to copy alphabetic letters by hand or by typing them. After 3 weeks of training, handwriting training was associated with better letter recognition than tying training, especially in older children. We also reviewed a neuroimaging study of preliterate children that found that printing, but not keyboarding letters, activated the neural network for reading (James & Engelhardt, 2012).

With increasing reliance on computers to generate text, keyboarding skill demands research to understand intervention methods that ensure criterion levels. If such levels can be achieved, keyboarding may represent an effective bypass tool for students with specific deficits in handwriting. For students with multiple deficits related to writing (handwriting, spelling, keyboarding, and/or composing skills), however, what has been learned in reading is likely to be true in writing: Integrating handwriting, spelling, or keyboarding into the actual process of writing is important. In younger students, writing may have more impact on learning to transcribe as well as learn phonological and orthographic relations essential to reading and spelling (Santangelo & Graham, 2016; Weiser & Mathes, 2011).

Teaching Composition

In producing a written composition, the student must simultaneously attend to the subject, the text, and the audience. It also calls upon students' problem-solving processes, oral language abilities, and reading skill, attention, and memory. In addition, studies (Berninger, 2004; Hayes & Flower, 1980; Hooper et al., 1994; Kim & Schatschneider, 2017) show that individuals who write well are goal-directed; understand the purpose of the writing assignment; have a good knowledge of the topic prior to writing; generate more ideas and use significant numbers of transitional ties; produce a more cohesive text and flow of ideas; and continuously monitor their written products for spelling and grammatical accuracy.

By contrast, Hooper et al. (1994) and others (e.g., De La Paz et al., 1998; Graham et al., 2017) reported that writers with LDs demonstrate deficits in deploying strategies during production of written text and also have problems in generating text, showing difficulties with expressive vocabulary, genre, and organization. As compared to students with average or strong writing skill, students with LDs produce shorter and less interesting essays, with poorly organized text at the sentence and paragraph levels, and they are less likely to review spelling, punctuation, grammar, or the body of their text to increase clarity (Hooper et al., 1994; Graham et al., 2017). These observations have led to interventions targeting specific aspects of the writing composition process. Some of the most effective of these interventions involve explicit instruction, the efficacy of which has also been demonstrated in reading and math interventions.

Self-Regulated Strategy Development:An Explicit Method

In their meta-analysis, Graham et al. (2012) summarized the effects of five explicit instructional interventions, four of which showed positive effects. Strategy instruction is one form of explicit intervention. The meta-analysis included 20 such studies. Most (14) taught students strategies for planning or drafting; four studies encouraged planning, drafting, or revision; and two just revising. Most (16) of the studies focused on genre-specific strategies. Ten included the full range of learners, but nine focused specifically on "struggling writers," which included low-performing students with and without LDs. All 20 studies reported positive effects on writing quality, with a combined effect size of 1.02. Effects, however, were significantly different, depending on the type of strategy instruction employed in the studies. For self-regulated strategy instruction (SRSD; 14 studies), the average effect size was 1.17; for other forms of strategy instruction the effect size was 0.59.

More specifically, Graham and Harris (2003) conducted a meta-analysis of 18 students with SRSDs. Most were identified with LDs, and students spanned grades 2–8 (the median grade was 5). As described in that meta-analysis, SRSD incorporates the following principles. (1) Explicit and extensive instruction is provided on writing strategies, self-regulation, and appropriate content knowledge. Self-regulation strategies include instruction on goal-setting, self-monitoring, self-instruction, and self-reinforcement. (2) Interactive learning and active collaboration from the students are encouraged. (3) Instruction is individualized to the student's needs using feedback and appropriate supports, and instruction is paced according to the students' needs, such that each student must achieve mastery of one stage of instruction before the next is introduced. (4) New strategies and new ways of using previously taught strategies are introduced continuously over the course of intervention.

Among students with LDs, SRSD was found to be effective for improving the quality of written composition (effect size = 1.14); writing elements (2.15); story grammar (3.52); and composition length (1.86). Other findings included the following. SRSD had moderate-to-strong effects on the revision skills of students with LDs, and these effects were maintained over time (although surface-level revisions were less well maintained). Students with LDs generalized their newly acquired writing and revising strategies to new writing genres and settings. Moreover, SRSD effects were strong when taught by researchers or school classroom teachers.

Finally, the 2003 meta-analysis indicated that SRSD's self-regulation component (i.e., goal setting, self-monitoring, self-recording, self-statements, and teacher modeling) significantly contributed to writing performance. This last finding aligns with the results of the more recent Graham et al. (2012) meta-analysis. Six studies that evaluated the contribution of explicit self-regulation instruction provided in the context of strategy instruction (five involved SRSD) produced an average effect size of 0.50.

The third explicit instructional practice evaluated in Graham et al. (2012)

was text structure instruction. This primarily involved stories (not informational text). The average effect size was a similar 0.59 standard deviations. For explicit instruction on creativity or imagery, the effect size increased to 0.70. For the last explicit instructional practice, however, effects were not supportive. The four studies explicitly teaching grammar resulted in a combined effect size of -0.41 on written composition quality.

Nonexplicit Instructional Methods

Returning to the Graham et al. (2012) meta-analysis, the literature also provides estimates of effectiveness for nonexplicit instructional techniques that are sometimes used to supplement explicit instruction and sometimes used instead of explicit instruction. The first is peer assistance in the writing process, for which Graham et al. reported an effect size of 0.89. Three of the four studies involved struggling writers. Three involved students assisting each other in the revision process, while the other study encouraged peer assistance throughout the writing process.

Such an approach is illustrated in Graham, Harris, and Mason's (2005) SRSD study. They taught struggling writers two genre-specific strategies that were embedded in a more general strategy for planning and writing a paper. The more general strategy reminded students to pick a topic, organize ideas into a writing plan, and use/upgrade this plan while writing. Within the second step of this general strategy (i.e., organize ideas into a writing plan), students were taught the two genre-specific strategies for generating ideas: the first for writing a story and the second for writing a persuasive essay. Further, students learned about the basic parts of a story and a persuasive essay, the importance of using words that make a paper more interesting, and self-talk to facilitate performance. Finally, a self-regulation component was overlaid on the instruction. With self-regulation, students set goals to write complete papers, monitored and graphed their success in achieving this goal, compared their preinstructional performance to their performance during instruction, and credited their success to the use of the target strategies.

At the same time, the study examined the effect of peer mediation in enhancing the effects of the strategy instruction, especially for the purpose of maintenance and generalization. In the peer-mediated condition, students worked together to promote strategy use, identifying other places or instances in which they could apply the strategies and brainstorming about how they might need to modify the strategies for the new application. They were then encouraged to remind each other to apply what they were learning to those transfer situations, and in the next session they identified when, where, and how they had applied the strategies.

Thus, this study incorporated three conditions: writers' workshop (control condition), SRSD, and SRSD with peer mediation. The control condition represents a popular process approach to expressive writing in many public schools. Seventy-two students, screened into the study because of difficulty with writing, were assigned to pairs (even though pairs were not relevant to writers' workshop or SRSD). Pairs were, however, used as the unit of analysis in the study analyses (to control for the influence students in the peer mediation condition had on each other). For this reason, in all three conditions, pairs (not individuals) were assigned randomly to the three conditions. Instructors worked with students three times weekly, for 20 minutes each time, with approximately 11 hours of total instruction across the two genres.

Results showed the advantage of both SRSD conditions over writers' workshop for planning and composing stories and persuasive essays. Students in the SRSD conditions wrote longer, more complete, and qualitatively better papers for both genres, with effect sizes ranging between 0.82 and 3.23. These effects were maintained over time for story writing and generalized to a third uninstructed genre, informative writing. Moreover, the peer mediation component augmented SRSD by increasing students' knowledge of planning and enhancing generalization to informative and narrative writing.

With encouraging effects at grade 3, Harris, Graham, Mason, and Friedlander (2008) moved to second grade, conducting a parallel study that incorporated the same three conditions. Results were again strong. Among the struggling writers, SRSD produced greater knowledge about writing and stronger performance in the two instructed genres (story writing and persuasive writing) as well as two uninstructed genres (personal narrative and informative writing). Effect sizes were similarly strong. Peer support again augmented SRSD by enhancing specific aspects of students' performance in the instructed and uninstructed genres. Across the two studies, findings revealed (1) the capacity of SRSD to enhance relatively young students' writing performance, even within the high-poverty communities where this series of studies was conducted; (2) the added value of a peer support component to help with generalization of the targeted genres to untaught genres; and (3) the superiority of a structured, explicit, systematic approach to writing instruction over the more popular (and more constructivist) writers' workshop.

Another possible supplement to explicit intervention involves prewriting activities. The Graham et al. (2012) meta-analysis examined the effects of four studies including this component; three were, however, conducted with the full range of learners. With prewriting activities, students made notes or drew pictures prior to writing, while one study had students gather relevant information from the Internet. The effect size was a moderate 0.54. Another instructional component evaluated in the meta-analysis was the use of product goals, in which teachers provide students with specific goals for their compositions (e.g., the number of reasons they should cite to support their thesis; the number of revisions they must conduct). Product goals resulted in a mean effect size of 0.76. Although the effect size was 0.71 for studies on the full range of learners but only 0.43 for struggling learners, this difference was not statistically significant. The final instructional practice evaluated in Graham et al. (2012) was writing assessment. This included teacher feedback on student papers or on student progress for a specific writing skill; it included peer feedback; and it included self-assessment (i.e., the student using rating scales introduced by the teacher). The mean effect size was 0.42, but adult feedback was significantly more effective than peer- or self-feedback (0.80 vs. 0.37).

Summary: Interventions for Written Expression Disabilities

Intervention studies in handwriting and spelling demonstrate how explicit instruction can produce better outcomes for students with LDs on transcription skills that are foundational to written composition. Results also suggest how work targeting these foundational skills may not only improve quality of written composition, but also may simultaneously enhance word attack, word recognition, and reading comprehension. Essential principles for teaching written expression, especially in students with LDs, are summarized in Table 9.1. Three findings from the Graham and Hebert (2011) meta-analysis on writing to read require emphasis: (1) comprehension of science, social studies, and language arts texts improves when students write about what they read. Their recommendation was to encourage students to respond to a text with personal reactions or analyses and interpretations; write summaries of text; write notes about text; and write answers about a text or generate written questions for self-response or response by other students. (2) Teaching the skills and processes involved in creating text improves reading skills and comprehension. On this basis, they recommended that teachers teach text structures for producing essays as well as paragraph or sentence construction skills to improve reading comprehension; teach spelling and sentence construction skills to improve reading fluency; and teach spelling to improve word-level reading skill. (3) Simply increasing the amount students write in structured contexts linked to what they are reading enhances text comprehension, with the obvious recommendation to require high levels of written text production.

TABLE 9.1. Interventions: Fundamental Principles for Teaching Written Expression LDs

1. For transcription difficulties in younger students, explicitly teach handwriting and spelling with deliberate, speeded practice.
2. In older students, minimize demands for motor output. Use adjuncts such as word processors, keyboards, and spell-checks.
3. For generation problems, teach written expression in terms of genres, using self-regulation learning strategies.
4. Permit oral expression (if the LD is specific to writing and not a more general language problem) and dictation as compensatory approaches for older students who have not responded to instruction.

CONCLUSIONS: WRITTEN EXPRESSION DISABILITIES

In contrast to our current understanding of oral language and reading disorders among students with LDs, less is known about the etiology, developmental course, prognosis, and treatment for disorders of written expression. The distinction between the transcription (handwriting, spelling, keyboarding) versus generation (composition) problems is important. This is because the distinction may lead to the identification of skill deficits involving handwriting, spelling, and composition, which are separable in some students with LDs. Core cognitive processes have been identified, but these tend to be shared with other disorders. Neurobiological research is emerging, but distinct neural networks related to written expression have not been identified. Genetic studies show high heritability of written expression skills, including spelling, but with genetic overlap across writing skills and with reading and reading-related processes. There is also evidence for unique genetic effects and some potential candidate genes for spelling.

A key for the future is to attempt to identify subgroups of children with disorders of written expression that have some independence from other language-based disorders. Handwriting is the obvious example. But even here, the comorbidity issue has not been addressed adequately, and the independence of ADHD and handwriting disorders have not been adequately researched (Barkley, 2015). It is apparent that transcription and composition, although separable, are interdependent. In particular, transcription constrains writing quality; many children with LDs involving reading or with oral language disorders cannot produce narrative text because they cannot spell (Bishop & Clarkson, 2003).

As in other academic domains, clinicians and teachers must be aware that written composition is a complex domain. It requires the integration of oral language, written language, cognitive skills, and motor skills. In this context, a combination of the different general and specific intervention methods discussed is likely to net greater improvement in the writing skills of students with LDs. The key is to identify the basis for the impairment (handwriting, spelling, composition) and provide explicit instruction using one or a combination of the evidence-based approaches outlined in this chapter. Evidence-based strategies derived from this research have been distilled into lesson plans in handwriting, spelling, and composition (Harris et al., 2008). In the area of written composition, the most effective and well-researched single strategy to date is SRSD. This method reflects the core cognitive impairment in executive functions that characterize many students who struggle with composing, by teaching children explicit strategies that focus on problem solving, planning, and self-regulation in the context of written composition. Application of the research-based methods in schools is limited, which reflects the translational issues discussed in Chapter 11.

CHAPTER 10

The Problem of Automaticity

Developing proficiency in reading, math, and writing requires that basic skills like decoding, simple arithmetic, and transcription become automatic processes that require minimal conscious effort. In this chapter, we discuss automaticity as a general issue affecting children and adults with LDs. In all three domains, the fundamental concept is the need to automatize basic skills, largely through explicit, structured, and supported practice, so that cognitive and linguistic resources do not need to be allocated to the basic skills required to achieve higher-order performance.

Conscious attention to decoding makes reading slow and laborious, reducing access to processes needed to construct meaning, especially WM and the development of situation models (see Chapter 7). Similarly, the need to use procedural strategies to calculate math facts (i.e., simple addition, subtraction, and multiplication problems) interferes with the allocation of resources for more complicated, multistep calculations, word problems, or algebraic problem solving. Even at this higher level of math performance, formulae need to be automatized for easy retrieval and application. In writing, the speed with which a person can transcribe determines his or her capacity for focusing on the thematic organization of an essay or story. Children and adults who struggle with transcription can often tell a stronger story than they can write.

In the previous edition, we discussed whether a specific subgroup of reading difficulties exists involving fluency. In the present edition, we instead provide a more general discussion of automaticity because across academic domains, an important component of intervention for students with LDs is the use of strategies and practice to improve fluent performance. One central question, which can not be resolved with the present level of evidence, is whether automaticity is a more general factor affecting many with LDs regardless of whether their primary academic skill deficit is in reading, math, or writing, where overlap (comorbidity) in the affected domains is clearly apparent. At the end of the chapter, we discuss some concepts from brain function that may help address this issue. This chapter does not discuss transcription difficulties in written expression, which is the source of automaticity difficulties in written expression because transcription is pivotal to written expression and was discussed in Chapter 9.

READING FLUENCY

In reading, the importance of automaticity was clearly outlined in a theoretical paper by LaBerge and Samuels (1974), who proposed that visual processing in reading becomes automatized through a series of information-processing stages in which accuracy, which requires conscious attention, becomes routinized so that conscious attention is not required for comprehension. In essence, this means that decoding becomes so automatic that the reader can go from the orthographic presentation of the word directly to its meaning. As we indicated in Chapter 6, this type of automaticity occurs because the ventral brain systems are hypothesized to become attuned to the pattern of letters and the statistical probabilities in which letters are organized in orthography. The need to automatize lower-order basic skills in order to allocate resources for the construction of meaning is echoed in other broad accounts of reading development and proficiency (e.g., Perfetti, 2007). The importance of automaticity is encapsulated in the dual-route and connectionist models of reading discussed in Chapter 6 (Taylor et al., 2013). Depending on the properties of the word, the need for sublexical, or phonological processing, is minimized, and the proficient reader moves directly from the orthographic representation of the word to its meaning by accessing the sight word repertoire of the word form area in the fusiform gyrus. When word recognition occurs without conscious attention to decoding, there is immediate access to the meaning of the word, and the higher-level, integrative comprehension processes described in Chapter 7 can be the primary source of resource allocation.

Academic Skill Deficits

The primary core academic skill deficit characterizing people with reading fluency problems is slow reading rate for words and text, which is a proxy for the automaticity of word and text reading. Contemporary views of reading fluency conceptualize it as more than just an outgrowth of word recognition skills, although, from a developmental perspective, this is clearly the progression that occurs. For example, the NRP (NICHD, 2000, p. 3–5) defined fluency as "the ability to read a text quickly, accurately, and with proper expression." Meyer (2002, p. 15) defined fluency as the "ability to read connected text rapidly, smoothly, effortlessly, and automatically with little conscious attention to decoding."

The importance of fluency, however, extends beyond the development of word recognition skills and involves the concept of automaticity (LaBerge & Samuels, 1974). When decoding is an automatic process, oral reading of text is effortless and requires little conscious attention, thereby permitting more resources to be allocated to higher-order processing of the meaning of the text (Wolf et al., 2003). Moreover, people develop fluency difficulties despite accurate word recognition because of difficulties with attention, executive functions, and other skills that influence the efficient allocation of resources. This is especially apparent in children with ADHD (Denckla et al., 2013). Fluent readers can perform multiple tasks simultaneously, likely because of the efficient use of cognitive resources that reflect the operation of these skills. Most definitions of fluency also include an emphasis on prosody, or the ability to read with correct expression, intonation, and phrasing. We do not discuss this component of fluency, because poor readers' lack of prosody is not usually regarded as a disability. It is instead viewed as secondary to the problem with automaticity.

The assessment of fluency difficulties is less daunting than measurement of reading comprehension (see Chapter 7). Excluding prosody, the latent construct is automaticity of word and text reading, so that fluency essentially boils down to the reading rate (adjusted for accuracy). As we discussed in Chapter 5, fluency can be assessed as the amount of time needed to accurately read single words, a list of words, short passages, or longer texts. These measures tend to be highly correlated. Jenkins, Fuchs, van den Broek, Espin, and Deno (2003) found that fluency for reading words on a list and in a passage were both sensitive to reading impairments in WLRD. The latent variable study by Cirino et al. (2013) identified two reading rate factors: one involving the timed reading of word lists and passages and the other the timed reading of sentences and passages, but with a component requiring responses indicating comprehension of the text (e.g., fill in a blank missing word). Moreover, identifying individuals with rate deficits is no more difficult than identifying people with word recognition difficulties. It involves a decision about cut points on a dimension. As we discussed in Chapter 3, problems will emerge when efforts are made to create a subgroup of inaccurate word readers, who also have fluency problems, and compare them with a subgroup primarily impaired in fluency. The skills are not independent and the subgroup impaired in both processes will always be the more severely impaired subgroup compared to the subgroup with impairment in just one domain (Schatschneider et al., 2002). We reviewed efforts to disassociate reading accuracy and fluency subtypes in Chapter 6.

Core Cognitive Processes

Here we return to the issue of rapid automatized naming (RAN) in fluency, which is where the involvement of speeded processing and other cognitive skills tends to emerge. Fluency is also likely related to the ability to process increasingly large sublexical units of words, which some consider an orthographic process. In terms of automaticity, a major question is the link between tasks that mimic the reading process, such as rapid naming of letters laid out left-to-right like text, and linguistic capabilities that support orthographic processing, and whether rapid naming is a proxy for any form of speeded processing.

Rapid Automatized Naming

As we discussed in Chapter 6, researchers argue whether RAN contributes to reading achievement independently of its phonological component (Norton & Wolf, 2012; Vellutino et al., 2004; Vukovic & Siegel, 2006). Any task that requires retrieval of information with an articulatory component has to involve phonological processing. As rapid naming tasks are moderately correlated with phonological awareness measures, this appears to be a reasonable conclusion. In this interpretation, naming speed is essentially a measure of how rapidly an individual can access phonologically based codes. Nonetheless, naming speed and phonological processing are dissociable; just as word recognition and fluency are dissociable (see Chapter 6).

The alternative view, especially relevant for reading fluency, is that measures of naming speed involve nonphonological processes that are also related to reading (Wolf & Bowers, 1999). To complete rapid naming tasks, a variety of cognitive processes may be implicated, including executive functions that involve response inhibition and set shifting and lexical processes that permit retrieval and naming (Georgiou & Parilla, 2013; Norton & Wolf, 2012; Wolf et al., 2003). These processes are also involved in fluent reading of text, begging the question of what rapid naming tests actually measure. Wolf et al. (2003, p. 361) noted that "the components of naming speed represent a mini-version or subset of the components of reading."

Naming speed is not the most predictive component of rapid letter naming. In a component analysis of rapid-naming tasks, both Neuhaus, Foorman, Francis, and Carlson (2001) and Wolf and Obregon (1992) found that the pause time between stimuli, which is when these other cognitive processes should be operating, was most strongly associated with reading difficulties. Of course, the interstimulus interval is when other cognitive processes, involving attention and executive functions, and lexical retrieval would be operating if a person was reading a passage out loud. Clark, Hulme, and Snowling (2005) found that rapid naming of letters and digits accounted for unique variance in exception-word reading when phonological skills were controlled. However, in evaluating the different components of naming, neither the average item duration nor the average pause duration uniquely predicted reading ability. Rather, the number of pauses in naming was the unique predictor. Thus, deficits of rapid naming were interpreted as "top-down" or strategic factors that reflect differences in reading practice and experience. This view is consistent with findings of Schatschneider et al. (2004), who suggested that rapid letter naming predicted reading fluency because it was a simple assessment of reading ability, noting that only letter and digit naming—not color and object naming—seemed to uniquely predict reading ability when a regression method identifying unique contributions of individual variables (dominance analysis) was used. In contrast, studies with children reading more transparent languages (Norwegian; Lervåg & Hulme, 2009) found that assessments of nonalphanumeric RAN in this more transparent language predicted early reading skills if measured well before the onset of formal reading instruction. If measured after reading instruction had begun, alphanumeric RAN was most predictive. These researchers suggested that RAN exercised a causal influence on later reading ability because rapid naming of objects requires neural circuits responsible for object naming and orthographic processing. Sideridis, Simos, Mouzaki, and Stamovlasis (2016) found that RAN number naming mediated the relation of decoding and reading fluency nonlinearly, such that children who fell below a threshold of RAN speed showed generally dysregulated reading fluency, again suggesting commonality of the processes involved in RAN and fluent text reading. What is clear is that performance on RAN tasks is not a simple measure of processing speed (Hulme, Nash, Gooch, Lervåg, & Snowling, 2015).

The processing speed issue leads to the question of the specificity of rapid-naming deficits to reading fluency difficulties. Some investigators have found evidence for deficiencies on any task involving speeded and/or serial processing (e.g., Waber et al., 2001; Wolff, 1993). Catts, Gillispie, Leonard, Kail, and Miller (2002) found that some poor readers have a general deficit in speeded processing that accounts for their rapid-naming deficits. Speed of processing also uniquely predicted reading outcomes, with Catts et al. suggesting that such measures represented an "extraphonological" influence in some children's reading difficulties. Waber et al. (2000) demonstrated that unlike phonological awareness tasks, rapid naming measures do not differentiate children who have learning difficulties in other areas. Children with ADHD often show difficulties on measures of rapid automatized naming (Tannock, Martinussen, & Frijters, 2000). Based on these types of data, Waber et al. (2001) argued that these difficulties reflect common brain-based problems with timing or rapid processing that occur across many learning impairments. The relation of measures of processing speed and RAN is unclear and is important because processing speed measures that involve an alphanumeric code seem to represent the most consistent source of overlap in comorbid reading and math disorders (Willcutt et al., 2013).

Orthographic Mapping

Another major correlate of developing fluent and automatic word and text reading likely involves the capacity to process increasingly large units of words (Foorman, 1994). As the child becomes able to instantaneously recognize increasingly larger units of words, perhaps through programming of the ventral components of the reading network, word recognition becomes automatized, which allows more efficient allocation of resources to comprehension processes.

The early phases of learning to read involve learning to match the orthographic units that are present in text to their phonological representations in speech. With phonological recoding, children can access many relations of the orthographic units they see with the sounds of words that exist in the spoken language (Foorman, 1994; Ziegler & Goswami, 2005). As reading develops, children learn more about the orthographic patterns based on letter shapes and the statistical properties of letter order and become able to process increasingly large units of words, recognizing many words by sight. The origin of this development, however, is in the phonological code. This reflects the relation of phonemes to print (Lukatela & Turvey, 1998).

Cross-Linguistic Issues

The issue of phonological–orthographic mapping is especially important for understanding how reading develops in different languages. English is a language characterized by often arbitrary relations of sound and print, particularly because many orthographic units have multiple pronunciations. Some conceptualizations of reading disabilities are narrow because of their focus on accuracy, which may be more related to learning to read in a less transparent language like English (Share, 2008; Wimmer, 2006). Other languages have much more transparent relations of phonology and orthography. For example, Finnish, Italian, and Spanish are languages in which the pronunciation of words is fairly reliably signaled by how the word appears in print. English has low transparency, while German tends to be in the middle. The question is whether these differences in the relation of phonology and orthography are related to the development of reading and whether reading problems are different by virtue of this variation.

Caravolas and colleagues (Caravolas, Volin, & Hulme, 2005; Caravolas et al., 2012; Caravolas, Lervåg, Defior, Seidlová, Málková, & Hulme, 2013) completed a series of studies comparing the early development of reading across several different European languages. They found that the complexity of syllables and the transparency of the orthographies influenced how rapidly children learned to read. In languages that had inconsistent orthographic structures in which the syllabic structure was also complicated, such as English and Danish, reading developed most slowly. In contrast, reading developed faster in a language like Italian or Spanish that has a relatively shallow and transparent orthography and simple syllabic structures. Thus, learning to read is affected by the complexity of the orthographic relations in print as children begin to read.

Ziegler and Goswami (2005) identified three factors that influenced reading development across languages: (1) the availability of phonological units that can be explicitly accessed before reading; (2) the consistency of orthographic units, which may have multiple pronunciations, and phonological units, which may have multiple spellings; and (3) the size of the orthographic unit that is available within the written language system, which they termed the “granularity problem.” They developed the “psycholinguistic grain size” theory to help explain differences in lexical organization and processing strategies that would characterize skilled reading across different language orthographies.

Even in languages like Chinese, children are sensitive to the phonological components that are expressed in Chinese logograms and attend to the regularity by which the phonological component of the Chinese logograms affects pronunciation (Hanley, 2005). Although skilled English readers have strong phonological awareness skills, Chinese readers tend to have better syllable and morpheme awareness (Tan et al., 2005). However, all of these linguistic skills are operating at a sublexical level with variations that reflect, in part, the relation of the phonological and orthographic units.

In examining differences in the manifestations of reading problems across different languages, phonological skills still seem to drive the acquisition of word recognition and fluency (Caravolas et al., 2005, 2012; Wimmer & Mayringer, 2002). However, in orthographies where the relation of phonological and orthographic units is less consistent, such as in English, many more readers who are inaccurate will emerge. Pseudoword reading is especially difficult. Aro and Wimmer (2003) compared pseudoword reading controlled for letter patterns, onsets, and rimes in German, Dutch, English, Swedish, French, Spanish, and Finnish speakers in grades 1–4. Only English was associated with a low rate of accuracy (about 90%) by the end of grade 1. In German poor readers, difficulties with pseudoword reading were less common; poor reading was often characterized by fluency and spelling problems when the primary correlate was phonological processing. However, some German poor readers had poor fluency with adequate spelling and had more difficulties with rapid naming (Wimmer & Mayringer, 2002).

In a longitudinal study, Wimmer, Mayringer, and Landerl (2000) composed German-speaking groups based on the double-deficit model involving patterns of performance on phonological and rapid naming tests (Chapter 6) and compared their reading and spelling development 3 years later. They reported that phonological awareness deficits earlier in development were weakly linked with phonological decoding in German, but more strongly related to spelling and foreign-word reading. In contrast, naming speed was related to reading fluency, spelling, and foreign-word reading. They suggested that when reading was taught with synthetic phonics methods in a language with a more regular relation of phonology and orthography, the acquisition of reading was less affected in earlier phases by phonological processing than in later phases that build up orthographic relations of words.

In contrast, Ziegler, Perry, Ma-Wyatt, Ladner, and Schulte-Körne (2003) found that patterns of reading strengths and weaknesses were similar in German and English speakers: Both groups showed more difficulties in reading pseudowords than real words and had slow reading speeds. In a study of Dutch children (de Jong & van der Leij, 2003), phonological awareness and rapid naming tasks were administered in kindergarten, grade 1, and grade 6. Rapid naming discriminated good and poor readers through grade 6; phonological awareness deficits diminished by grade 6. This finding may have reflected a ceiling effect because, in a second study, poor Dutch readers struggled with phonological processing if task demands were increased. Caravolas et al. (2005) found that phonological awareness was a unique predictor of reading in Czech- and English-speaking children and that good and poor readers in both languages had similar phoneme awareness difficulties.

In another review of studies across several countries and languages involving phonology, Goswami (2002) argued that the core deficit in WLRD and the manifestation of this difficulty varies depending on the orthography of the language. More recently, Caravolas et al. (2012) compared early literacy development in English, Spanish, Slovak, and Czech in a 10-month longitudinal study. They found comparable predictive validity for phonological awareness, letter-sound knowledge, and rapid naming at the onset of instruction for both reading and spelling skills. WM did not account for unique variance, similar to other longitudinal studies in English (Wagner et al., 1997). In a subsequent analysis, Carovalas et al. (2013) reported that growth trajectories were slower in English than in the more transparent Czech and Spanish languages, but phonological awareness, letter-sound knowledge, and rapid naming were still the best predictors. The latter studies provide particularly strong evidence for the relative contributions of phonological awareness and rapid naming tests to early reading in English, where accuracy is the best predictor. However, if the outcome is reading fluency, rapid naming in kindergarten is a more dominant predictor (Schatschneider et al., 2004). It is likely that phonological awareness, letter-sound knowledge, and rapid naming all measure components of the early ability to represent written letters and sounds at a phonological level, which becomes less important as these skills are automatized.

RAN and Orthographic Processing

RAN has been extensively studied in cross-linguistic research, often as a proxy for reading fluency. Wolf et al. (2003) linked rapid naming to orthographic processing, relying in part on cross-linguistic research. Georgiou, Parilla, and Liao (2008) examined the relation of RAN digit and color naming with assessments of reading accuracy and rate in Chinese, English, and Greek. Across all three languages, they found stronger relations of RAN digit naming than color naming with assessments of rate and concluded that differences across languages may reflect the use of rate versus accuracy measures. However, in English and Greek, pause time between digits and colors was a stronger predictor of accuracy and fluency than in Chinese; in Greek, articulation time was a stronger predictor than in English and Chinese.

Even in English, Manis, Doi, and Bhada (2000) found that rapid naming accounted for significant variability in reading even when phonological awareness and vocabulary were controlled. The unique contribution of rapid naming was stronger for orthographic processing; phonological awareness was more closely related to pseudoword reading. Other studies that address relations of rapid naming and orthographic processing found that both phonological and orthographic processing are related to rapid naming, questioning whether the rapid-naming component is specific to orthographic processing (Holland, McIntosh, & Huffman, 2004). Georgiou and Parilla (2013) concluded that there was no conclusive evidence regarding the basis for the strong, cross-linguistic relation of RAN and reading, identifying 13 hypothesized mechanisms.

In a meta-analysis of the association of RAN tasks and four domains of reading (decoding real words and pseudowords, sentence and text reading, and reading comprehension), Araújo et al. (2015) found little evidence of cross-linguistic differences. RAN performance was significantly correlated with all four domains (average $r = .43$). The correlations were higher for real-word decoding and text reading than for pseudoword decoding and reading comprehension. Fluency measures were also more highly correlated with RAN tasks than accuracy measures. Letter and digit naming were more strongly related to reading domains than nonalphanumeric stimuli. The relation was stronger with fluency in grades 1–8, but diminished over time in relation to accuracy.

These findings are not surprising given that RAN is a simple version of a text-reading task with little semantic information. As Araújo et al. (2015) and

Georgiou and Parilla (2013) observed, there is little evidence for a single, unitary explanation of why RAN tasks are related to reading, including explanations based on phonological and processing speed mechanisms, and orthographic depth or processing. RAN tasks simply mimic the text-reading process and are timed, which is why the strongest relations are between timed assessments of serial naming of alphanumeric stimuli and reading fluency. In many respects, especially the need to consider differentiated instruction, the rate–accuracy dissociation proposed by Lovett (1987) may be more meaningful. Nonetheless, understanding why RAN relates to reading fluency should reveal common mechanisms for stimuli that in the case of RAN do not require extensive experience to learn.

Interventions

Interventions that address the automaticity of reading skills have been developed. These interventions do not focus on improving RAN task performance, for which there is little evidence outside of interventions addressing letter-naming fluency in young children (e.g., Struiksma, van der Leij, & Stoel, 2009). In the best examples, fluency interventions are paired with interventions that address the development of decoding skills, epitomized by the pairing of decoding skills with strategy instruction (e.g., Lovett et al., 2013, 2017). Even if a person has specific difficulties with fluency and not accuracy, interventions of this sort can be applied in isolation of a decoding component to address fluency issues. However, if the person has a history of WLRD, it is likely that problems with complex multisyllabic words and spelling will be present. It is also clear that without an explicit comprehension component, generalization of fluency training to comprehension will be less likely, although the evidence for this transfer has not been adequately studied in poor readers.

Most attempts to intervene in the fluency area usually involve students who began with problems with word recognition, and typically include both word recognition and fluency components in the intervention. The most frequently studied interventions provide supported reading practice. This includes repeated reading or guided oral reading, along with increasing the amount of time students spend independently reading scaffolded books geared at their instructional level.

Empirical Syntheses

The NRP (NICHD, 2000) reviewed classroom and tutorial studies addressing intervention studies involving fluency. The panel identified 16 studies that included 398 students who were poor readers and 281 students who were good readers. The NRP found comparable, moderate effect sizes (around 0.50) for both poor readers and average readers. Although a variety of intervention programs were examined, the only domains in which they could be characterized as effective involved repeated reading and other supported reading interventions. In general, these types of interventions involved repeated oral readings with a model or with a peer or parent. The NRP was not able to document relations of simple, unstructured practice or exposure, described as silent, sustained reading, with improved reading ability.

Kuhn and Stahl (2003) followed the NRP report by including studies that involved repeated reading, assisted reading in clinical settings, and approaches to fluency development that involved the entire classroom. This synthesis did not compute effect sizes, but we include it here because it was a follow-up to the NRP report. This review confirmed the NRP finding: practice-based interventions for fluency are efficacious. However, gains were generally lower in students with reading difficulties. Approaches that involved some form of assistance, such as reading with a model or listening during reading, appeared more effective than approaches that did not involve assistance, such as silent sustained reading. These findings suggest that adult guidance and monitoring is a critical component of fluency instruction. Kuhn and Stahl noted that evidence was less supportive of simple repeated reading of passages and stories. This suggests that time spent in oral reading of connected text, as opposed to repetition, may be responsible for the effect of repeated reading on fluency and comprehension.

In a response to the inability of the NRP to document a relation of unstructured reading exposure to improved reading, Lewis and Samuels (2003) conducted a meta-analysis of a broad array of studies examining the effects of simple unguided practice and exposure on reading ability. This unpublished study is important because many in the reading community believe that simple exposure is not only effective, but the primary pathway to proficient reading (Krashen, 2004). It is clear that proficient readers read more and less proficient readers read less, which leads to hypotheses about the Mathew effect (Chapter 6).

Lewis and Samuels (2003) examined 10 reviews of the literature on reading practice programs that largely relied on sustained silent reading; six reviews concluded that independent reading time had positive effects on reading comprehension. They found 43 studies that examined the relation of time spent reading and level of proficiency. These studies largely utilized outcome assessments in which votes were tabulated on whether the intervention was effective. Collapsing across what is clearly a weak metric, Lewis and Samuels concluded that a strong, positive relation exists between reading exposure and achievement. In 49 additional studies, 17 were correlational and 25 were quasi-experiments (about half of which had some sort of comparison group). Across studies, there was a positive relation, with a cross-study correlation of .10 between exposure and reading. This correlation, while significant, is low. Altogether, while exposure seems to exert a positive relation on achievement, the effects are small and much less robust than those reported in controlled studies involving some form of structured practice. Similar conclusions are apparent for methods that involve reading aloud simultaneously (sometimes referred to as the neurological impress method) or reading with a tape recorder (Therrien, 2004).

In a synthesis of intervention studies addressing students with LDs, Chard et al. (2002) found 24 published and unpublished studies that reported specific findings involving fluency. These studies, which included repeated reading, both with and without a model, and sustained silent reading, evaluated issues involving the number of repetitions, text difficulty, and the extent of improvement. Chard et al. (2002) found 21 studies that addressed whether repeatedly reading text resulted in improved reading fluency in students defined with LDs. These studies yielded an average effect size in the moderate-to-large range (0.68). In 14 studies, almost all single cases involving modeling by an adult, positive effect sizes in the small-to-large range were also reported. Peer modeling was also associated with small-to-moderate effect sizes, as was the case for modeling with an audiotape or computer. A variety of factors influenced effect size estimates, including the amount of text, text difficulty, number of repetitions, types of feedback, and criteria for repeated reading. Chard et al. concluded that an emphasis on fluency building as part of either classroom or tutorial interventions is essential to improving performance in this domain.

Therrien (2004) conducted a meta-analysis examining the effects of repeated reading on fluency and comprehension. From the 19 studies, there were 28 effect sizes for repeated reading in which students read the same passage multiple times (nontransfer) and 27 effect sizes for studies in which the intervention involved repeated reading of different passages several times (transfer).

As in the NRP, Therrien found that repeated reading improved reading fluency and comprehension in good and poor readers. For good readers, the effect size was 0.76 for fluency and 0.48 for comprehension; for poor readers, the effect sizes were moderate: 0.77 for fluency and 0.59 for comprehension. The effects were stronger for nontransfer methods (0.83 for fluency, 0.67 for comprehension) than for transfer methods (0.50 for fluency, 0.25 for comprehension). However, if the transfer involved reading aloud to an adult, the effects were larger (fluency, 1.37; comprehension, 0.71). Therrien concluded that regardless of nontransfer or transfer approaches, repeated reading should be done with adults who cue the passage (i.e., "read as fast and accurately as you can"), and indicated that reading the same passage three to four times was most effective. Corrective feedback on errors was also associated with larger effect sizes.

Interventions for Struggling Readers

As Torgesen et al. (2001; see Figure 6.6) dramatically demonstrated, a common finding in remedial approaches for students with word recognition deficiencies is improvement in word reading and comprehension, but little change in fluency. Although early intervention will help address some of these difficulties for many students, the reduced efficacy of many remedial approaches may be due to persistent word recognition difficulties that could have been reduced through earlier intervention. Early intervention programs do impact fluency as well as word recognition (Torgesen, 2002).

This finding may well reflect the earlier access to print afforded by early intervention and more rapid development of decoding skills. Specifically, early intervention promotes the opportunity for the repeated exposures to words that facilitates rapid processing at a larger orthographic level. Nonetheless, many remedial studies show that students who respond to instruction in the alphabetic principle continue to have fluency difficulties. In turn, many of these students may be unable to comprehend primarily because their slow reading rate places too many demands on their ability to process what they have read. In addition, students who are not fluent do not enjoy reading, so they are less likely to read, which contributes to the failure to build sight word vocabulary, a key to the development of accurate and fluent reading skills.

Supported Reading

The methods identified as most consistently effective in the meta-analyses involve adult-supported repeated reading of the same passage with error correction and cuing. Less clear are the effects of reading different passages, also with adult support. Stahl, Huebach, and Cramond (1997) developed fluency-oriented reading instruction, a classroom approach to facilitate automatic word recognition and fluency with three components: (1) a redesign of the basal reading lesson to include specific components involving fluency; (2) a period involving free reading in school; and (3) reading at home. The redesign of the basal reader differentiated instruction by dividing students into two groups based on their reading levels, with modifications of fluency instruction based on the amount of assistance needed. The school and home components were designed to increase the amount of time spent reading connected text, but these components were structured and monitored by adults.

An initial evaluation of the program (four teachers, two schools, eventually expanded to 10 teachers and three schools) showed positive results. On average, students gained about 2 years in overall reading growth on an informal inventory. Of particular importance was the finding that over the 2-year period, even struggling readers improved in fluency, with only two of about 105 students reading below second-grade level by the end of the year. Reading practice clearly improved fluency.

Kuhn et al. (2006) conducted a large-scale investigation of fluency-oriented reading instruction that included control groups. The study involved eight schools and 24 classrooms across two sites. Fluency-oriented reading instruction was compared with a program that emphasized repeated reading of a wide range of materials. A third group served as a classroom curriculum control and was followed over time with no researcher-provided intervention. Historical controls were also included. Kuhn et al. reported that both interventions resulted in better outcomes than the historical and curriculum controls, with no systematic differences between the two treatments.

In that study, results were especially dramatic for struggling readers, who also received supplementation of the fluency program using Direct Instruction principles to address decoding weaknesses. In evaluating performance relative to students who had been in the same school programs in the past (historical controls), improvements in word recognition, oral reading rate and accuracy, and comprehension were apparent. The effects on struggling readers, many of whom likely had LDs, were especially interesting.

The key to both approaches may involve the scaffolding of texts to the readers' instructional level. Stahl (2004) suggested that scaffolding may explain why approaches like the two interventions reviewed above improve fluency and comprehension and why sustained silent reading (e.g., drop everything and read) fail to improve reading performance. In addition, as reported in the Therrien (2004) and the NRP (NICHD, 2000) studies, supported approaches are more effective than approaches based on simple exposure.

These conclusions were supported by O'Connor, White, and Swanson (2007) who compared repeated reading with a method based on reading a wide range of material. The participants were children identified as struggling readers, most of whom were also identified by the schools as students with LDs. In both experimental conditions, the students practiced with an adult for 15 minutes three times weekly for 14 weeks. They read aloud and received error corrections as they read. The material was selected to match the student's instructional level in reading. In comparison to a business-as-usual comparison group, both treatment groups showed significant growth in fluency and comprehension over the 14 weeks, with no differences on outcomes between the two treatment groups.

Multicomponent Programs

Read Naturally. A commercial program specifically targeting fluency is Read Naturally (Ihnot et al., 2001). In Read Naturally, students read nonfiction passages designed for students in grades 1–8. Students practice oral reading of short, interesting passages (i.e., repeated reading), read along with a recording of the passage at a challenging pace, and time and graph their reading rates (e.g., words correct per minute) so they are constantly aware of their progress. A comprehension component involves discussing passages with the teacher and answering questions about what the students read. Few studies have been conducted on Read Naturally. Exceptions are Hasbrouck, Ihnot, and Rogers (1999) and Denton et al. (2006b). Hasbrouck et al. reported cases that had benefited from Read Naturally, but these were not controlled evaluations. Denton et al. found that 8 weeks of instruction (1 hour per day) based on Read Naturally led to significant improvement in reading fluency skills in students who had failed to respond to an earlier intervention, but with little improvement in decoding or comprehension. More research on the effectiveness of this approach to reading fluency is needed.

Retrieval, Automaticity, Vocabulary Elaboration, and Enrichment with Language–Orthography. Not surprisingly, fluency has emerged as a major emphasis in the remedial area, with newer efforts perhaps best characterized by the Retrieval, Automaticity, Vocabulary Elaboration, and Enrichment with Language–Orthography (RAVE-O) program developed by Wolf et al. (2002). It is designed to facilitate the development of automaticity in reading subskills, to facilitate fluency in decoding and comprehension processes, and to enhance interest and engagement in reading and language use in students with LDs.

RAVE-O is based on a developmental model of fluency (Wolf et al., 2003) that emphasizes the multiple contributions to proficient comprehension made by the student's familiarity with common orthographic patterns, as well as the student's knowledge of a word's meaning(s), morpheme parts, and grammatical uses. A major premise is that the more a student knows about a word, the faster the student will retrieve and read it. The game-like format includes intensive work on rapid orthographic pattern recognition; building word webs; learning word retrieval and comprehension strategies; playing games with language through computer games enhanced with animation; and rapid, repeated reading of short (1-minute) mystery stories that incorporate the multiple meanings and syntactical uses of core words.

This program is typically used in conjunction with a word recognition program and was evaluated along with the PHAST Track Reading Program by the Morris, Wolf, and Lovett research group described in Chapter 6. Morris et al. (2012) found that RAVE-O enhances word recognition, fluency, and comprehension better than instruction based only on the decoding skills programs. To date, there is no strong evidence from these studies that RAVE-O produces larger gains in fluency at the word level than a program like PHAST, which teaches strategies for generalizing from the alphabetic principle to larger sublexical units at the morphosyntactic level. This finding highlights the importance of including instruction focused on increasingly large sublexical units for people with decoding and fluency difficulties.

One question is whether a program like RAVE-O leads to greater improvement in reading connected text, as well as comprehension, as compared with a program that emphasizes the generalization of word recognition strategies, such as the WIST component of PHAST. This possibility would reflect the focus of programs like RAVE-O on fluency at the sublexical, word and connected text levels. Previous theories of how fluency emerges focused on accurate and fluent word recognition, which is supported by the comparable word-level fluency results of programs like WIST. Findings that RAVE-O leads to stronger gains in fluency at the connected text level and in reading comprehension would support a more comprehensive approach to intervention at the text level of proficiency and comprehension. In addition, it would be interesting to compare the effects of programs like RAVE-O and WIST in children identified as having fluency, but not accuracy, difficulties.

Transfer Effects to Words and Text. Although programs like RAVE-O focus more broadly on fluency at the text level, most theories of how fluency emerges also focus on accurate and fluent word recognition. This is supported by programs like PHAST that include generalization strategies like WIST. To illustrate, in a series of remedial studies specifically addressing fluency deficits (Levy, 2001), English-speaking students identified with fluency difficulties, most of whom had concurrent word recognition problems, received a variety of interventions. The studies were specifically designed to evaluate whether transfer in fluency is mediated at the level of word recognition or at the level of text reading. In general, these studies showed that the reading fluency of poor readers is limited by their slow rate of processing at the level of the individual word, which is consistent with the nature of fluency difficulties for speakers of more transparent languages. Levy found that simple practice in a “repetition of names” game led to significant gains in word recognition skills, particularly for poor readers. Words were learned best through word-training study, in which the student was taught to read a list of words as fast as possible. The alternative involved having students read a story four times in succession that contained the same word. For the poor readers, transfer to improved reading speed occurred regardless of whether a similar or different story context was used.

Levy (2001) reported that context was not an essential component of the experience and that teaching automaticity of word reading was possible for poor readers and also made them more successful. There was clear evidence for transfer across linguistic levels in context. In other studies, there appeared to be little additional benefit of highlighting shared orthographic units. Yet, blocking according to the orthographic unit, which has the effect of making the orthographic relation more explicit, resulted in greater automaticity. These results are consistent with the premises of RAVE-O, showing that grouping words into similar orthographic patterns accelerates fluency.

Levy (2001) noted that many poor readers were very slow in generalizing across words. In other languages, one of the few studies targeting students with specific fluency difficulties (Thaler, Ebner, Wimmer, & Landerl, 2004) provided computerized training on repeated reading of 32 words over 25 days to a sample of 20 German-speaking children. Each word was designed to emphasize the onset segment and was presented up to six times per day. Although fluency for reading the trained words improved over the 5-week period, there was only a slight improvement in reading untrained words. In a different study of poor readers who spoke Dutch, de Jong and Vrielink (2004) trained grade 1 students to rapidly name serially presented letters. There was little evidence of improvement when rapid serial naming of letters was directly trained. Thus, training students on orthographic processing even in more transparent languages does not generalize strongly to new words and is difficult to achieve.

Speeded Practice. If the goal is to simply help the reader become faster readers, why not use the popular speed reading methods that are widely available? These methods essentially teach skimming skills and offer faster reading times with little effect on comprehension. In a comprehensive review of reading and the effects of interventions based on speed, Rayner, Schotter, Masson, Potter, and Treiman (2016) found clear evidence for an accuracy-speed trade-off. If the proficient reader processes 250 words per minute, increasing speed to 500–750 words per minute clearly results in reduced comprehension of the material. They noted that reading speed was not the primary limitation. Rather, language skills such as vocabulary delimit comprehension ability. Reading speed can improve through practices that improve language comprehension because "language skill is at the heart of reading speed." (Rayner et al., 2016, p. 4). Applications designed to move text more rapidly are constrained by the limits of visual information processing and the span with which eye movements can process text. Comprehension monitoring and other strategies for correcting comprehension failure detract from speed, but can be improved through stronger language skills along with more time practicing reading. But taking poor readers who lack automaticity and putting them in trainings or applications designed to simply boost speed has not been found effective.

Summary: Reading Fluency

Reading fluency is an important reading skill that is correlated with but also independent of word recognition. The core cognitive correlates involve rapid naming, orthographic mapping, and other cognitive skills that regulate attention, inhibitory processing, and lexical retrieval. A major issue concerns what is actually measured by rapid-naming tests and whether they are proxies for text-reading fluency. There is clear evidence for a dissociation of reading fluency deficits from those involving word recognition and comprehension, but these are correlated processes and reading fluency grows out of proficiency with decoding in more and less transparent orthographies. It is essential to differentiate needs for code-based instruction from those involving lower-level decoding skills and those associated with the development of fluency and automaticity.

Fluency interventions have focused largely on procedures that lead to repeated exposures to words and text. These approaches are likely to be maximally effective if the reading material is scaffolded to the child's reading instructional level, adults are involved, and error correction is provided. Table 10.1 provides a summary of practices that may be helpful in promoting automaticity, especially in terms of engagement and scaffolding text to the reader's instructional level. As a primary reading method, fluency-based interventions do not substitute for explicit training in decoding or the need to explicitly teach comprehension skills. However, to develop fluency, children need to be engaged with print as soon as they begin to read. One reason that children who learn to decode much later in their development remain slow readers even after intervention may be the cumulative effects of lack of experience, which prevents the development of a sight-word vocabulary (Torgesen et al., 2001). Multicomponent interventions based on broader views of fluency that extend beyond repeated reading and pair the intervention with decoding instruction (Morris et al., 2012), such as with the RAVE-O program, are emerging. As the importance of learning to process larger orthographic units becomes more fully appreciated, it seems likely that approaches to reading and spelling instruction that explicitly focus on these opportunities will be linked to fluency.

TABLE 10.1. Building Reading Fluency through Independent Reading

1. Reader should be able to read text with 90% accuracy.
2. Ratio of known and unknown words should be below 1:20 to facilitate vocabulary acquisition.
3. Content of independent reading should relate to classroom content.
4. Follow-up activity and discussion should be based on independent reading.
5. Teacher and student should share understanding of the purpose of the reading assignment.

MATH AUTOMATICITY

In the domain of mathematics, research on automaticity focuses dominantly on math facts (simple addition, subtraction, and multiplication problems that can be retrieved from long-term memory as declarative information). Assessments of these skills are straightforward and can be accomplished with short, timed assessments of math fact fluency. The major achievement tests reviewed in Chapter 5 all have short assessments of math fluency, and these skills can also be assessed by curriculum-based measures.

In this section, we focus on addition and subtraction, where much of this research has occurred. We begin by reviewing the process by which competence in retrieving answers to such problems develops among typically developing children, and we outline what's known about the processes that support such development. Then we explain the developmental delays and differences students with mathematics LDs experience and discuss intervention strategies designed to alter those trajectories. Finally, we discuss areas where additional work on automaticity in the domain of mathematics is required, with implications for students with LDs.

Academic Skill Deficits and Cognitive Correlates

A primary academic skill deficit involves difficulty retrieving simple arithmetic facts, a developmental process involved in typical children that is often delayed in children who struggle with math. Studies of these processes have largely occurred in the domain of math, but correlations with more general domain-general learning and memory components may be important.

Retrieval of Addition and Subtraction Facts in Typically Developing Children

By the time children enter first grade, most have a rudimentary understanding about number, the counting sequence, and the number line. They also comprehend concepts about addition and subtraction and can count to solve these problems (Geary, 1993). For addition, young children typically count both addends. For subtraction, they represent the beginning quantity (the minuend) with objects and sequentially separate the number of objects to be subtracted (the subtrahend); then they count the remaining set (e.g., Groen & Resnick, 1977). As understanding of cardinality and the counting sequence continues to develop, children discover the number-after rule for adding with 1. They also come to understand that the sum of 5 + 2 cannot be 6 but instead is two numbers beyond 5.

In this way, children discover the efficiency of counting from the first addend and start relying on more efficient counting procedures. For addition, the most efficient counting procedure involves starting with the cardinal value of the larger addend and counting up the number of times equal to the smaller addend (e.g., 3 + 4 = "four: five, six, seven"). For subtraction, the most efficient strategy involves counting up from the subtrahend to the minuend (e.g., 5 − 2 = "two: three, four, five"; the number of counts, i.e., the distance or difference between 5 and 2 is the answer). Frequent use of efficient counting procedures reliably produces the correct association between problem and answer, which results in long-term memories that pair problem stems with answers (Fuson & Kwon, 1992; Siegler & Shrager, 1984). This enables direct retrieval of answers, and the commutativity of addition facilitates retrieval of related addition problems. Subtraction, which is not commutative, is more difficult, but can be facilitated by retrieval of related addition facts (e.g., 8 − 5 = 3, based on 5 + 3 = 8), once children have come to understand the inverse relation between addition and subtraction (Geary et al., 2008b).

Automatic retrieval of math facts thus depends on two processes. One involves understanding concepts and principles about number, addition, and subtraction (e.g., counting on for addition; taking away and difference for subtraction, commutativity). Such understanding is thought to depend on a synergy between earlier forms of foundational mathematics competence as well as domain-general cognitive processes (domain-general because they apply across many academic domains). In the development of automatic retrieval of addition, these include visuospatial memory (Fuchs et al., 2016a) and RAN (Fuchs et al., 2016a; Koponen, Aunola, Ahonen, & Nurmi, 2007; Koponen, Salmi, Eklund, & Aro, 2013) as well as attentive behavior and reasoning (Fuchs et al., 2016a; Geary et al., 2012b). Each cognitive process appears to increase children's capacity to produce the correct responses to simple arithmetic problems—the many correct responses that are required to form associations in long-term memory. Note that these processes are not dissimilar from those that involve reading fluency.

The second process involves the capacity to form and fluently retrieve from memory arbitrary associations between the visual symbolic and phonological forms. This may reflect the functional integrity of the hippocampal-dependent memory system, which engages the prefrontal, parietal, and medial temporal areas during the early phases of learning (Qin et al., 2014) and has been shown to be important in learning arithmetic facts (De Smedt, Holloway, & Ansari, 2011; Qin et al., 2014).

Automatic retrieval of answers to simple addition and subtraction problems is a hallmark of mathematics achievement in the primary grades. Most children achieve such competence by the end of second grade. Every increase in automatic retrieval of one unit at end of second grade is associated with an increase of .20 standard deviation units at end-of-third-grade mathematics calculation skill (Fuchs et al., 2016b), and every increase in automatic retrieval of one unit at end of fourth grade is associated with an increase of .09 standard deviation units at end-of-fourth-grade word-problem skill (Fuchs et al., 2016b). The National Mathematics Advisory Panel (2008) concluded that strong reliance on retrieval when solving simple arithmetic problems is a critical step toward eventual mastery of high school algebra, a gateway for later entry into mathematics-intensive fields.

Accuracy versus Fluency

Recent work provides insight into the contribution of early arithmetic fluency over accuracy. Although a large literature demonstrates a role for one or the other, we identified only two studies that included both variables in the same models to estimate the contribution of one while controlling for the effects of the other. Carr and Alexeev (2011) concluded that both dimensions of second-grade calculation skill contribute to fourth-grade performance on a general mathematics achievement test. Fuchs et al. (2016b) extended that study by examining how fluency and accuracy affect two specific types of higher-order mathematics performance, while controlling for early number knowledge. This latter study identified a more complicated pattern. The role of accuracy versus fluency depended on the type of higher-order mathematics outcome. Although intermediary calculation accuracy played a role in both outcomes, its direct path coefficient was twice as large in predicting the prealgebra outcome than in forecasting future word-problem performance. Moreover, intermediary calculation fluency provided added value in predicting fourth-grade word-problem-solving, but the effect of intermediary calculation fluency on fourth-grade prealgebraic skill was not significant. These findings indicate the need for additional research to identify the types of complex mathematics outcomes that depend on accuracy versus automaticity.

Developmental Delays and Differences in Children with LDs

Research demonstrates that students with mathematics LDs show consistent delays in the adoption of efficient counting procedures, make more counting errors when executing those efficient counting procedures, and fail to make the shift toward memory-based retrieval (e.g., Geary et al., 2012b; Goldman et al., l988). Most of these children eventually catch up to peers in skilled use of counting procedures, but difficulty with retrieval tends to persist (Geary et al., 2012b; Jordan et al., 2003). Students with mathematics LDs retrieve fewer answers from memory and when they do retrieve answers, they commit more errors (e.g., Geary et al., 2007).

Intervention Studies

Research shows that skills involving retrieval of math facts and simple problems can be improved in students with LDs, although effects of randomized control trials are not consistent. L S. Fuchs et al. (2006b) conducted a randomized control trial to assess the effects of providing students with practice in quick correct responding. An addition or subtraction problem with its answer briefly flashed on a computer screen; then students generated the problem and answer from short-term memory. This occurred for 10 minutes, twice weekly across 18 weeks. Compared to an analogous computer-assisted practice condition in spelling, arithmetic practice produced significantly better performance for addition but not subtraction; effect sizes were 0.95 and –0.01. Other outcomes were not assessed.

Two randomized control trials combined number knowledge tutoring with practice and assessed a broader range of outcomes. In Fuchs et al. (2005a), tutoring occurred three times per week for 16 weeks. Each session included 30 minutes of tutor-led instruction designed to build number knowledge plus 10 minutes of computerized arithmetic practice, as just described. Results favored tutoring over a no-tutoring control group on measures of concepts and applications (0.67), procedural calculations (0.40–0.57), and word problems (0.48), but effects were not reliable on simple arithmetic (0.15–0.40). Bryant et al. (2011) integrated tutoring on number knowledge with practice (four times per week for 19 weeks). In each session, 20 minutes were devoted to number knowledge and 4 minutes to practice, which focused on arithmetic problems, as well as reading numerals, counting on/back, writing dictated numerals, and writing three-number sequences. Effects were significantly stronger for tutoring compared to a no-tutoring control group on simple arithmetic (0.55), place value (0.39), and number sequences (0.47). But tutoring did not enhance word-problem outcomes (–0.05 and 0.07).

In the only randomized control trial to focus exclusively on number knowledge, Smith et al., (2013) evaluated Math Recovery, in which tutors adapt lessons to meet student needs as reflected on Math Recovery assessments. Tutors introduce tasks and have students explain their reasoning, but practice is not provided. Tutoring was designed to occur four or five times per week, 30 minutes per session, across 12 weeks. At end of first grade, effects favored Math Recovery over the control group on fluency with simple arithmetic (0.15), concepts and applications (effect size = 0.28), quantitative concepts (0.24), and math reasoning (0.30). Effects were stronger for students who began tutoring below the 25th percentile (0.31–0.40), but are generally smaller than in the other studies. Comparisons are, however, difficult because this study allowed fidelity to vary, whereas the other studies tried to ensure fidelity.

These randomized control trials suggest the potential of tutoring programs for enhancing some forms of mathematics learning, including fluency with math facts, among at-risk first graders. In perhaps the largest study, L. S. Fuchs et al. (2013b) investigated the effects of first-grade number knowledge tutoring with contrasting forms of practice. Tutoring occurred three times per week for 16 weeks. In each 30-minute session, the major emphasis (25 minutes) was number knowledge; the other 5 minutes provided practice in one of two forms: nonspeeded practice, which reinforced relations and principles addressed in number knowledge tutoring, or speeded practice, which promoted quick responding and use of efficient counting procedures to generate many correct responses. At-risk first graders were randomly assigned to number knowledge tutoring with speeded practice (n = 195), number knowledge tutoring with nonspeeded practice (n = 190), and control (no tutoring, n = 206). Each tutoring condition produced stronger learning than in the control condition on all four mathematics outcomes. In terms of the distinctions between the two practice conditions, *speeded practice produced stronger learning than nonspeeded practice* on fluency with simple arithmetic problems and fluency with two-digit calculations; effects were comparable on number knowledge and word problems. Importantly, effects of both practice conditions on arithmetic were partially mediated by increased reliance on automatic retrieval.

The bulk of evidence therefore suggests that first-grade intervention among children who are at-risk for mathematics LDs can alter the trajectory of these children's automatic retrieval of math facts. These studies also suggest that effects transfer to more complex forms of calculations but not necessarily to performance in other mathematics domains.

Summary: Math Automaticity

As reflected in this section, research on automaticity in the domain of mathematics has focused predominantly on the development of automatic retrieval of simple math facts in young children, with several studies showing strong efficacy. L. S. Fuchs et al. (2013b) also demonstrates the added value of speeded practice, in which children are supported to use strategic behavior to answer problems (if confident of the answer, say it; if unsure, use the efficient counting procedure to count up) with a time penalty for guessing (i.e., as time elapses, the student immediately corrects an error, using the efficient counting procedures taught in the program), for producing stronger outcomes over number knowledge tutoring with nonspeeded practice provided in a game format. In evaluating these studies, it is important to keep in mind that in older children, math fact skills can be improved as well by allocating 5 minutes to strategic math-fact practice within every 25-minute multicomponent intervention session focused on word-problem solving (reviewed in Chapter 8). This proved as effective as providing 25-minute intervention sessions focused solely on math facts. Questions remain, however, about the extent to which automaticity of larger units of mathematics operations or automaticity with strategies occurs and, if so, what is the relation of such automaticity to more advanced forms of mathematics competence. Future studies might, for example, address the following questions: Does automaticity in solving simple algebraic expressions (e.g., $x = 3$; $2x + y = 10$; $y = ?$) support more advanced mathematics achievement in representative samples? Does the automaticity with which individuals solve such expressions distinguish students with mathematics LDs versus students without mathematics LDs? Does intervention designed to promote automaticity in solving simple algebraic expressions eventuate in stronger performance on advanced mathematics topics for students with versus without mathematics LD?

These questions can be posed with other forms of automaticity, such as automatic retrieval of fraction equivalencies for benchmark fractions (e.g., $\frac{2}{4}$ and $\frac{3}{6}$ and $\frac{4}{8}$ for ½), automatic retrieval of conversions between improper and mixed numbers, and automatic reliance on efficient strategies to compare fraction magnitudes. Understanding the answer to these questions, in combination with additional research probing the types of complex mathematics outcomes that depend on accuracy versus automaticity, are important for gaining insight into methods for promoting better outcomes for students with mathematics LDs. For further discussion, see Chapter 9.

NEUROBIOLOGICAL FACTORS

Brain Structure and Function

Few studies have focused on subgroups with isolated automaticity difficulties in reading or math. However, there is evidence for the role of specific brain regions in developing automaticity in both domains.

Reading

Norton et al. (2014) used rhyme detection and lexical decision tasks to evaluate children identified with double and single deficits in phonological awareness and RAN in an fMRI study. There were no major differences in brain activation for the two groups impaired in phonological awareness, with reduced activation involving the dorsal components of the reading network (left inferior frontal and left inferior parietal regions). Activation was, however, more reduced in the double deficit group. In contrast, in the rapid naming subgroup, there was reduced activation of the right cerebellar lobule (Area VI), as well as reduced connectivity of cerebellar–frontal connections. These results support the dissociation of phonological processing and rapid naming observed earlier in this section. However, the RAN deficit group was small ($n = 10$) and the tasks themselves did not manipulate fluency.

Some studies have imaged RAN tasks. Wiig et al. (2002), using regional cerebral blood flow assessments of hemodynamic changes in brain activation, found that rapid-naming tasks involving objects and objects blended with colors activated the parietal lobes. Color naming did not result in reliable changes in brain activation. Misra, Katzir, Wolf, and Poldrack (2004) used fMRI to assess brain activation in response to rapid-naming tasks for objects and letters. They found that the network typically implicated in word reading was activated (see Figure 6.2), with some differences when the letter and color tasks were used. Moreover, there was additional activation of areas involving eye movements and attention, which would be expected in a task requiring serial processing of stimuli. In a study of adults, Lymberis, Christodoulou, O'Loughlin, Del Tufo, and Gabrieli (2009) had adults with and without dyslexia perform RAN tasks for letters and numbers in an fMRI paradigm. In contrast to typical adult readers, who engaged the ventral network early in the naming tasks, those with dyslexia showed underactivation of the ventral stream and more engagement of the dorsal and frontal components of the reading network.

Other studies have manipulated the speed with which reading stimuli are presented. Rimrodt et al. (2009) evaluated fMRI in sentence-reading tasks controlling for word reading. There was more activation in children with WLRD in the dorsal network than in controls, with the latter showing more activation of the ventral pathways. Behavioral performance on fluency tasks was more strongly related to activation of the ventral pathways, while poorer performance was related to the dorsal pathways. In an MSI study of sentence reading involving serial presentation of contextually presented words, Simos et al. (2011a) found widespread bilateral underactivation of dorsal and ventral regions, as well as the left posterior cingulate, in children with WLRD. Latency analyses showed that later engagement of the temporoparietal and ventral regions predicted silent reading of sentences in typically developing children, but not in those with WLRD.

Langer, Benjamin, Minas, and Gaab (2015) manipulated reading speed for sentences in children with WLRD and controls and found increased activation across the reading network in both groups. However, children with WLRD showed less activation in the left fusiform gyrus that correlated with reading fluency task performance. Christodoulou et al. (2014) compared typical adults and adults with WLRD on sentence-reading tasks using a word-by-word presentation that manipulated speed of presentation. As in Rimrodt et al. (2009) and Langer et al. (2015), they found increased activation in both groups as speed increased. Differences were most apparent when behavioral differences were equated, especially in the left prefrontal and superior temporal regions, both associated with semantic and phonological processing.

Finally, other studies have examined changes in patterns of neural activation in relation to reading fluency. In their neuroimaging study of response to a phonologically mediated intervention with children identified with decoding problems (Blachman et al., 2004), B. A. Shaywitz et al. (2004) observed significant changes in the occipitotemporal regions of the brain that they related to improvements in fluency. Similarly, in a sample of poor decoders who received an intervention emphasizing decoding and fluency (Denton, Ciancio, & Fletcher, 2006a), Simos et al. (2007b) used MSI and found normalization of latency of responses in the ventral regions specifically on a task designed to assess the fluency of word reading. Such changes were less apparent on a pseudoword-decoding task that resulted in more changes in the temporoparietal regions.

We identified only a handful of relevant structural MRI studies. Eckert et al. (2003) found that greater rightward asymmetry of the right cerebellar hemisphere and left inferior frontal gyrus correlated with poorer performance on RAN tasks. He et al. (2013) found that gray matter volumes in the dorsal pathways predicted phonological decoding, while naming speed was associated with volumes in occipital, parietal, temporal, and frontal regions. Chang et al. (2007) evaluated adults with periventricular nodular heterotopia and compared them with adults with dyslexia. The group with heterotopias had isolated difficulties with reading fluency, although this is hardly unusual in a group with brain injury. Fluency and rapid-naming deficits were related to the degree of white matter integrity in this group and more generally with the organization of white matter tracts connecting different brain regions.

Mathematics

Although it is difficult to mimic reading fluency tasks in functional neuroimaging studies, tasks involving math fact retrieval are easier to create. De Smedt et al. (2011) examined brain activation for single-digit addition and subtraction in children 10–12 years of age who varied in automaticity of these skills. They found that complex problems and subtraction engaged the frontal–parietal network described in Chapter 7, particularly the intraparietal sulcus. In contrast, addition and simple problems showed more engagement of the hippocampus, especially for problems that required retrieval of math facts. Children low in fact retrieval skills showed less hippocampal activity and more reliance on the right intraparietal sulcus, suggesting that math fact retrieval in these children was not consolidated into declarative memory systems. Bugden, Price, McLean, and Ansari (2012) found more left intraparietal sulcus activity in children with stronger arithmetical fluency. In a study of high school students, brain activation during single digit arithmetic correlated with activation of the frontotemporal components of the math network; greater activation of the right intraparietal sulcus was associated with lower math scores (Price, Mazzocco, & Ansari, 2013). These studies are sparse, but suggest that reduced automaticity of math facts is more associated with brain regions involved in higher-level math calculations and less so with systems involving declarative memory. At the same time, there are also relations of hippocampal volume and math facts in an intervention context (Supekar et al., 2013), so this issue is far from resolved.

Genetic Factors

Reading and math automaticity have been studied in genetically sensitive designs. The findings support the general finding that automaticity assessments of reading and math are heritable and can be distinguished from other reading and math processes, but also share variance with these components, reflecting the fact that all these processes are correlated.

Reading

Although there are no genetically sensitive studies of a reading fluency subgroup, there is evidence for common and separate heritability of the accuracy and fluency of word-reading skills and RAN when treated as dimensions. Davis et al. (2001) found that rapid-naming measures had significant heritability even when reading measures were included in the model. In a study of 800 twin pairs, Compton et al. (2001) found evidence of a common set of genes for phonological processing, rapid naming, and reading in affected twins. This group also showed evidence for genetic influences that were specifically involved in the relation of rapid naming and reading. In contrast, a control group of unaffected twins also revealed common genetic influences for phonology, rapid naming, and reading, but no evidence of an independent relation of rapid naming and reading. There was little evidence of shared environmental influences in the affected group, which included children 8–18 years of age. In a similar sample of mostly older children, Tiu, Wadsworth, Olson, and DeFries (2004) found that measures of phonological processing and rapid-naming skills both made significant genetic contributions to reading. In a study of younger twin pairs, Petrill et al. (2006) found significant heritability of rapid naming of letters that was not explained by phonological measures and that generally had a smaller relation with the environment than phonological measures.

Among a set of reading measures, Petrill et al. (2007, 2010) found much higher heritability estimates for a rapid-naming measure than a word recognition measure. Hart et al. (2016) used a twin sample to evaluate relations of reading fluency, spelling, and earlier assessments of reading comprehension on later reading comprehension in a large sample of students in grades 3 and 4. Reading fluency had a large genetic component and a small unshared environmental component. In a multivariate analysis, they found evidence for a general genetic factor across the four measures. This finding is consistent with the strong correlations among these measures and suggests that they share a common cognitive locus or the same genes. These findings support the hypothesis that naming speed is etiologically distinct from phonological awareness. In this respect, Raskind et al. (2005) compared the heritability of component skills involving accuracy and fluency of pseudoword reading. Using a variety of genetic association methods and a genomewide scan, the researchers found evidence for involvement of chromosome 2 for fluency, but not accuracy, of pseudoword decoding. There was also clear evidence for a shared genetic etiology for these two correlated processes.

Mathematics

Most studies of the heritability of math fluency come from the Case Western Reserve twin sample study led by Stephen Petrill. The assessment is a timed measure of simple single-digit addition and subtraction. In these studies, math fluency has a strong heritable component and is the only math ability assessed that has unique genetic influences. At the same time, there is genetic overlap of math fluency with reading fluency as well as untimed calculation and math problem-solving measures (Hart et al., 2009). In subsequent studies, math fluency continues to emerge as a distinct component of math ability. Petrill et al. (2012) reported that about two-thirds of the variance in math fluency performance was independent of performance on other untimed math measures and that this assessment was stable over two time points.

Summary: Neurobiological Factors of Reading and Math Fluency

Most studies of reading and math fluency are based on functional neuroimaging. Although not studied simultaneously in the same study, comparisons suggest some variations in the reading network when the task involves a requirement for automaticity. This is largely reflected in quantitative changes in the degree of activation of different components of the network. In reading, fluency demands lead to increased overall activation of the network. There is underactivation of the ventral stream and relatively more reliance on the dorsal stream in poor readers. The Norton et al. (2014) finding of cerebellar involvement in an isolated naming speed group is intriguing, but consistent with the role of the cerebellum in automatizing cognitive functions. In math, the greater reliance on the intraparietal sulcus and the underinvolvement of memory-based hippocampal circuits implies lack of memorization of math facts.

Behavior genetic studies show strong and potentially unique heritability of reading and math fluency. The distinctness of math fluency is especially interesting. However, these measures are correlated, especially when the underlying constructs are assessed. There is strong overlap and evidence of general genetic influences, although math fluency still seems to emerge as genetically distinct.

CONCLUSIONS: ARE THERE DOMAIN-GENERAL AUTOMATICITY FACTORS?

Automaticity of reading, math, and writing skills are separable dimensions of performance within each domain. They are likely related to the developmental progression of basic skill development, although it is certainly possible that difficulties with automaticity emerge independently of basic skill difficulties, as the rate–accuracy and double deficit hypotheses imply. Note that similar hypotheses have not emerged in math and written expression, likely because of the absence of an influential variable like RAN. As Norton and Wolf (2012) indicate, RAN is a microcosm of the reading process and analogues to RAN for math and writing do not exist. Treating RAN as an interchangeable process with other assessments of reading fluency for words is hard to understand since RAN is so highly correlated with and predictive of reading rate (Norton & Wolf, 2012). But fluency itself is an important reading skill that is correlated with, but also independent of, word recognition. The neuroimaging and genetics studies show clear evidence for independence of phonological processing/decoding, and rapid naming/fluency as well as shared skills and genes.

A limitation of research is the reliance on the English language (Share, 2008; Wimmer, 2006), with research on other languages more strongly suggesting that children can have specific difficulties involving fluency, especially of single words. In English, the research base addressing children with specific fluency difficulties is sparse, and children with isolated rapid naming/reading fluency deficits occur infrequently compared to those with decoding problems in English. In other languages with more transparent orthographies, difficulties with timed reading tasks are more frequent. There are few studies of math fluency in other languages, but cross-linguistic spelling studies are available (see Chapter 9).

We reviewed automaticity as a separate chapter because it is a problem that tends to characterize many children with LDs. In many respects, it can be the most challenging aspect of intervention, especially if the child does not master basic skills early in development. Multi-component interventions in all three domains seem most effective in addressing both basic skill deficiencies and automaticity issues, but even with the improvements realized via intervention, automaticity problems can persist, especially as the nature of reading text becomes more complicated, as the unit of automaticity required for math success increases, and as the complexity of writing assignments grows.

In reading, it is clear that the ventral systems are not adequately programmed for automaticity due to lack of access to print, reflecting the early problems of children with WLRD with phonological representations of words. This programming involves repeated exposures to orthographic patterns, which is a form of perceptual learning based not only on the forms of the letters, but also on the statistical probabilities of letter orderings. Efforts to explain poor decoding in terms of low-level visual and auditory deficits have not been robust, largely because such hypotheses have to construct a theory of reading that is often incompatible with the conspicuous problem with identifying and spelling isolated words accurately and fluently. Many of the problems identified as deficits are consequences of the inability to read, such as perceptions of crowding (Chapter 6).

In a recent paper from John Gabrieli's laboratory, Perrachione et al.

(2016) introduced the concept of "neural adaptation" in a series of fMRI studies of dyslexia. Neural adaptation is a neurophysiological learning process derived from animal models in which the brain adapts to repeated exposures to the same stimulus. The auditory and visual cortices adapt to and refine the perception of repeated stimuli. Thus, when listening to a person with a strong dialect, comprehension improves over time, which would be less apparent than listening to different people speaking the same dialect.

As Perrachione et al. (2016) point out, individuals with dyslexia show less adaptation on perceptual learning tasks. For example, performing visual and auditory tasks with perceptual noise is more difficult, indicating less rapid neural adaptation (Sperling, Lu, Manis, & Seidenberg, 2006; Ziegler, Pech-Georgel, George, & Lorenzi, 2009). Difficulties with maintaining perceptual consistency and responding to stimulus regularities have been observed in studies with dyslexia (see Chapter 6; Ahissar, 2007). In their fMRI study, Perrachione et al. (2016) used neural adaptation paradigms to evaluate the effects of repeated presentations of objects, spoken and written words, and faces in adults and children with dyslexia and controls. For all four tasks, adaptation was reduced, and there were significant correlations with the reading performance in the group with dyslexia. There was also less activation of brain regions associated with processing of these stimuli; in the condition where written words were presented, the fusiform gyrus showed less adaptation in those with dyslexia than controls, while those with dyslexia tended to show more adaptation in brain regions associated with semantic processing. This systemwide processing problem suggests that people with dyslexia are less able to benefit from repeated exposures to stimuli. This could be related to automaticity difficulties, where structured exposure and practice are critical.

More research is needed to evaluate this idea of a domain-general problem with neural adaptation. Understanding would be enhanced by comparisons with children who have math disabilities. The role of the cerebellum, critical for automaticity, is also not adequately understood. Another candidate for a domain-general deficit is "processing speed." It is interesting that simple assessments of processing speed, such as reaction time to a visual stimulus or crossing out numbers as fast as possible, are not strongly related to rapid naming. This contradicts Kail and Hall's (1994) hypothesis that RAN tasks simply index processing speed (see Norton & Wolf, 2012).

At the same time, studies of comorbid reading and math disorders identify processing speed as a common source of difficulty (Willcutt et al., 2013). The task employed in that study, however, does not represent processing speed as defined by Kail and Hall (1994). It instead involves learning an alphanumeric code in a compressed amount of time. It is also important to note that implicit learning is also impaired in people with reading and math disabilities. Fletcher, Taylor, Morris, and Satz (1982) refuted what at the time was a long-held association of finger agnosia deficits with dyslexia, showing that the difficulties on tasks requiring identification of an unseen stimulated finger occurred not because of simple perceptual problems, but because the child had to learn a code for numbering the fingers. The inability to learn the simple code in kindergarten was highly predictive of future reading problems. Again, this implies that a domain-general problem may explain the difficulties many individuals with LDs experience in developing the capacity for rapid execution of tasks. In explaining automaticity deficits, it may in turn help account for the comorbidity of reading and math disabilities.

Interventions for automaticity in older, hard-to-treat individuals with LDs are perhaps the most important future directions of research on LDs. Automaticity problems can be prevented or mitigated in many children through early intervention and aggressive efforts to teach basic skills in the context of multicomponent interventions that begin in the general education classroom with differentiated instruction. But considerable intensity is required when children fail to respond to early intervention and when responsiveness to remediation in older children and adults is not strong, largely because of the automaticity issue. Developing methods for enhancing automaticity should be a priority and will likely be aided by advances in neurobiological research.

CHAPTER 11

Translating the Results of Scientific Research into Educational Practice

The preceding chapters focused on reviewing and integrating scientific research relevant to the identification, assessment, and instruction of individuals with LDs. To do so, we drew on international research across multiple disciplines, including educational intervention, cognitive development, neuroscience, and genetics. In addition to research specifically devoted to people with LDs, we also relied on research depicting normative development of academic and cognitive skills, as well as their neurobiological bases. In undertaking the second edition, we asked several questions:

1. Has the research since 2007 improved in overall scientific quality and to a point where a second edition was needed?
2. Has the research led to the development of screening, assessment, and instructional methods that can assist practitioners in teaching individuals with LDs?
3. Has the research been implemented in schools and classrooms with success?

We believe that the answers to the first two questions are affirmative. This second edition shows that the evidence base on LDs has expanded dramatically over the past 10 years in quality and in quantity, filling many of the gaps we identified in the first edition. The extent to which the scientific knowledge base has increased across academic domains is noteworthy (see Chapter 2). This includes the strong research base that has emerged in reading comprehension, mathematics, and written language relative to the first edition of this book. Although the research base on word-level disorders remains the largest and most mature, scientific understanding in other academic domains continues to grow, especially in areas involving assessment and intervention.

More integrated service-delivery frameworks have also continued to evolve. These frameworks are epitomized by the use of MTSS as a schoolwide approach to enhancing academic outcomes, especially when MTSS includes all students, including those with disabilities. Intervention studies have been enhanced by the increased application of the rigorous designs we advocated in the previous edition, including randomized controlled trials and regression discontinuity designs. This trend was evident in the Scammacca et al. (2016) meta-analysis of 100 years of reading research. In the last 5 years of this meta-analysis (2010–2014), all since the first edition of this book was published, there was a significant increase in the rigor of intervention studies.

With respect to question 3, it is apparent that when assisted by researchers, schools can implement the service delivery frameworks, assessments, and interventions developed through scientific research (Denton et al., 2010; Gersten et al., 2009). However, without this type of support, efforts to scale evidence-based frameworks have either been poorly implemented and/or not sustained over time (Balu et al., 2015). Many schools continue to find it difficult to scale research to the point of effective implementation, and for this reason often do not obtain the results documented in research studies. The reasons for this research-to-practice gap include problems achieving fidelity and diluted efforts due to a lack of resources, insufficient professional development, and an inability to sustain efforts because of unrealistic expectations of immediate results (Hess, 2008). The latter may be most significant. Implementation of the interventions described at the core-classroom level, which is essential to MTSS frameworks, and the MTSS service delivery system itself, requires years of sustained effort and a long-term view. This can be difficult to sustain as school leaders, education reforms, and governmental policies change so frequently.

To put the third question in context, consider Table 11.1. This table was created from the last chapter of the previous edition of the book (Fletcher et al., 2007), where we concluded by identifying seven barriers to the implementation of scientifically based research into educational and clinical practice. We added an additional item to this list and separated screening and progress monitoring to permit clearer discussion. We first discuss progress on the first seven barriers. Then we address in detail the last barrier shown in Table 11.1, reliance on clinical experience and craft knowledge over scientific evidence. This issue is perhaps the most persistent, pervasive, and major hurdle in improving practice for students with LDs. Finally, we identify future directions for research and practice to address the needs of students with LDs.

TABLE 11.1. Barriers for Moving Research into Practice

1. Inadequate implementation.
2. Insufficient reliance on screening and progress monitoring.
3. Inadequate attention to prevention.
4. Failure to implement research-based methods for intensifying intervention.
5. Insufficient consideration of multifaceted problems.
6. Need for integration across instructional components.
7. Lack of sufficient engagement and practice.
8. Reliance on clinical experience and craft knowledge over scientific evidence.

SEVEN BARRIERS TO TRANSLATION OF RESEARCH: AREAS OF PROGRESS AND REMAINING NEEDS

Over the past decade, there has been inconsistent progress in addressing the first four barriers in Table 11.1. For the first barrier, *implementation of research* remains inconsistent. Many instructional approaches and interventions demonstrate efficacy when they are studied in controlled environments. However, when the interventions are translated into everyday practice in complex school and classroom settings, fidelity suffers. Contextual variables such as teacher preparation and commitment to the intervention, composition of students, and adequacy of resources dilute the efficacy that is apparent in a more controlled research setting (Denton, Vaughn, & Fletcher, 2003). As schools begin to implement modified service delivery frameworks as represented by MTSS in conjunction with validated interventions, the implementation problems will be increasingly formidable. Yet there are relatively few examples of school districts that have been able to scale MTSS with success (Jimerson et al., 2015).

Screening and progress monitoring tools have improved through research, showing progress on the second barrier. To implement any form of early intervention, or framework based on MTSS, different kinds of assessments need to be routinely implemented in schools. The first type is universal screening for academic and behavior problems, where the methods have continued to evolve; in reading, universal screening is required by legislation in many U.S. states, although implementation of methods with established reliability and validity is inconsistent. As with curricula, many tools are aggressively marketed through commercial organizations with little or no evidence of reliability and validity. In addition to screening, progress monitoring of students who show risk characteristics or who demonstrate disabilities is critical. We noted in the previous edition that the technology for progress monitoring is well developed in reading, math, and behavior through the elementary grades and into middle school (Stecker et al., 2005). It has expanded to secondary environments. Educators are more successful at administering screening and progress-monitoring measures than they have been in using them systematically to make educational and instructional decisions. As we showed in Chapter 5, when these measures are used to access progress and make decisions, better outcomes emerge. We return to the topic of ongoing progress monitoring below under the topic of intensive intervention.

Related to the first two barriers in Table 11.1, *early intervention* continues to be more of a slogan than a practice. We expressed frustration in the first edition with the widespread lack of implementation of core and supplemental intervention programs in general education, which can prevent unnecessary development of LDs. Unfortunately, while there are some glimmers of improvement, widespread, high-quality implementation of preventive intervention is not apparent. We reiterate that even the best prevention programs will still prove inadequately effective for some students. Here is where the link of instructional response and disability is critical. If we have strong prevention programs in place to decrease the number of children who struggle due to poor instruction and inadequate preventive intervention, we can be more confident that those who continue to struggle have LDs. We can then provide the level of intensity required to begin to remediate these students. Prevention programs may lead to *gains by all students* and allow more intensive targeting of those who struggle (Torgesen, 2009).

The barrier that we added, which was embedded in the previous edition, is the *need to implement intensive research-based interventions*, especially for students with disabilities. We interpret this as a significant problem attributable in part to the goal of including students with disabilities full time in general education settings. This has resulted in inclusion per se becoming the intervention, often with little special education and/or intensive intervention support. Although the goal of full inclusion is positive because segregated special education programs that operate in isolation of general education are not effective (see Chapter 5), inclusion is not an intervention for students with LDs, who usually exhibit school-related problems because they struggle in the general education classroom. Students with LDs need intensive interventions that are supplemental and often beyond the capacity of general education. Inclusion is a process, not an outcome or intervention. Unfortunately, inclusion is frequently treated as a primary educational goal, even for students with LDs, where the results on inclusion efforts have not been associated with improved academic outcomes (see Chapter 5). To illustrate, consider Figure 11.1, which shows the relation of NAEP reading scores at basic and proficient levels for students with disabilities. Consistent with the evidence reviewed in Chapter 5, there is a steady increase in the number of students with disabilities who spend most of their time in an inclusive environment. Reading proficiency scores are flat. As indicated in Chapter 2, intensive interventions are associated with improved academic outcomes for students with LDs and are a necessary component of successful treatment plans. Many schools perceive that providing a Tier 2-type intervention is adequate; however, many students do not respond adequately to these less intensive interventions and require more extensive and intensive treatments. For these students, Tier 3 intensive interventions that are remedial and designed with a specific goal of improving their academic outcomes with accelerated gains is necessary. Often, Tier 3 intensive interventions need to be undertaken in a timelier manner than occurs even when schools have MTSS frameworks in place. It is not always necessary for students to pass through all of the intervention tiers before special education services are provided; some children are so far behind that they should move immediately to Tier 3 or some other intensive intervention (Vaughn & Fletcher, 2012).

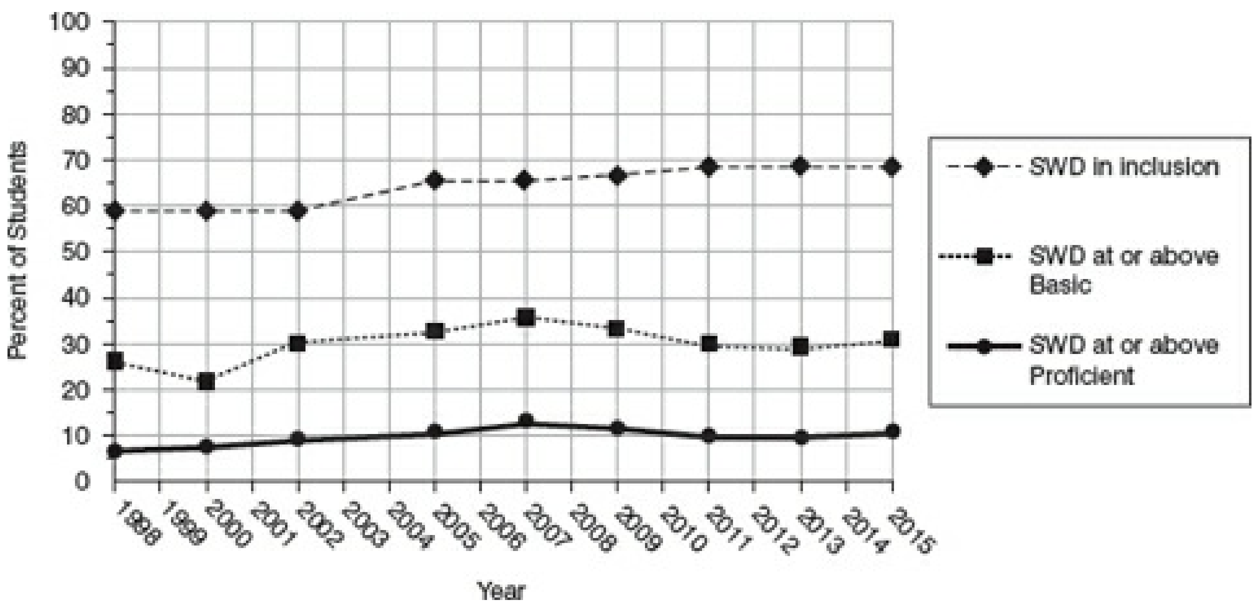

FIGURE 11.1. Reading proficiency levels of students with disabilities (SWD) on the NAEP in relation to the number of students with disabilities in inclusive placements. Reading scores are flat as the number of students with disabilities in inclusion rises. Figure constructed from different public data sources by Douglas Fuchs and provided as a courtesy.

Within Tier 3 (intensive intervention), the technology currently available for progress monitoring still requires widespread adoption to provide the data necessary to individualize intensive interventions, including special education, for students with LD and to provide accountability by iteratively monitoring responsiveness and adjusting instruction (L. S. Fuchs & D. Fuchs, 1998). Implemented effectively, progress monitoring for students with identified LDs promises to have more impact than any other single scaling component because it provides immediate and ongoing feedback on student progress in ways that allow for carefully, individually tailored instruction (D. Fuchs, L. S. Fuchs, & Vaughn, 2014b). Implementation of data-based individualization for intensive intervention is still often absent from schools and rarely appears in Individual Education Plans. Indeed, special education placement has often served to remove the student from accountability systems and less progress monitoring occurs after students are identified with disabilities than they might have received if they had not been identified (see Figure 3.8).

In contrast to these first four barriers, progress is more apparent for the fifth, sixth, and seventh barriers identified in Table 11.1. In terms of the fifth barrier, *insufficient consideration of multifaceted problems*, there is greater understanding of comorbidity at multiple levels of analysis, including identification, cognition, brain structure and function, and genetics. Interventions increasingly take into account comorbidity either through provision of programming that addresses two different problems (e.g., the programming for decoding and ADHD in Tamm et al., 2017) or through increased attention to the oral language problems apparent in many children with LDs in reading, math, and writing. Across these domains, there is better scientific understanding of core and Tier 2 instruction and progress across the board in enhancing instructional interventions for children with LDs. The progress in scientific understanding of the cognitive, neural, and genetic factors that underlie LDs is comparable, international in scope, and highly promising. Even so, additional research on interventions designed to explicitly address comorbid problems requiring intervention across academic domains is required to provide a stronger evidence base on the effectiveness of such interventions and to promote feasible delivery in the schools.

Delivering multicomponent interventions within integrated service delivery systems, such as MTSS, recognizes the complex presentation and multifaceted nature of LDs. For the sixth barrier, the *need for better integration of instructional methods,* more integrated programs have emerged and been evaluated. These advances include development of more complex, multicomponent intervention programs, which can be delivered in the context of integrated service delivery systems like MTSS. In reading, core classroom programs such as the approach developed by Connor and Morrison (2016) and the complex remedial programs developed by Lovett et al. (2013) and Wolf et al. (2002), in which cognitive strategy instruction for word identification and reading comprehension are linked with explicit decoding programs (Direct Instruction; see Chapter 6), are excellent examples. In math, the core programs described in Chapter 8 and interventions like Pirate Math (L. S. Fuchs et al., 2009), in which explicit instruction in word-problem solving is linked with instruction and practice in foundational arithmetic, are strong examples. There is also more integration across levels of instruction in core classroom instruction and Tier 2 supplemental instruction (e.g., Foorman et al., 2017; L. S. Fuchs et al., 2008b).

With respect to the seventh barrier, *lack of sufficient engagement and practice*, interventions developed over the past decade incorporate more opportunities for engagement and practice. In addition to cumulative review, specific types of practice are built into instruction, such as speeded practice in decoding and math interventions. In reading, there is much greater awareness at all levels of the importance of building automaticity through structured engagement where the reading of text is scaffolded to the reader's instructional level. In math and written expression, students typically work in the context of their problem areas and engage in additional exercises to increase engagement and practice.

It therefore appears that much progress has been made in understanding the complex nature of students with LD via multicomponent interventions that meet the complex learning needs of these students than in addressing the first four barriers, all of which are about implementation and program delivery: preventive intervention, with universal screening and progress monitoring, in the context of MTSS service delivery frameworks. While progress has been minimal addressing the first three barriers, in the United States, systemic efforts were mounted to address the first three barriers. The most noteworthy reform was Reading First, a joint effort by President George W. Bush and Congress, with strong bipartisan support. The focus was early identification of kindergarten, grade 1, and grade 2 children at risk for reading difficulties, through universal screening programs, with timely delivery of early reading preventive interventions and progress monitoring to determine the adequacy of response to intervention. The goal was to have every child reading on-level by grade 3, an ambitious and perhaps unrealistic goal, but one where considerable improvement in outcomes was supportable from a scientific perspective (Hess, 2008). In the 10 years since we wrote the first edition, Reading First has come and gone, but the problems in appropriately instructing young students to read have not. Initiatives are reemerging as dyslexia legislation at multiple state levels reflects grassroots advocacy concerns over the need to appropriately instruct and meet the needs of students with word-level reading problems. This involves an emphasis on early screening and intervention targeted to the needs of students with specific learning characteristics. The key to improving outcomes with students at risk for dyslexia remains enhanced general education instruction as outlined in Chapter 6, which should reduce the number of students identified with dyslexia, but also improve academic outcomes for all students.

RELIANCE ON CLINICAL EXPERIENCE AND CRAFT KNOWLEDGE OVER SCIENTIFIC EVIDENCE

This leads to the eighth barrier, where we see little progress: *reliance on scientific evidence for decision making in education*. As just noted, although federal and state governments invested resources in education practices focused on the first three barriers over the past 15 years, scaling scientific research in education has proven difficult. These efforts and the efforts of many scientists were oriented toward reading, which historically has been a focus of national and international concern, with numerous attempts to try and improve reading outcomes in public schools. Dyslexia, the most common reading LD, has been a major component of these efforts.

Scaling and Translating Educational Research

Why is scientific research in education difficult to scale? Many in education discuss the need to implement methods based on scientific research, but there is disagreement about what constitutes scientific research. This includes whether education research can be considered scientific and whether the quality of education research is adequate to be considered scientific. Many researchers have strong opinions and may not be completely objective when interpreting their own results. They may also not be clear in presenting their own results or identifying weaknesses or contradictory results across studies

(Hess, 2008). Hopefully, the synthesis of research presented in this book helps reduce concerns about the quality of educational research. For the first concern about the scientific nature of educational research, the National Research Council report on Scientific Research in Education (Shavelson & Towne, 2002) concluded that educational research was scientific and functioned like other domains of scientific inquiry:

> Scientific research, whether in education, physics, anthropology, molecular biology, or economics, is a continual process of rigorous reasoning supported by a dynamic interplay among methods, theories, and findings. It builds understandings in the form of models or theories that can be tested. Advances in scientific knowledge are achieved by the self-regulating norms of the scientific community over time, not, as sometimes believed, by the mechanistic application of a particular scientific method to a static set of questions. (p. 2)

The report identified six principles of scientific research that are universal in all domains and summarized in Table 11.2. All science is empirical, whether qualitative and observational or quantitative and experimental. Common across disciplines is an effort to rigorously and systematically observe the phenomena of interest, whether through a survey, observation, or experiment. In educational research, this empirical work can be quantitative, qualitative, or a combination of both approaches, but how the phenomena are observed and measured determines the degree to which generalizations can be made. Inferences about causality are on a continuum, with studies that manipulate the conditions under which children learn (e.g., randomized controlled trials) associated with the strongest inferences. Research that is observational or correlational does not support strong causal inferences, but may help describe or establish associations with mechanisms that are causal. But no single study is sufficient. As Table 11.2 shows, scientific research begins with the formulation of a question that can be addressed empirically. The question has to be linked to theory and it has to be evaluated using methods appropriate to the question. The findings need to connect with other studies *as part of a chain of reasoning* and be reproducible, which is most important for understanding causality. It must then be subjected to different forms of scrutiny, such as peer review and the use of empirical synthesis (meta-analysis) to explicitly address the potential bias associated with small and/or individual studies.

TABLE 11.2. Universal Principles of Scientific Research

1. Pose significant questions that can be investigated empirically.
2. Link research to relevant theory.
3. Use methods that permit direct investigation of the question.
4. Provide a coherent and explicit chain of reasoning.
5. Replicate and generalize across studies.
6. Disclose research to encourage professional scrutiny and critique.

Note. Based on Shavelson and Towne (2002).

We would submit that the research on LDs synthesized in this book has many of these characteristics (see Chapter 2). Although the evidence from research studies is incomplete and requires additional replication, we have considerable knowledge and the questions are important and empirical. A large body of research has been evaluated by methods that vary in rigor. The pieces connect into a larger body of evidence on LDs, academic learning, and cognitive learning, and include neuroscience. Studies have been scrutinized via peer review. Most importantly, the science has yielded findings in multiple domains, especially assessment and intervention, which can be implemented in nonresearch settings. Yet the application of the scientific findings presented in the preceding chapters has not been actualized in school settings for decision making about students, programs, and outcomes. Indeed, rarely are new programs evaluated, with the results used to drive school practice. Inadequate research-based decision making also plagues many commercial programs marketed to the education community, often in the name of science.

Four Hurdles to the Implementation of Scientific Research

Epistemology

Fletcher, Foorman, Denton, and Vaughn (2006) identified four hurdles to the implementation of scientific research in schools. The first involves *epistemological issues*, which have reduced the impact of scientific research on decision making in schools for many years (Carnine, 2000; Seidenberg, 2017). If research is cited, there is often a minimal attempt to evaluate its quality.

Decision making in public education is often driven by the marketing of reading instructional programs by individuals or commercial publishers of textbooks used in reading instruction. Schools often adopt reading programs based on program costs, without regard to evidence that a particular program or approach works. Researchers contribute to this problem because they ineffectively communicate research findings to practitioners, who cannot be expected to read peer-reviewed journals and have had little to no guidance on implementing research findings in their classrooms.

Another epistemological problem, thoroughly addressed by Seidenberg (2017), is that there are differences in the educational community about the value of scientific research and of different types of scientific research (Shavelson & Towne, 2002; Stanovich & Stanovich, 2003). Many researchers place an emphasis on descriptive, observational studies of teachers and students in classroom settings. These types of qualitative studies are appropriate and important, but are not capable of establishing the effectiveness of practices or particular programs and often provide an inadequate basis for broad applicability (Vaughn & Dammann, 2001). There is often a lack of attention to reliability and replication of the research findings, which is often justified because of the complex sociocultural milieu of schools and the richness of the individual student and teacher experience (St. Pierre, 2000). In this book, we focused on validated intervention programs established by at least one strong randomized trial. Often there are multiple trials, such as in the validation of Pirate Math (Chapter 8), the seven randomized trials supporting the core reading approach of Connor and Morrison (Chapter 6), and the multiple trials by Lovett et al. (2013) on PHAST. This validation does not ensure that the programs can be implemented, but the multiple trials do demonstrate that the interventions are effective across differing time points, settings, and children.

On the other hand, many quantitatively oriented researchers do not understand or value the types of research questions and studies that ethnographic and observational research addresses. This is due, in part, to a lack of recognition that qualitative research methods are essential for determining interactions within a classroom and classroom climate variables that may not be suitable for quantitative methods. Both quantitative and qualitative methods are equally scientific and important when they are applied appropriately to the research question at hand (Shavelson & Towne, 2002).

Education is certainly not alone in not embracing scientifically derived knowledge to guide clinical practice. This is reflected in the following quote about the origins of medical practice (Thomas, 1983):

> The history of the profession has never been a particularly attractive subject in professional education, and one reason for this is that it is so unrelievedly deplorable a story. For century after century all the way into the remote millennia of its origins, the profession got along by sheer guesswork and the crudest sort of empiricism. It is hard to conceive of a less scientific enterprise among human endeavors. Virtually anything that could be thought up for treatment was tried out at one time or another, and once tried, lasted decades or even centuries before giving it up. It was, in retrospect, the most frivolous and irresponsible kind of human experimentation, based on nothing but trial and error, and usually resulting in precisely that sequence. (pp. 2–3)

Thomas's observations on early medicine reflect some of the trial-and-error practices seen in education in the past and today. It is encouraging that medicine now is usually seen as connected to evidence-based practices. The journey from anecdote and untested assumptions to science was a difficult one in medicine, and it was not until the turn of the 20th century that a series of scientific breakthroughs altered the value that both physicians and their patients placed in scientific research. Laboratory and clinical research had exposed ineffective and often harmful outcomes of common medical treatments such as blistering, and bleeding and brought evidence-based practices, including antiseptic surgery, vaccination, and public sanitation to the forefront (Beck, 2004; Flexner, 1910). In 1881, President James Garfield died after an assassination attempt because of sepsis despite the availability of work by Joseph Lister, prominent and widely available at the time, on the importance of clean hands, sanitary practices, and other procedures now routinely used to prevent infection. This suggests that the timeline to widespread implementation of research-based practices is affected by social and cultural influences outside the realm of science, a point we return to below in discussing the attempts to implement research-based reading practices at scale.

Teachers and administrators have typically found it difficult to discriminate between research findings that are valid and replicable from those that are not, due to scientific jargon combined with a lack of training within colleges of education on the principles of scientific research evidence (Lyon & Chhabra, 2004). Early generations of reading research were weak, and educators have been frequently assaulted by the next “research-based” instructional “magic bullet” without having had the preparation necessary to distinguish between warranted claims of effectiveness and weak commercially motivated claims of efficacy. When such magic bullets fail, as they invariably do, many teachers and administrators lose trust in the capacity of research to inform their teaching (Hess, 2008). This should not be a surprise. It is difficult for teachers to make use of findings from education research because it has historically been of poor scientific quality, lacks the authority of valid evidence, and is not communicated clearly (Lyon & Esterline, 2007).

Inertia

The second hurdle to implementing scientific research on reading involves *investments in the current system and change for the sake of change.* These two types of inertia affect decision making in education in different ways. It is sometimes difficult to effect change because decision makers have investments in particular approaches that are based on ideology rather than scientific evidence. There is also insufficient accountability for results. Decision makers change frequently, and fads and trends in education emerge from interesting ideas that have not been evaluated. Shortcuts are often promoted that offer simple solutions to chronic, systemic problems. These approaches are appealing, but lack scientific support. Unfortunately, given the need to address how well programming works from a scientific perspective, effective programs are often abandoned because administrators or policies change. Administrators and policymakers in education at both the federal and state levels sometimes have an underdeveloped understanding of the role scientific evidence plays in education policy development and implementation, partly because scientists have not always done an adequate job of communicating its importance and helping to translate results. Although scientific research is generally recognized as critical to other policy environments (e.g., public health, agriculture, and commerce), education has typically been viewed as value-driven, and primary policy input has been obtained from politicians and diverse special interest groups rather than education scientists. Thus, education policies have often been forged within a political, rather than a scientific, context (Lyon & Esterline, 2007).

Financial Resources

The third hurdle to implementing scientific research on reading is *financial resources.* Implementing approaches based on scientific research may come with additional costs, particularly if educational decision makers are not willing to "let go" of current approaches and are trying to "layer on" research-based practices with existing practices that have little or no research basis. Financial resources in the last decade have decreased in many schools, partly due to the 2008 recession, but also because of political choices. This is especially apparent in the need to develop the capacity for providing evidence-based interventions early in a child's schooling. In many schools, the primary alternative to general classroom instruction is special education. As we have emphasized throughout this book, early intervention can prevent academic difficulties, reducing the need for later remediation for some children. Special education can play a critical role in addressing the needs of students with LDs; however, few would argue that special education resources should be allocated for all students who have not first been provided adequate academic instruction in general education settings. The hope is that prevention based on scientific research will decrease overidentification of students for special education, and thereby improve the effectiveness of general education and special education in meeting the needs of student at risk for or with identified with LDs.

Underdeveloped Research Base on Scaling

The fourth hurdle is an *underdeveloped research base on how to bring validated education practices to scale* (L. S. Fuchs & D. Fuchs, 1998; Denton et al., 2003). In scaling research, attempts are made to implement and translate a practice on a large-scale basis. Initial efforts focus on a limited number of schools and then move to increasingly larger units, such as an entire district. This process requires more systematic study to facilitate the translation of educational research into educational practice at scale (Elmore, 1996). Unfortunately, even the most effective programs can fail in the classroom because the fundamental principles of how to implement and scale a new and unfamiliar instructional program or initiative have not been considered. School administrators and teachers are frequently left to their own devices in determining the critical conditions under which implementation and scaling can be achieved with fidelity (L. S. Fuchs & D. Fuchs, 1998; Gersten, Vaughn, Deshler, & Schiller, 1997). Over the last 15 years, these issues have played out in attempts to improve reading skills in the United States. Despite significant investments in schools in the reading area, policies of the sort we have proposed were attempted and not scaled successfully. In the next section we consider Reading First as an example of an attempt to scale research-based reading instruction at a national level.

SCALING RESEARCH ON READING INSTRUCTION AT THE NATIONAL LEVEL: READING FIRST

Research on reading development, reading difficulties, and reading instruction has a long scientific history, and this content domain has been the primary focus for legislative actions due, in part, to the amount of scientific evidence, but also to the recognition that reading proficiency is the skill most fundamental to academic learning and success in school, to occupational and vocational opportunities, and to health outcomes (see Chapter 2). By the early 1990s, lawmakers and the nation at large were confronted with the evidence that an unacceptable number of children in the United States could not read proficiently (see the report of the National Commission on Excellence in Education (1983), *A Nation at Risk: The Imperative for Educational Reform*). The National Assessment of Educational Progress (NAEP) reports from 1992 (Mullis, Campbell, & Farstrup, 1993) and 1994 (Campbell, Donahue, Reese, & Phillips, 1996) showed that in the fourth grade alone, 37% of students nationally read below the basic level. Only 31% of students were reading at or above the level of proficiency; the achievement gap for African American and Hispanic and Latino students was disturbingly large. To be clear, it is not race or ethnicity that portended this significant underachievement in reading. The major culprit is poverty coupled with inadequate instruction. Minority students are overrepresented among impoverished families and are often attending schools with fewer resources and inadequate opportunities to learn. This trend in national statistics has not shown dramatic improvement in the ensuing decades.

Scientific evidence or the lack thereof is not the catalyst that typically leads to the attention that governments give to educational issues. Rather, as Song, Coggshall, and Miskel (2004) pointed out, for new policy directions and actions to occur, "a societal condition must capture policymakers' attention and be recognized as a problem that demands action" (p. 2.). Similarly, McDaniel, Sims, and Miskel (2001) reported, "The importance of improving the reading abilities of American school children has likely evolved into a permanent national concern" (p. 445). What emerged from these concerns was a series of task forces that produced controversy. The controversy ostensibly stemmed from two sources: belief that these reports represented advocacy for narrow, phonics-based instruction, as well as for explicit instruction in reading, and the reports' reliance on scientific inquiry and quantitative syntheses of research. We briefly discuss two task force reports, the policy initiative that followed (Reading First), and the lessons learned about scaling scientific research on reading to practice.

Consensus Reports

The National Research Council (NRC) report *Preventing Reading Difficulties in Young Children* (Snow et al., 1998) signaled an attempt to underscore the critical role of converging evidence in understanding reading development and preventing reading failure. The Preventing Reading Difficulties Committee was convened by the National Research Council of the National Academy of Sciences and supported by the National Institute of Child Health and Human Development (NICHD) and the U.S. Department of Education. A broad scientific consensus about the development of beginning reading and reading instruction was forged by highly respected researchers representing diverse perspectives. The conclusion reached by the committee is summarized in the following quote:

> All members agreed that reading should be defined as a process of getting meaning from print, using knowledge about the written alphabet and about the sound structure of oral

> language for the purpose of achieving understanding. All thus also agreed that early reading instruction should include direct teaching of information about sound/symbol relationships to children who do not know about them and that it must also maintain a focus on the communicative purposes and personal value of reading. (Snow et al., 1998, p. 6)

Despite the many strengths of the previous report, there was concern that although the committee had cogently summarized the skills that were critical for beginning reading proficiency, they did not focus on how those skills could be most effectively taught. Moreover, systematic review criteria were not established by the committee and the studies reviewed ranged in scientific quality. Following a series of briefings, Congress convened the National Reading Panel (National Institute of Child Health and Human Development, 2000) to provide the first evidence-based summary of the effectiveness of different reading instructional approaches and methods. This panel was charged with providing a report that "should present the panel's conclusions, an indication of the readiness for application in the classroom of the results of this research, and, if appropriate, a strategy for rapidly disseminating this information to facilitate effective reading instruction in the schools" (NICHD, 2000, p. 1).

The NRP report provided strong evidence that many struggling readers could learn to read if their teachers were adequately trained to implement effective scientifically validated instruction. Six NRP subgroups reviewed the studies that were considered by the panel to be methodologically sound: a subgroup to establish rigorous methodology followed by content subgroups, including alphabetics (phonemic awareness and phonics), fluency, comprehension, teacher education, and technology.

Both consensus reports made clear that a comprehensive, scientifically based approach to reading instruction is necessary if children were to learn to read efficiently. Although criticized as a phonics-only document (Allington, 2006), the NRP called for a comprehensive approach to reading instruction that emphasized the importance of systematic explicit instruction in decoding, vocabulary and comprehension, and fluency (see Stuebing et al., 2008, for an analysis and response to different critiques of the NRP).

From Consensus Reports to Educational Policy

The NRP report became the basis of the Reading First legislation initiated by President George W. Bush. It was included in the No Child Left Behind Act (2001). Reading First was a major education reform devoted to scaling research on beginning reading. The scope of this reform created substantial challenges, which included the sheer amount of required technical assistance the U.S. Department of Education had to provide each state and the low level of readiness for implementation of scientifically based assessments and instruction at the classroom level.

Indeed, it was more the rule than the exception that during the first 2 years of Reading First implementation in districts and schools, teachers were learning to understand, administer, and use the results of assessments to support instructional decision making. Moreover, as teachers were learning to apply these new concepts in their classrooms, they were also taking part in state reading academies to learn more about the foundations of scientifically based reading research (SBRR) in five areas of reading in kindergarten through grade 3. They were also expected to integrate core program instruction with additional interventions to meet individual student needs, implement new center activities, and group students for instruction based on need. It goes without saying that teachers were confronted with and asked to learn and implement an enormous amount of information on assessment and instructional concepts, methods, and strategies in a short period of time (Hess, 2008).

The mandatory evaluation of Reading First, called the Reading First Impact Study (RFIS; Gamse, Jacob, Horst, Boulay, & Unlu, 2008), reported mixed findings and less impact than had been anticipated. The popular press and education publications widely publicized the message that Reading First was ineffective because it did not improve reading comprehension (Glod, 2008; Manzo, 2008). The weaknesses in the report's methods were immediately identified (Reading First Advisory Committee, 2008), and its findings were eventually contradicted by analyses of state-level data (see Barbash, 2008; Bean, Draper, Turner, & Zigmond, 2010; Beck, 2010; Carlisle, Cortina, & Zeng, 2010; Center for Educational Policy, 2008; Dole, Hosp, Nelson, & Hosp, 2010; Foorman, Petscher, Lefsky, & Toste, 2010). Nonetheless, funding was significantly reduced in 2008, and the program was eliminated after 6 years of implementation.

This RFIS study examined the impact of the Reading First funding in 17 school districts across 12 states and in one statewide program, all selected *after* Reading First was initiated. As the authors of the study pointed out, the study schools were not a national probability sample but did share many characteristics with the national population of schools receiving Reading First grants. The evaluation compared outcomes for a group of 125 schools that were selected to receive a Reading First grant with a group of 123 schools from the same districts that did not receive a grant. The study employed a regression discontinuity design that took advantage of the fact that school districts and states rank-ordered schools on a set of independent criteria to choose those that would qualify for funding. This design was necessary because schools were already implementing Reading First and a randomized control design was not possible.

The results indicated that on average, Reading First increased the amount of time that teachers spent on the five essential components of reading instruction required by the program. Reading First did not meaningfully improve students' reading comprehension test scores when Reading First schools were compared to control schools, but there was a positive and significant impact on first-grade decoding. In contrast, when the data were disaggregated for sites awarded earlier and later in the 6-year cycle in late-award sites Reading First produced consistently stronger and positive effects on teachers' instructional practice *and* on students' reading comprehension test scores. Compared with early-award sites, the late-award sites tended to receive larger Reading First grants, serve lower-achieving students, and, but for the Reading First funding, would have spent less time on the five components of reading (as demonstrated by what happened in the non-Reading First schools). Importantly, however, the study design did not provide a reliable means of determining what might cause differences in Reading First impacts.

In response to the RFIS report, the Reading First Advisory Committee (2008) articulated several concerns with the study's methodology, sampling practices, delays, and conclusions drawn. Many schools were selected from the same district, and some districts came up with funding to support schools that were not eligible for Reading First grants, thereby blurring the distinction between Reading First and control schools. Gamse et al. (2008) noted that a limitation of the regression discontinuity analysis was that it did not include (by design) the most disadvantaged schools receiving Reading First funding.

In 2010, findings counter to the RFIS were published in four reports in the *Journal of Literacy Research.* These reports, assessing the impact of the Reading First initiative, were peer-reviewed with two commentaries by distinguished reading scholars who vetted the state evaluations for methodological robustness, the accuracy and the veracity of the obtained data, and the appropriateness of the conclusions provided in each of the papers. The four states that provided the reports were Florida, Michigan, Pennsylvania, and Utah and the two respondents to the reports were David Pearson from the University of California at Berkeley and Isabelle Beck from the University of Pittsburgh. Each of the states was geographically diverse albeit very similar in what was taught and what was assessed through the Reading First program.

In the Florida evaluation (Foorman et al., 2010; Torgesen, 2009), 5ive years of reading comprehension data in Florida Reading First schools were analyzed to address questions regarding student improvement, reduction in the achievement gap, efficacy of site visits to schools making no achievement gains, and effects of student mobility on growth in reading comprehension. Participants were 120,000 students (about 30,000 each in grades K–3) in the 318 schools in the first cohort of Florida Reading First from 2003 to 2008. Outcome measures were the reading comprehension scores on the Stanford Achievement Test and the Florida Comprehensive Assessment Test. The percentage of students on grade level (at or above the 40th percentile) increased, and the percentage of students at high risk (below the 20th percentile) decreased over the 5 years. Racial/ethnic minority, economically disadvantaged, and English language learner groups improved performance as well, but there was no evidence of narrowing the achievement gap. Reduction in risk for students with LDs was noteworthy, with declines in identification rates. Increased support to low-performing schools was associated with improved performance. Finally, there were significant reductions in growth in reading comprehension associated with leaving a Reading First school.

The evaluation in Pennsylvania (Bean et al., 2010) also revealed encouraging results. It examined student achievement outcomes for third graders in Reading First schools in Pennsylvania over the 5 years of implementation for the group as a whole, for disaggregated groups of third graders, and for third graders who received reading instruction in Reading First schools for 1, 2, and 3 years. Results indicated that third-grade students in these Reading First schools were making substantially more progress in third-grade reading than were students in other schools in Pennsylvania. There was an increase of nearly 24% of students performing at proficient or advanced levels over the 5 years, and the students in the below-basic range on the achievement measure declined nearly 18%. Further, the disaggregated data showed that the achievement gap, though not closed, was reduced for all groups of third graders. Finally, nearly 80% of the Reading First schools were successful in accomplishing the two goals of Reading First: an increase in the percentage of students at grade level, and a reduction in the percentage of students who were seriously below grade level.

Reading First data in Michigan were likewise encouraging. The report assessed improvements in the percentage of first-, second-, and third-grade students in Reading First schools who performed at or above grade level on a standardized test of reading comprehension. The study reported results for 140 schools that participated in the Reading First initiative between 2002 and 2008. Results showed that significant progress in improving the percentage of students reading at or above grade level in grades 1 and 2, but not in grade 3. While not the primary focus of the study, additional analyses showed significant decreases in the percentage of students reading substantially below grade level in Reading First schools, particularly for first and second grades. The effects of poverty and student and teacher attrition significantly and negatively affected schools' progress in improving reading comprehension achievement. The results suggest that the Reading First initiative in Michigan was effective in improving the reading comprehension of students in schools that were less hampered by problems associated with very high poverty.

The Utah report presented impact data using the results on the Utah Language Arts Criterion-Referenced Tests (CRTs; Dole et al., 2010). This evaluation found that Reading First schools made greater gains than comparison schools or the state totals in percentage of students who performed proficiently. Tests of whether there were significant differences in achievement as a function of the amount of time spent in Reading First showed that those who spent 3 years in Reading First averaged the highest proportion of students who scored at the proficient level.

In reviewing these studies, P. D. Pearson's (2010) observations were positive despite reservations for the Reading First program:

> The consistent message is that Reading First worked, at least as measured by the criteria used to evaluate its impact on school programs, professional development, teacher practices, and student achievement. By and large, teachers appreciated the various programs (all variations on the "Big Five" theme—phonemic awareness, phonics, fluency, vocabulary, and comprehension—of the National Reading Panel Report), liked and responded well to the staff development provided, and implemented the key components of the enabling legislation—the Big Five and a progress monitoring system in all four states plus a minimum time period for literacy instruction in Utah (270 minutes) and Michigan (90 minutes). Even more important, students appeared to benefit from the programmatic changes that were implemented. (P. D. Pearson, 2010, p. 101)

Beck (2010) also had a positive, but sobering review of the four studies:

> When I received the four evaluation studies . . . I assumed some folks who read this issue would compare the four state evaluation studies to the Reading First Impact Study Final Report (RFISFR) . . . so I reread that report. For those of you who have not looked at RFISFR, the two big findings were that there was (a) a positive and significant impact on first-grade decoding, the only grade at which it was measured; and (b) no effect on comprehension at grades one, two, and three. . . . What came to mind was, "Hmm . . . half-empty or half-full?" The half-empty interpretation would be something like, "Who cares that they can decode, if they can't comprehend what they read." The half-full interpretation would be that improved decoding puts a reader in position for improved comprehension if the other components (e.g., fluency, vocabulary) are in good shape or repaired.
>
> A serious problem with educational research is that when something does not work out completely, the field too often fails to build on what did work, or to figure out why a component might not have worked, fix it, and measure again. The half-full, half-empty notion turned out to be a lens through which I started to read the four studies. Upon completion, I rated each of them as half-full. My reason for the half-full categorization is that there were many positive trends, some of which were shown to be significant, that the field could build on and move the glass toward beyond half-full. (p. 94)

WHAT INTERFERES WITH OR FACILITATES IMPLEMENTATION?

The NRP conclusions about the characteristics of effective beginning reading programs have held up over time, but the progress hoped for in improving children's reading and preventing reading difficulties has been less than hoped for, especially in narrowing achievement gaps for minority students and students with LDs. As states revisit the difficulties faced by their students in reading and other academic domains, we address some barriers and facilitators that are relevant to implementing scientifically based research in any academic domain, whether that occurs at the level of school district, state, or nation. In this vein, our intent is not to discuss the merits or weaknesses of the Reading First program or its evaluation per se, but rather to use Reading First as a vehicle for addressing issues related to the scaling of scientific, research-based practices.

Enlisting relevant experts by engaging them from a wide range of professional disciplines and perspectives in the conceptualization and implementation of complex educational programs is critical. Having a diverse set of eyes looking at complex implementation, even if there is disagreement among them about concepts, methods, and philosophies of teaching and learning, produces stronger decisions. For example, for Reading First, engaging implementation scientists with experience in culture change, new initiatives, and very large-scale initiatives would have been productive. The failure to be more inclusive in the involvement of educational researchers with different perspectives also contributed to vulnerability in both implementation efforts and sustainability.

Sticking to the scientific evidence is a necessary ingredient to produce successful outcomes. When Reading First began, there were very few reading programs that met stringent effectiveness criteria. As such, the Congress and the U.S. Department of Education made the decision to soften the criteria and allow funding for programs that were based upon *principles of scientific evidence* rather than based on *evidence of program-specific effectiveness*. Given the softening of eligibility criteria, several programs from multiple publishers and vendors without sufficient evidence of effectiveness were adopted by states and implemented in local districts using federal funds. It is difficult to address the central question of whether effective programs remain effective at scale, but some of the programs supported via Reading First had no prior evidence of effectiveness even under ideal conditions (e.g., in an efficacy RCT); this problem was exacerbated because of weak local evaluations of efficacy. When taking efficacious programs to scale, it is important to implement and evaluate the effects of *all* programs.

This brings us to the point that *feasibility issues are related to implementation and scaling*. Congress and the U.S. Department of Education typically expect complex programs to be in place, in full operation, as soon as legislation is passed. When school districts invest in new programs, there is also the expectation of a fast roll-out. Yet, most initiatives that involve scaling efficacious programs involve learning numerous new concepts and procedures, many of which likely represent large departures from current practice (and sometimes philosophies). For example, while it is difficult to implement a new program, it is often even more difficult to move implementers away what they had been previously trained to do. A strong example is the persistence of multiple-cue approaches that permit students to guess at words or to use pictures for context when the development of phonics-based skills requires students to sound out words. In this example, the implementer is enabling previously taught strategies that are counterproductive for the students, who become as confused as the implementers! A similar example is persistence in teaching a student to sound out words when these skills are clearly developed and need to be practiced in text and not just isolation, and with comprehension as a goal. Thus, the provision of technical assistance and professional development is important. In Reading First, for example, ensuring that all involved at every level (state, district, school, classroom) understood the essential conditions needed to coordinate and implement a massive and unique program and to anticipate the need for customizing some features based on individual district and school characteristics would have enhanced opportunities for success (see Constas & Sternberg, 2006; Welch-Ross & Fasig, 2007). The Reading First program did not address administrative and organizational features of schools—factors that are critical in bringing about significant and sustained school-level gains in academic achievement. This is an oversight that must be corrected in future initiatives.

Related to this issue is the need for comprehensive professional development of teachers, a key piece of Reading First that was inconsistently implemented. Many professional development efforts did not concentrate sufficiently on vocabulary and reading comprehension instruction, particularly given the background of the children receiving Reading First programs. There was a failure to press home the "necessary but not sufficient" concept of explicit decoding instruction when talking about reading development, partly because the findings of the NRP report were misrepresented. While word-level skills and fluency are necessary and nonnegotiable, they clearly are not sufficient for ensuring children's understanding of what they read (see Chapter 7). Relatedly, there was less high-quality research on vocabulary development and reading comprehension instruction at that time—but what was available should have received greater emphasis in professional development. In some cases, the assessment strategies and programs used in Reading First schools also did not adequately address vocabulary and reading comprehension instruction. Within a data-based accountability system, there is a clear need to teach what is important and to measure what is taught.

Independent and timely program evaluation tied to the design of program implementation is also critical. Although the Reading First legislation provided specific language that an independent evaluation of every state program be carried out each year over a 6-year period, the Reading First national evaluation (published 6 years after states began receiving Reading First funding) was neither timely nor sufficient to obtain the necessary data that could inform modifications and adjustments to state and local Reading First implementation on a year-to-year basis. Evaluations of education initiatives whether at the local, state, or federal level, need to be built and implemented on schedule to consistently monitor and improve programs. There needs to be progress monitoring, if you will, for programs implemented at scale.

FUTURE DIRECTIONS AND CONCLUSIONS

In Chapter 11, we have focused on the scaling issues, outlining barriers to translation and using Reading First as an example of the difficulties with large-scale implementation of scientific research. Although these examples are drawn from the United States, similar disputes in other countries have occurred, notably the arguments over the efficacy of Reading Recovery in New Zealand (Chapman et al., 2015) and the dyslexia debate in the United Kingdom (Elliott & Grigorenko, 2014). Scaling and translation of research are fundamental issues in many branches of science, and education is no exception. *The two single-most important issues are the needs for early intervention and explicit instruction.* Pedagogical disputes aside, the evidence for the importance and key role of these two issues for students with LDs is overwhelming.

There is also a need for continued research on LDs. Table 11.3 provides 12 areas, all derived from the systematic review of research in this book, that represent future directions for research. At the top of the list is the scaling issue. We need to better understand and prioritize the issues involved in *scaling educational research*, which in many respects is a question for multidisciplinary research that includes expertise on sociocultural and political facilitators and barriers to change.

TABLE 11.3. Areas of Research That Need More Emphasis

1. Implementation and scaling of research.
2. Development of automaticity in the domains of reading, math, and written expression, especially in relation to intervention.
3. Scientific understanding of inadequate responders to intervention and the cognitive, neural, genetic, and instructional factors underlying individual differences in instructional response; development and valuation of intensive interventions.
4. Neurobiological research, emphasizing neural systems in SRCDs and written expression; candidate genes and epigenetics; longitudinal research.
5. Unpacking effective interventions: components and their organization, duration, and intensity; links of prevention and remediation; effects of prior intervention; and transfer and generalization.
6. Relations of oral language processing and impairments across LDs.
7. Role of visual and attention processes as shared factors across LDs.
8. Long-term effects of interventions across LDs.
9. Methods for predicting long-term individual student response to intervention.
10. Older school-age students and adults with LDs.
11. Compensatory adjuncts and accommodations.
12. Research on written expression LDs.

The second issue requiring attention is research on the *development of automaticity* in reading, math, and writing. As we showed in Chapter 10, automaticity is a general issue for all domains of LDs. What is not clear is whether this is a general issue common to all forms of LDs or whether it is specific to each academic domain. There have been suggestions that processing speed or WM represent domain-general deficits in LDs that might be linked to difficulties in developing automaticity, but how we address these issues within academic interventions requires additional study. It may also be important to consider automaticity and links to the neurobiological research on factors common to LDs at the level of the brain and genes. For example, the ventral occipitoparietal system is implicated in reading and math LDs. Because of its role in automatizing behavior, the cerebellum could also be a link to fluency deficits. The cognitive control system in the frontal regions could be related to automaticity deficits in math and reading comprehension because of their regulatory role. The evidence for generalist genes that affect all domains of LDs (and ADHD comorbidity) is also relevant in identifying common factors across LDs. This research may be fruitful and lead to even more interdisciplinary integration. Intervention programs that facilitate the development of automaticity are needed and should be embedded in current multicomponent programs. This will require an expectation for longer periods of intervention in remedial situations, although there is strong evidence suggesting that automaticity problems are alleviated with early identification and intervention. Another avenue for research in this area involves manipulating the conditions of learning and practice using principles from cognitive science (e.g., retrieval-based practice, optimal spacing of practice, interleaved practice) (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013) that are associated with increased processing speed, reduced forgetting, and resistance to interference in typical learners (Racsmány, Szőllősi, & Bencze, 2017).

The third issue is research on *inadequate responders*. Researchers need to study children who have not responded to quality core and supplemental instruction. We need to understand the basis for individual differences in instructional response based on cognitive, neural, and genetic data to fully understand biological constraints on learning. This research will lead to the development of intensive and more effective interventions for children with LDs, which are, at present, best developed for word-level disorders, but not well developed in other domains. We need to test the effects of programs and approaches that permit more sustained, individualized, intensive intervention for Tier 3 and special education.

Conducting more *neurobiological research,* the fourth issue, requires continued emphasis. In particular, a better understanding of the neural networks characteristics of SRCD and how they are different from WLRD is important. Although work on candidate genes in dyslexia is promising, molecular biology approaches to math and written expression are desperately needed. This would help address the comorbidity issue and the extent to which the characteristics of LDs are shared. More longitudinal research is also needed to map out behavior and brain trajectories over time (Black et al., 2017).

The fifth general need reiterates a call from the first edition, which is to *unpack effective interventions.* We need to understand why interventions are effective in terms of which component or procedure, or which combination or sequence of components or procedures, are critical to promote learning gains. In addition, more research is needed on the duration and intensity of intervention, especially at Tier 3; the effects of prior intervention, especially the link between preventive and remedial efforts; and teacher and contextual variables critical to effective implementation. Generalization and transfer are also important issues: between taught and untaught materials within a domain; between basic, foundational skills and higher-order skills; and between the intervention context and the classroom.

The sixth issue concerns *links of oral language*, with and without concurrent speech and language impairments, on comorbid forms of LDs. Reading relies on language, but to what extent does oral language impairment underlie other forms of LDs, especially SRCD, math problem solving, and composition. Oral language impairments may hold the key to understanding some comorbidities and inadequate responders.

By the same token, the seventh issue, the *role of visual and attention processes,* is poorly understood at the levels of behavior and the brain. This lack of understanding is not just from a behavioral comorbidity perspective, but also from a cognitive neuroscience perspective. The ventral system and the fusiform gyri seem to have important roles in reading and math: Are these common sources of shared difficulty that help explain shared characteristics of LDs? Similar areas of the brain also seem to be implicated in reading and math LDs, and also in ADHD. Research to tease apart these areas of possible overlap and differentiation are needed.

The eighth issue concerns examining the *long-term effects of interventions across LDs.* Research to date has focused almost exclusively on effects at the end of intervention, while studies are just now beginning to identify fade-out effects for many programs that work in the short term. In some respects, it is not that the effects really fade, but that the students stay at the level attained at the end of the intervention and then begin to fall behind again (Blachman et al., 2014; Vellutino et al., 2006). Future studies need to include a long-term longitudinal framework and test the effects of manipulations (e.g., different types and schedules of retrieval practice; "booster shots" of instruction after the summer gap) that may be associated with better long-term retention (Dunlosky et al., 2013) and higher levels of transfer to new learning.

Relatedly, researchers must address the ninth item, which is to identify *reliable methods for predicting long-term individual student response* to validated forms of intervention, so that empirically derived criteria can be employed to determine which students can be released from intervention to return with success to the general education program and which students, often presumed to have LDs, require more sustained, individualized, intensive intervention.

Research on *older students and adults with LDs* constitutes the tenth research need. Such research has exploded over the past decade, but we still understand little about effective practices for older students with LDs as well as possible "late-emerging" LDs. For adults, we do not have good evidence of effective interventions and know little about how adults adapt to LDs. We did not address this issue in the book because of insufficient research, especially for adults, although we did address the weak effects of intervention for adults in Chapter 6. The entire area of adult adaptation in well-characterized samples needs more research.

The related eleventh issue, also not addressed in this book, is how to use *compensatory adjuncts and accommodations* for older students. How well do assistive reading devices, computational aides, and computer interfaces for writing actually facilitate the adaptation of older students and adults with LDs? Similarly, while accommodations such as extended time, calculators, and test modifications are commonly implemented, there is scant evidence for their effectiveness (Lovett & Lewandowski, 2015).

Finally, the twelfth issue is *written expression,* which remains the least well understood form of LDs. There is especially a need to understand phenotypic variability and the extent to which written expression LDs occur as prototypes and in conjunction with other LDs and ADHD. This means devoting more careful attention to definition and classification issues. Emerging research on the link between transcription and composition needs to be maintained.

Research has and we hope will continue to flourish in ways that build our confidence about instructional approaches to identifying and treating individuals with LD. This is what motivates a second edition of this book. But additional research is still required. Systematic, empirical education research continues to emerge, including research clearly characterizing people who have LDs in reading, written language, and mathematics, and the need to integrate instructional design across special and general education on challenging standards. In the last edition, we expressed the hope and the belief that the practice of waiting for students to fail before identifying LDs and then placing them in educational environments not capable of closing the gap would begin to change, with a new and greater emphasis on preventing disabilities through effective general education.

Without wanting to in any way to downplay the improved knowledge we have about assessing and treating students with LD, we regret that there is little consistent evidence for a shift toward applying this knowledge to scaling effective practices in schools for this population. The key is to see students with LDs as a shared responsibility of general education and special education: all students are general education students first. Yet we must recognize the need to provide intensive interventions for these students when enhanced general education instruction is not sufficient. There may be a larger group of students requiring intensive interventions than those requiring special education per se. Parental involvement needs to be maximized and not just through the due process requirements of special education. But special education should be available for a smaller subset of inadequately responsive students. We suggest that special education could then harness the power of the legislation underlying special education to allow special educators to individually design and deliver more intensive interventions that cannot be provided in the classroom or through small-group instruction in general education.

The research reviewed in this book illustrates that many practices exist that would work with many students if effectively implemented. But research needs to be strongly linked to implementation. Scientific evidence is informing instruction, but implementation problems need to be understood more fully and addressed more effectively for the promises of this research to be realized. We should be optimistic about the future. Research is flourishing and there are no shortage of programs, practices, and procedures that can be implemented. There is progress in understanding and treating LDs in research and the integration across disciplines is a model for other areas of science. Many parents and advocates are highly aware of the research and work toward its implementation in schools. In the end, our society needs effective partnerships of scientists, educators, and parents united toward achieving the common goal of enhancing educational outcomes for all students.

References

Aaron, P. G., Kuchta, S., & Grapenthin, C. T. (1988). Is there a thing called dyslexia? *Annals of Dyslexia, 38*, 33–49.

Abbott, R. D., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology, 102*, 281–298.

Aboud, K. S., Bailey, S. K., Petrill, S. A., & Cutting, L. E. (2016). Comprehending text versus reading words in young readers with varying reading ability: Distinct patterns of functional connectivity from common processing hubs. *Developmental Science, 19*, 632–656.

Ackerman, P. T., Anhalt, J. M., & Dykman, R. A. (1986). Arithmetic automatization failure in children with attention and reading disorders: Associations and sequelae. *Journal of Learning Disabilities, 19*, 222–232.

Adams, G., & Carnine, D. (2003). Direct instruction. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 403–416). New York: Guilford Press.

Adams, G. L., & Engelmann, S. (1996). *Research on Direct Instruction: 25 years beyond DISTAR*. Portland, OR: Educational Achievement Systems.

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

Ahissar, M. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends in Cognitive Sciences, 11*, 458–465.

Ahissar, M., & Oganian, Y. (2008). Response to Ziegler: The anchor is in the details. *Trends in Cognitive Sciences, 12*, 245–246.

Ahmed, Y., Francis, D. J., York, M., Fletcher, J. M., Barnes, M. A., & Kulesz, P. (2016). Validation of the direct and inferential mediation (DIME) model of reading comprehension in grades 7–12. *Contemporary Educational Psychology, 45*, 68–82.

Ahmed, Y., Wagner, R. K., & Lopez, D. (2014). Developmental relations between reading and writing at the word, sentence, and text levels: A latent change score analysis. *Journal of Educational Psychology, 106*, 419–434.

Alarcon, M., DeFries J. C., Light, J. C., & Pennington, B. F. (1997). A twin study of mathematics disability. *Journal of Learning Disabilities, 30*, 617–623.

Allington, R. L. (2006). Reading lessons and federal policy making: An overview and introduction to the special issue. *Elementary School Journal, 107*, 3–15.

Al Otaiba, S., & Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective?: An experimental and longitudinal study. *Journal of Learning Disabilities, 39*, 414–431.

Altarelli, I., Leroy, F., Monzalvo, K., Fluss, J., Billard, C., Dehaene-Lambertz, G., . . . Ramus, F. (2014). Planum temporal asymmetry in developmental dyslexia: Revisiting an old question. *Human Brain*

Mapping, 35(12), 5717–5735.

Alves, R. A., Limpo, T., Fidalgo, R., Carvalhais, L., Pereira, L. Á., & Castro, S. L. (2016). The impact of promoting transcription on early text production: Effects on bursts and pauses, levels of written language, and writing performance. *Journal of Educational Psychology, 108*, 665–675.

American Academy of Pediatrics. (2009). Joint statement—Learning disabilities, dyslexia, and vision. *Pediatrics, 124*, 837–848.

American Academy of Pediatrics. (2011). Clinical practice guidelines for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics, 128*(5), 1–16.

American Association on Intellectual and Developmental Disabilities. (n.d.). Definition of intellectual disability. Retrieved from *http://aaidd.org/intellectual-disability/definition*.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2013a). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.

American Psychiatric Association. (2013b). DSM-5 fact sheets: Specific Learning Disorder. Retrieved from *www.psychiatry.org/psychiatrists/practice/dsm/educational-resources/dsm-5-fact-sheets*.

Andersson, U. (2010). Skill development in different components of arithmetic and basic cognitive functions: Findings from a 3-year longitudinal study of children with different types of learning difficulties. *Journal of Educational Psychology, 102*, 115–134.

Ansari, D., & Lyons, D. M. (2016). Cognitive neuroscience and mathematics learning: How far have we come? Where do we need to go? *ZDM Mathematics Education, 48*, 379–383.

Apel, K., Wolter, J. A., & Masterson, J. J. (2006). Effects of phonotactic and orthotactic probabilities during fast mapping on 5-year-olds' learning to spell. *Developmental Neuropsychology, 29*, 21–42.

Araújo, S., Reis, A., Petersson, K. M., & Faísca, L. (2015). Rapid automatized naming and reading performance: A meta-analysis. *Journal of Educational Psychology, 107*(3), 868–883.

Arfé, B., Dockrell, J. E., & De Bernardi, B. (2016). The effect of language specific factors on early written composition: The role of spelling, oral language and text generation skills in a shallow orthography. *Reading and Writing, 29*, 501–527.

Arndt, E. J., & Foorman, B. R. (2010). Second graders as spellers: What types of errors are they making? *Assessment for Effective Intervention, 36*, 57–67.

Arnett, A. B., Pennington, B. F., Peterson, R. L., Willcutt, E. G., DeFries, J. C., & Olson, R. K. (2017). Explaining the sex difference in dyslexia. *Journal of Child Psychology and Psychiatry, 58*(6), 719–727.

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics, 24*, 621–635.

Arsalidou, M., & Taylor, M. J. (2011). Is 2 + 2 = 4?: Meta-analyses of brain areas needed for numbers and calculations. *NeuroImage, 54*(3), 2382–2393.

Arter, J. A., & Jenkins, R. (1979). Differential diagnosis: Prescriptive teaching: A critical appraisal, *Review of Educational Research, 49*, 517–555.

Ashkenazi, S., Black, J. M., Abrams, D. A., Hoeft, F., & Menon, V. (2013). Neurobiological underpinnings of math and reading learning disabilities. *Journal of Learning Disabilities, 46*(6), 549–569.

Aylward, E. H., Richards, T. L., Berninger, V. W., Nagy, W. E., Field, K. M., Grimme, A. C., . . . Cramer, S. C. (2003). Instructional treatment associated with changes in brain activation in children with dyslexia. *Neurology, 22*, 212–219.

Babayiĝit, S., & Stainthorp, R. (2010). Component processes of early reading, spelling, and narrative writing skills in Turkish: A longitudinal study. *Reading and Writing, 23*, 539–568.

Bach, S., Richardson, U., Brandeis, D., Martin, E., & Brem, S. (2013). Print-specific multimodal brain activation in kindergarten improves prediction of reading skills in second grade. *NeuroImage, 82*, 605–615.

Badian, N. A. (1999). Reading disability defined as a discrepancy between listening and reading comprehension: A longitudinal study of stability, gender differences, and prevalence. *Journal of Learning Disabilities, 32*, 138–148.

Bailey, S., Hoeft, F., Aboud, K., & Cutting, L. (2016). Anomalous gray matter patterns in specific reading comprehension deficit are independent of dyslexia. *Annals of Dyslexia, 66*, 256–274.

Balu, R., Zhu, P., Doolittle, F., Schiller, E., Jenkins, J., & Gersten, R. (2015). *Evaluation of response to intervention practices for elementary school reading*. Washington, DC: National Center for Education Evaluation and Regional Assistance.

Barbaresi, W. J., Katusic, S. K., Colligan, R. C., Weaver, A. L., & Jacobsen, S. J. (2005). Math learning disorder: Incidence in a population-based birth cohort, 1976–82, Rochester, Minn. *Ambulatory Pediatrics, 5*, 281–289.

Barbash, S. (2008). The Reading First controversy: Promises and perils of federal leadership. *Reading Next, 8*, 47–53.

Barkley, R. A. (2015). *Attention-deficit hyperactivity disorder, fourth edition: A handbook for diagnosis and treatment*. New York: Guilford Press.

Barnes, M. A. (2015). What do models of reading comprehension and its development have to contribute to a science of comprehension practice and assessment for adolescent students? In K. L. Santi & D. K. Reed (Eds.), *Improving comprehension of middle and high school students* (pp. 1–18) New York: Springer.

Barnes, M. A., Ahmed, Y., Barth, A., & Francis, D. J. (2015). The relation of knowledge–text integration processes and reading comprehension in 7th- to 12th-grade students. *Scientific Studies of Reading, 19*(4), 253–272.

Barnes, M. A., Dennis, M., & Haefele-Kalvaitis, J. (1996). The effects of knowledge availability and knowledge accessibility on coherence and elaborative inferencing from six to fifteen years of age. *Journal of Experimental Child Psychology, 61*, 216–241.

Barnes, M. A., Klein, A., Swank, P., Starkey, P., McCandliss, B., Flynn, K., . . . Roberts, G. (2016a). Effects of tutorial interventions in mathematics and attention for low-performing preschool children. *Journal of Research on Educational Effectiveness, 9*, 577–606.

Barnes, M. A., Raghubar, K. P., English, K. P., Agostino, A., Taylor, H., Landry, S. B., . . . Smith-Chant, B. (2011). Mathematical skills in 3- and 5-year olds with spina bifida and their typically developing peers: A longitudinal approach. *Journal of the International Neuropsychological Society, 17*, 385–392.

Barnes, M. A., Raghubar, K. P., English, L., Williams, J. M., Taylor, H., & Landry, S. (2014). Longitudinal mediators of achievement in mathematics and reading in typical and atypical development. *Journal of Experimental Child Psychology, 119*, 1–16.

Barnes, M. A., Stuebing, K. K., Fletcher, J. M., Barth, A. E., & Francis, D. J. (2016b). Cognitive difficulties in struggling comprehenders and their relation to reading comprehension: A comparison of group selection and regression-based models. *Journal of Research on Educational Effectiveness, 9*(2), 153–172.

Barquero, L. A., Davis, N., & Cutting, L. E. (2014). Neuroimaging of reading intervention: A systematic review and activation likelihood estimate meta-analysis. *PLOS ONE, 9*(1), e83668.

Barrett, B. T. (2009). A critical evaluation of the evidence supporting the practice of behavioural vision therapy. *Ophthalmic and Physiological Optics, 29*, 4–25.

Barth, A. E., Barnes, M., Francis, D., Vaughn, S., & York, M. (2015). Inferential processing among adequate and struggling adolescent comprehenders and relations to reading comprehension. *Reading and Writing, 28*, 587–609.

Barth, A. E., Denton, C. A., Stuebing, K. K., Fletcher, J. M., Cirino, P. T., Francis, D. J., & Vaughn, S. (2010). A test of the cerebellar hypothesis of dyslexia in adequate and inadequate responders to reading intervention. *Journal of the International Neuropsychological Society, 16*, 526–536.

Barth, A. E., Stuebing, K. K., Anthony, J. L., Denton, C. A., Mathes, P. G., Fletcher, J. M., & Francis, D. J. (2008). Agreement among response to intervention criteria for identifying responder status. *Learning and Individual Differences, 18*, 296–307.

Barth, A. E., Stuebing, K. K., Fletcher, J. M., Cirino, P. T., Romain, M., Francis, D. J., & Vaughn, S. (2012). Reliability and validity of oral reading fluency median and mean scores among middle grade readers when using equated texts. *Reading Psychology, 33*, 133–161.

Bates, T. C., Castles, A., Coltheart, M., Gillespie, N., Wright, M., & Martin, N. G. (2004). Behaviour genetic analyses of reading and spelling: A component processes approach. *Australian Journal of Psychology, 56*, 115–126.

Bean, K., Draper, I., Turner, G., & Zigmond, N. (2010). Reading First in Pennsylvania: Achievement findings after five years. *Journal of Literacy Research, 42*, 5–26.

Beattie, R. L., Lu, Z. L., & Manis, F. R. (2011). Dyslexic adults can learn from repeated stimulus presentation but have difficulties in excluding external noise. *PLOS ONE, 6*(11), e27893.

Beck, A. (2004). The Flexner Report and the standardization of American medicine. *Journal of the American Medical Association, 291*, 2139–2140.

Beck, I. (2010). Half full or half empty. *Journal of Literacy Research, 42*, 94–99.

Benítez-Burraco, A. (2010). Neurobiology and neurogenetics of dyslexia. *Neurología, 25*, 563–581.

Benner, G. J., Nelson, J. R., Ralston, N. C., & Mooney, P. (2010). A meta-analysis of the effects of reading instruction on the reading skills of students with or at risk of behavioral disorders. *Behavioral Disorders, 35*, 86–102.

Benton, A. L. (1975). Developmental dyslexia: Neurological aspects. *Advances in Neurology, 7*, 1–47.

Bentum, K. E., & Aaron, P. G. (2003). Does reading instruction in learning disability resource rooms really work?: A longitudinal study. *Reading Psychology, 24*, 361–382.

Bergmann, J., & Wimmer, H. (2008). A dual-route perspective on poor reading in a regular orthography: Evidence from phonological and orthographic lexical decisions. *Cognitive Neuropsychology, 25*, 653–676.

Berkeley, S., Scruggs, T. E., & Mastropieri, M. A. (2010). Reading comprehension instruction for students with learning disabilities, 1995–2006: A meta-analysis. *Remedial and Special Education, 31*(6), 423–436.

Berninger, V. W. (1994). *Reading and writing acquisition: A developmental neuropsychological perspective* (Brown & Benchmark's developmental psychology series). Madison, WI: Brown & Benchmark.

Berninger, V. W. (2004). Understanding the graphia in developmental dysgraphia: A developmental neuropsychological perspective for disorders in producing written language. In D. Dewey & D. Tupper (Eds.), *Developmental motor disorders: A neuropsychological perspective* (pp. 189–233). New York: Guilford Press.

Berninger, V. W., & Abbott, R. D. (2010). Listening comprehension, oral expression, reading comprehension, and written expression: Related yet unique language systems in grades 1, 3, 5, and 7. *Journal of Educational Psychology, 102*, 635–651.

Berninger, V. W., Abbott, R. D., Augsburger, A., & Garcia, N. (2009a). Comparison of pen and keyboard transcription modes in children with and without learning disabilities. *Learning Disability Quarterly, 32*, 123–141.

Berninger, V. W., Abbott, R. D., Jones, J., Wolf, B. J., Gould, L., Anderson-Youngstrom, M., . . . Apel, K. (2006a). Early development of language by hand: Composing, reading, listening, and speaking connections; three letter-writing modes; and fast mapping in spelling. *Developmental Neuropsychology, 29*, 61–92.

Berninger, V. W., & Amtmann, D. (2003). Preventing written expression disabilities through early and continuing assessment and intervention for handwriting and/or spelling problems: Research into practice. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 345–363). New York: Guilford Press.

Berninger, V. W., & Chanquoy, L. (2012). What writing is and how it changes across early and middle childhood development. In E. L Grigorenko, E. Mambrino, & D. D. Preiss (Eds.), *Writing: A mosaic of new perspectives* (pp. 65–84). New York: Psychology Press.

Berninger, V. W., & Fuller, F. (1992). Gender differences in orthographic, verbal, and compositional fluency: Implications for assessing writing disabilities in primary grade children. *Journal of School Psychology, 30*, 363–382.

Berninger, V. W., & Hart, T. (1992). A developmental neuropsychological perspective for reading and writing acquisition. *Educational Psychologist, 27*, 415–434.

Berninger, V. W., Nagy, W., & Beers, S. (2011). Child writers' construction and reconstruction of single sentences and construction of multisentence texts: Contributions of syntax and transcription to translation. *Reading and Writing, 24*, 151–182.

Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Gender differences in severity of writing and reading disabilities. *Journal of School Psychology, 46*, 151–172.

Berninger, V. W., Richards, T. L., Stock, P. S., Abbott, R. D., Trivedi, P. A., Altemeier, L. E., & Hayes, J. R. (2009b). fMRI activation related to nature of ideas generated and differences between good and poor writers during idea generation. *BJEP Monograph Series II, Number 6, Teaching and Learning Writing, 77*, 77–93.

Berninger, V. W., Rutberg, J. E., Abbott, R. D., Garcia, N., Anderson-Youngstrom, M., Brooks, A., . . . Fulton, C., (2006b). Tier 1 and Tier 2 early intervention for handwriting and composing. *Journal of School Psychology, 44*, 3–30.

Berninger, V. W., Vaughan, K. B., Abbott, R. D., Abbott, S. P., Rogan, L. W., Brooks, A., . . . Graham, S. (1997). Treatment of handwriting problems in beginning writers: Transfer from handwriting to composition. *Journal of Educational Psychology, 89*, 652–674.

Berninger, V. W., Vaughan, K., Abbott, R. D., Begay, K., Coleman, K. B., Curtin, G., . . . Graham, S. (2002). Teaching spelling and composition alone and together: Implications for the simple view of writing. *Journal of Educational Psychology, 94*, 291–304.

Berninger, V. W., & Winn, W. (2006). Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. In *Handbook of writing research* (pp. 6–114). New York: Guilford Press.

Besner, D., Risko, E. F., Stolz, J. A., White, D., Reynolds, M., O'Malley, S., . . . Robidoux, S. (2016). Varieties of attention: Their roles in visual word identification. *Current Directions in Psychological Science, 25*, 162–168.

Biancarosa, C., & Snow, C. E. (2006). *Reading next—A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York* (2nd ed.). Washington, DC: Alliance for Excellent Education.

Birsh, J. (Ed.). (1999). *Multi-sensory teaching of basic language skills*. Baltimore: Brookes.

Bishop, D. V. M. (2007). Curing dyslexia and attention-deficit hyperactivity disorder by training motor coordination: Miracle or cure? *Journal of Pediatrics and Child Health, 43*, 653–655.

Bishop, D. V., & Adams, C. (1992). Comprehension problems in children with specific language impairment: Literal and inferential meaning. *Journal of Speech, Language, and Hearing Research, 35*(1), 119–129.

Bishop, D. V. M., & Clarkson, B. (2003). Written language as a window into residual language deficits: A study of children with persistent and residual speech and language impairments. *Cortex, 39*, 215–237.

Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin, 130*, 858–886.

Blachman, B. A., Schatschneider, C., Fletcher, J. M., Francis, D. J., Clonan, S. M., Shaywitz, B. A., . . . Shaywitz, S. E. (2004). Effects of intensive reading remediation for second and third graders and a 1-year follow-up. *Journal of Educational Psychology, 96*, 444–461.

Blachman, B. A., Schatschneider, C., Fletcher, J. M., Murray, M. S., Munger, K. A., & Vaughn, M. G. (2014). Intensive reading remediation in grade 2 or 3: Are there effects a decade later? *Journal of Educational Psychology, 106*, 46–57.

Black, J. M., Xia, Z., & Hoeft, F. (2017). Neurobiological bases of reading disorder: Part I. The importance of developmental considerations in typical and atypical readers. *Language and Linguistics Compass, 11*(10), e12252.

Blashfield, R. K. (1993). Models of classification as related to a taxonomy of learning disabilities. In G. R. Lyon, D. B. Gray, J. F. Kavanagh, & N. A. Krasnegor (Eds.), *Better understanding learning disabilities: New views from research and their implications for education and public policies* (pp. 17–26). Baltimore: Brookes.

Boden, C., & Giaschi, D. (2007). M-stream deficits and reading-related visual processes in developmental dyslexia. *Psychological Bulletin, 133*, 346–366.

Bornstein, M. H., Hahn, C. S., Putnick, D. L., & Suwalsky, J. T. D. (2014). Stability of core language skill from early childhood to adolescence: A latent variable approach. *Child Development, 85*(4), 1348–1356.

Bos, C. S., & Anders, P. L. (1990). Interactive teaching and learning: Instructional practices for teaching content and strategic knowledge. In T. E. Scruggs & B. Y. L. Wong (Eds.), *Intervention research in learning disabilities* (pp. 161–185). New York: Springer-Verlag.

Boulay, B., Goodson, B., Frye, M., Blocklin, M., & Price, C. (2015). *Summary of research generated by Striving Readers on the effectiveness of interventions for struggling adolescent readers* (NCEE 2016-4001). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Bourassa, D. C., & Treiman, R. (2008). Morphological constancy in spelling: A comparison of children with dyslexia and typically developing children. *Dyslexia, 14*, 155–169.

Bradley, R., Danielson, L., & Hallahan, D. P. (Eds.). (2002). *Identification of learning disabilities: Research to practice*. Mahwah, NJ: Erlbaum.

Brankaer, C., Ghesquière, P., & De Smedt, B. (2014). Children's mapping between non-symbolic and symbolic numerical magnitudes and its association with timed and untimed tests of mathematics achievement. *PLOS ONE, 9*(4), e93565.

Brasseur-Hock, I. F., Hock, M. F., Kieffer, M. J., Biancarosa, G., & Deshler, D. D. (2011). Adolescent struggling readers in urban schools: Results of a latent class analysis. *Learning and Individual Differences, 21*(4), 438–452.

Breier, J. I., Fletcher, J. M., Denton, C., & Gray, L. C. (2004). Categorical perception of speech stimuli in children at risk for reading difficulty. *Journal of Experimental Child Psychology, 88*, 152–170.

Breier, J. I., Fletcher, J. M., Foorman, B. R., & Gray, L. C. (2002). Perception of speech and nonspeech stimuli by children with and without reading disability and attention deficit hyperactivity disorder. *Journal of Experimental Child Psychology, 82*, 226–250.

Breier, J. I., Simos, P. G., Fletcher, J. M., Castillo, E. M., Zhang, W., & Papanicolaou, A. C. (2003). Abnormal activation of temporoparietal language areas during phonetic analysis in children with dyslexia. *Neuropsychology, 17*, 610–621.

Brem, S., Bach, S., Kucian, K., Kujala, J. V., Guttorm, T. K., Martin, E., . . . Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter–speech sound correspondences. *Proceedings of the National Academy of Sciences of the USA, 107*, 7939–7944.

Brown Waeschle, J. S., Schatschneider, C., Maner, J. K., Ahmed, Y., & Wagner, R. K. (2011). Examining agreement and longitudinal stability among traditional and RTI-based definitions of reading disability using the affected-status agreement statistic. *Journal of Learning Disabilities, 44*, 296–307.

Bruck, M. (1987). The adult outcomes of children with learning disabilities. *Annals of Dyslexia, 37*, 252–263.

Bryant, B. R., & Wiederholt, J. L. (2001). *Gray Oral Reading Tests* (4th ed.). Austin, TX: PRO-ED.

Bryant, D. P., Bryant, B. R., Roberts, G., Vaughn, S., Pfannenstiel, K. H., Porterfield, J., & Gersten, R. (2011). Early numeracy intervention program for first-grade students with mathematics difficulties. *Exceptional Children, 78*, 7–23.

Bugden, S., & Ansari, D. (2016). Probing the nature of deficits in the "Approximate Number System" in children with persistent developmental dyscalculia. *Developmental Science, 19*, 817–833.

Bugden, S., Price, G. R., McLean, D. A., & Ansari, D. (2012). The role of the left intraparietal sulcus in the relationship between symbolic number processing and children's arithmetic competence. *Developmental Cognitive Neuroscience, 2*(4), 448–457.

Burns, M. K., Petersen-Brown, S., Haegele, K., Rodriguez, M., Schmitt, B., Cooper, M., . . . VanDerHeyden, A. M. (2016). Meta-analysis of academic interventions derived from neuropsychological data. *School Psychology Quarterly, 31*, 28–42.

Burt, C. (1937). *The backward child*. London: University of London Press.

Butler, F. M., Miller, S. P., Crehan, K., Babbitt, B., & Pierce, T. (2003). Fraction instruction for students with mathematics disabilities: Comparing two teaching sequences. *Learning Disabilities Research and Practice, 18*, 99–111.

Butterworth, B. (2005). Developmental dyscalculia. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 455–467). New York: Psychology Press.

Cain, K., & Barnes, M. A. (2017). Reading comprehension: What develops and when? In K. Cain, & D. L. Compton, & R. K. Parrila (Eds.), *Theories of reading development* (pp. 257–282). Amsterdam: Benjamins.

Cain, K., & Oakhill, J. V. (1999). Inference making and its relation to comprehension failure in young children. *Reading and Writing, 11*, 489–503.

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and the relation to knowledge. *Memory and Cognition, 29*, 850–859.

Cain, K., Oakhill, J. V., & Bryant, P. (2000). Phonological skills and comprehension failures: A test of the phonological processing deficits hypothesis. *Reading and Writing, 13*, 31–56.

Cain, K., Oakhill, J. V., & Bryant, P. (2004a). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*, 31–42.

Cain, K., Oakhill, J. V., & Lemmon, K. (200b). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology, 96*, 671–681.

Calkins, L. (2003). *Launching the writing workshop*. Portsmouth, NH: Heinemann.

Camilli, G., Wolfe, P. M., & Smith, M. L. (2006). Meta-analysis and reading policy: Perspectives on teaching children to read. *Elementary School Journal, 107*, 27–36.

Campbell, J. R., Donahue, P. L., Reese, C. M., & Phillips, G. (1996). *NAEP 1994 Reading Report Card: Findings from the National Assessment of Educational Progress*. Washington, DC: National Center for Education Statistics.

Canning, P. M., Orr, R. R., & Rourke, B. P. (1980). Sex differences in the perceptual, visual–motor, linguistic and concept-formation abilities of retarded readers? *Journal of Learning Disabilities, 13*, 563–567.

Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G., & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science, 24*(8), 1398–1407.

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., . . . Seidlová-Málková, G. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science, 23*, 678–686.

Caravolas, M., Volin, J., & Hulme, C. (2005). Phoneme awareness is a key component of alphabetic literacy skills in consistent and inconsistent orthographies: Evidence from Czech and English children. *Journal of Experimental Child Psychology, 92*, 107–139.

Carey, S. (2004). Bootstrapping and the origin of concepts. *Daedalus, 133*, 59–68.

Carlisle, I. E., Cortina, K. S., & Zeng, J. (2010). Reading achievement in Reading First schools in Michigan. *Journal of Literacy Research, 42*, 49–70.

Carnine, D. (2000). *Why education experts resist effective practices (Report of the Thomas B. Fordham Foundation)*. Washington, DC: Thomas B. Fordham Foundation.

Carr, M., & Aleexev, N. (2011). Fluency, accuracy, and gender predict developmental trajectories of arithmetic strategies. *Journal of Educational Psychology, 103*, 617–631.

Carretti, B., Caldarola, N., Tencati, C., & Cornoldi, C. (2014). Improving reading comprehension in reading and listening settings: The effect of two training programmes focusing on metacognition and working memory. *British Journal of Educational Psychology, 84*, 194–210.

Carroll, J. M., Maughan, B., Goodman, R., & Meltzer, H. (2005). Literacy difficulties and psychiatric disorders: Evidence for comorbidity. *Journal of Child Psychology and Psychiatry, 46*, 524–532.

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition, 47*, 149–180.

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

Castro-Caldas, A., Petersson, K. M., Reis, A., Stone-Elander, S., & Ingvar, M. (1998). The illiterate brain: Learning to read and write during childhood influences the functional organization of the adult brain. *Brain, 121*, 1053–1063.

Cattinelli, I., Borghese, N. A., Gallucci, M., & Paulesu, E. (2013). Reading the reading brain: A new meta-analysis of functional imaging data on reading. *Journal of Neurolinguistics, 26*, 214–238.

Catts, H. W., Adlof, S. M., Hogan, T. P., & Weismer, S. E. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language, and Hearing Research, 48*, 1378–1396.

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research, 49*, 278–293.

Catts, H. W., Compton, D., Tomblin, B., & Bridges, M. S. (2012). Prevalence and nature of late-emerging poor readers. *Journal of Educational Psychology, 10*, 166–181.

Catts, H. W., Gillispie, M., Leonard, L. B., Kail, R. V., & Miller, C. A. (2002). The role of speed processing, rapid naming, and phonological awareness in reading achievement. *Journal of Learning Disabilities, 35*, 510–525.

Catts, H. W., Nielsen, D. C., Bridges, M. S., & Liu, Y. S. (2016). Early identification of reading comprehension difficulties. *Journal of Learning Disabilities, 49*, 451–465.

Center for Educational Policy. (2007). *Reading First: Locally appreciated, nationally troubled*. Washington, DC: Author.

Center for Effective Reading Instruction. (2018). Knowledge and practice standards. Retrieved June 5, 2018, from *https://effectivereading.org/knowledge-and-practice-standards/*.

Chan, D. W., Ho, C. S. H., Tsang, S. M., Lee, S. H., & Chung, K. K. (2006). Exploring the reading–writing connection in Chinese children with dyslexia in Hong Kong. *Reading and Writing, 19*, 543–561.

Chang, B. S., Katzir, T., Liu, T., Corriveau, K., Barzillai, M., Apse, K. A., . . . Walsh, C. A. (2007). A structural basis for reading fluency: White matter defects in a genetic brain malformation. *Neurology, 69*, 2146–2154.

Chapman, J. W., Greaney, K. T., & Tunmer, W. E. (2015). Is Reading Recovery an effective early literacy intervention programme for children who most need literacy supports? In W. E. Tunmer & J. W. Chapman (Eds.), *Excellence and equity in literacy education: The case of New Zealand* (pp. 41–70). London: Palgrave Macmillan.

Chapman, J., & Tunmer, W. (2016). Is Reading Recovery an effective intervention for students with reading difficulties?: A critique of the i3 Scale-Up Study. *Reading Psychology, 37*, 1025–1042.

Chard, D. J., Vaughn, S., & Tyler, B. (2002). A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of Learning Disabilities, 35*, 386–406.

Chen, Q., & Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychologica, 148*, 163–172.

Cheyne, D., & Papanicolaou, A. C. (2017). Magnetoencephalography and magnetic source imaging. In A. C. Papanicolaou (Ed.), *The Oxford handbook of functional brain imaging in neuropsychology and cognitive neurosciences* (pp. 13–42). New York: Oxford University Press.

Christodoulou, J. A., Del Tufo, S. N., Lymberis, J., Saxler, P. K., Ghosh, S. S., Triantafyllou, C., . . . Gabrieli, J. D. (2014). Brain bases of reading fluency in typical reading and impaired fluency in dyslexia. *PLOS ONE, 9*(7), e100552.

Christodoulou, J. A., Murtagh, J., Cyr, A., Perrachione, T. K., Chang, P., Halverson, K., . . . Gabrieli, J. D. (2016). Relation of white-matter microstructure to reading ability and disability in beginning readers. *Neuropsychology, 31*, 508–515.

Christopher, M. E., Hulslander, J., Byrne, B., Samuelsson, S., Keenan, J. M., Pennington, B., . . . Olson, R. K. (2015). Genetic and environmental etiologies of the longitudinal relations between prereading skills and reading. *Child Development, 86*, 342–361.

Christopher, M. E., Keenan, J. M., Hulslander, J., DeFries, J. C., Miyake, A., Wadsworth, S. J., . . . Olson, R. K. (2016). The genetic and environmental etiologies of the relations between cognitive skills and components of reading ability. *Journal of Experimental Psychology: General, 145*, 451–466.

Christopher, M. E., Miyake, A., Keenan, J. M., Pennington, B., DeFries, J. C., Wadsworth, S. J., . . . Olson, R. K. (2012). Predicting word-reading and comprehension with executive function and speed measures across development: A latent variable analysis. *Journal of Experimental Psychology: General, 141*(3), 470–488.

Cirino, P. C., Romain, M. A., Barth, A. E., Tolar, T. D., Fletcher, J. M., & Vaughn, S. (2013). Reading skill components and impairments in middle school struggling readers. *Reading and Writing, 26*, 1059–1086.

Cirino, P. T., Fletcher, J. M., Ewing-Cobbs, L., Barnes, M. A., & Fuchs, L. S. (2007). Cognitive arithmetic differences in learning difficulty groups and the role of behavioral inattention. *Learning Disabilities*

Research and Practice, 22, 25–35.
Cirino, P. T., Fuchs, L. S., Elias, J., Powell, S. R., & Schumacher, R. (2015). Cognitive and mathematical profiles for different forms of learning difficulties. *Journal of Learning Disabilities, 48*, 156–175.
Cirino, P. T., Israelian, M. K., Morris, M. K., & Morris, R. D. (2005). Evaluation of the double-deficit hypothesis in college students referred for learning difficulties. *Journal of Learning Disabilities, 38*, 29–43.
Clark, D. B., Hulme, C., & Snowling, M. (2005). Individual differences in RAN and reading: A response timing analysis. *Journal of Research in Reading, 28*, 73–86.
Clark, D. B., & Uhry, J. K. (1995). *Dyslexia: Theory and practice of remedial instruction*. Baltimore: York Press.
Clark, K. A., Helland, T., Specht, K., Narr, K. L., Manis, F. R., Toga, A. W., & Hugdahl, K. (2014). Neuroanatomical precursors of dyslexia identified from pre-reading through to age 11. *Brain, 137*(12), 3136–3141.
Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading-comprehension difficulties: A randomized controlled trial. *Psychological Science, 21*, 1106–1116.
Clay, M. M. (1993). *Reading Recovery: A guidebook for teachers in training*. Portsmouth, NH: Heinemann.
Clay, M. M. (2002). *An observation survey of early literacy achievement* (2nd ed.). Portsmouth, NH: Heinemann.
Clements, S. D. (1966). *Minimal brain dysfunction in children* (NINDB Monograph No. 3). Washington, DC: U.S. Department of Health, Education and Welfare.
Cohen, S. S., & Vaughn, S. (1994). Gifted students with learning disabilities: What does the research say? *Learning Disabilities: A Multidisciplinary Journal, 5*, 87–94.
Coltheart, M. (2005). Modeling reading: The dual-route approach. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 6–23). Oxford, UK: Blackwell.
Compton, D. L., DeFries, J. C., & Olson, R. K. (2001). Are RAN and phonological awareness deficits additive in children with reading disabilities? *Dyslexia, 3*, 125–149.
Compton, D. L., Miller, A. C., Elleman, A. M., & Steacy, L. M. (2014). Have we forsaken reading theory in the name of "quick fix" interventions for children with reading disability? *Scientific Studies of Reading, 18*(1), 55–73.
Connelly, V., Campbell, S., MacLean, M., & Barnes, J. (2006). Contribution of lower order skills to the written composition of college students with and without dyslexia. *Developmental Neuropsychology, 29*, 175–196.
Connor, C. M. (2016). A lattice model of reading comprehension. *Child Development Perspectives, 10*, 269–274.
Connor, C. M., & Morrison, F. J. (2016). Individualizing student instruction in reading: Implications for policy and practice. *Policy Insights from the Behavioral and Brain Sciences, 3*(1), 54–61.
Connor, C. M., Morrison, F. J., Fishman, B., Crowe, E. C., Al Otaiba, S., & Schatschneider, C. (2013). A longitudinal cluster-randomized controlled study on the accumulating effects of individualized literacy instruction on students' reading from first through third grade. *Psychological Science, 24*(8), 1408–1419.
Connor, C. M., Morrison, F. J., Fishman, B. J., Schatschneider, C., & Underwood, P. (2007). The early years: Algorithm-guided individualized reading instruction. *Science, 315*(5811), 464–465.
Connor, C. M., Piasta, S. B., Fishman, B., Glasney, S., Schatschneider, C., Crowe, E., . . . Morrison, F. J. (2009). Individualizing student instruction precisely: Effects of child by instruction interactions on first graders' literacy development. *Child Development, 80*, 77–100.
Constas, M., & Sternberg, R. (Eds.). (2006). *Translating education theory and research into practice*.

Mahwah, NJ: Erlbaum.
Cornoldi, C., Mammarella, I. C., & Fine, J. G. (2016). *Nonverbal learning disabilities*. New York: Guilford Press.
Corriveau, K. H., Goswami, U., & Thompson, J. M. (2010). Auditory processing and early literacy skills in a preschool and kindergarten population. *Journal of Learning Disabilities, 43*, 369–382.
Cortiella, C., Gamm, S., Rinaldi, C., & Goodman, S. (2014). RTI-based SLD identification toolkit: Criterion 2: Insufficient progress to adequate instruction. *RTI Action Network e-newsletter*. Retrieved from *www.rtinetwork.org/getstarted/sld-identification-toolkit/ld-identification-toolkit-criterion-2*.
Costa, L. J. C., Edwards, C. N., & Hooper, S. R. (2016). Writing disabilities and reading disabilities in elementary school students: Rates of co-occurrence and cognitive burden. *Learning Disability Quarterly, 39*, 17–30.
Critchley, M. (1970). *The dyslexic child*. Springfield, IL: Charles C Thomas.
Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology, 99*, 311–325.
Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Reading comprehension of scientific text: A domain-specific test of the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology, 102*, 687–700.
Cutting, L. E., Clements-Stephens, A., Pugh, K. R., Burns, S., Cao, A., Pekar, J., . . . Rimrodt, S. L. (2013). Not all reading disabilities are dyslexia: Distinct neurobiology of specific comprehension deficits. *Brain Connect, 3*(2), 199–211.
D'Agostino, J. V., & Murphy, J. A. (2004). A meta-analysis of Reading Recovery in United States schools. *Educational Evaluation and Policy Analysis, 26*(1), 23–28.
Danelli, L., Berlingeri, M., Bottini, G., Ferri, F., Vacchi, L., Sberna, M., & Paulesu, E. (2013). Neural intersections of the phonological, visual magnocellular and motor/cerebellar systems in normal readers: Implications for imaging studies on dyslexia. *Human Brain Mapping, 34*, 2669–2687.
Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin and Review, 3*, 422–433.
Davis, C. J., Gayan, J., Knopik, V. S., Smith, S. D., Cardon, L. R., Pennington, B. F., . . . DeFries, J. C. (2001). Etiology of reading difficulties and rapid naming: The Colorado Twin Study of Reading Disability. *Behavior Genetics, 31*, 625–635.
Davis, N., Barquero, L., Compton, D. L., Fuchs, L. S., Fuchs, D., Gore, J. C., . . Anderson, A. (2011). Functional correlates of children's responsiveness to intervention. *Developmental Neuropsychology, 36*, 288–301.
de Jong, P. F., & van der Leij, A. (2003). Developmental changes in the manifestation of a phonological deficit in dyslexic children learning to read a regular orthography. *Journal of Educational Psychology, 95*, 22–40.
de Jong, P. F., & Vrielink, L. O. (2004). Rapid automatic naming: Easy to measure, hard to improve (quickly). *Annals of Dyslexia, 54*, 65–88.
De La Paz, S., Swanson, P. M., & Graham, S. (1998). The contribution of executive control to the revising by students with writing and learning difficulties. *Journal of Educational Psychology, 90*, 448–460.
De Smedt, B., Holloway, I. D., & Ansari, D. (2011). Effects of problem size and arithmetic operation on brain activation during calculation in children with varying levels of arithmetical fluency. *NeuroImage, 57*, 771–781.
De Smedt, B., Noël, M., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical

skills?: A review of evidence from brain and behavior. *Trends in Neuroscience and Education, 2*, 48–56.
De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numeral magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology, 103*, 469–479.
DeFries, J. C., & Fulker, D. W. (1985). Multiple regression analysis of twin data. *Behavior Genetics, 15*, 467–478.
Dehaene, S. (2009). *Reading in the brain: The new science of how we read*. London: Penguin.
Dehaene, S., & Cohen, L. (1997). Cerebral pathways for calculation: Double disassociation between rote verbal and quantitative knowledge of arithmetic. *Cortex, 33*, 219–250.
Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews Neuroscience, 16*, 234–244.
Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences, 9*(7), 335–341.
Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., . . . Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science, 330*, 1359–1364.
Dekker, T. M., Mareschal, D., Johnson, M. H., & Sereno, M. (2014). Picturing words?: Sensorimotor cortex activation for printed words in child and adult readers. *Brain and Language, 139*, 58–67.
Denckla, M. B., Barquero, L. A., Lindstrom, E. R., Benedict, S. L., Wilson, L. M., & Cutting, L. E. (2013). Attention-deficit hyperactivity disorder, executive function, and reading comprehension: Different but related. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 155–169). New York: Guilford Press.
Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., & Fletcher, J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society, 15*, 331–343.
Deno, S. L., & Marston, D. (2001). *Test of Oral Reading Fluency*. Minneapolis: Educational Testing Service.
Denton, C. A. (2011, Winter). Physical exercise and movement-based interventions for dyslexia. *Perspectives on Language and Literacy*, pp. 21–24, 27–32.
Denton, C. A., Ciancio, D. J., & Fletcher, J. M. (2006a). Validity, reliability, and utility of the Observation Survey of Early Literacy Achievement. *Reading Research Quarterly, 41*, 8–34.
Denton, C. A., Cirino, P. T., Barth, A. E., Romain, M., Vaughn, S., Wexler, J., . . . Fletcher, J. M. (2011). An experimental study of scheduling and duration of "Tier 2" first grade reading intervention. *Journal of Research on Educational Effectiveness, 4*, 208–230.
Denton, C. A., Enos, M., York, M. J., Francis, D. J., Barnes, M. A., Kulesz, P. A., Fletcher, J., & Carter, S. (2015). Text processing differences in adolescent adequate and poor comprehenders reading accessible and challenging narrative and informational text. *Reading Research Quarterly, 50*, 393–416.
Denton, C. A., Fletcher, J. M., Anthony, J. L., & Francis, D. J. (2006b). An evaluation of intensive intervention for students with persistent reading difficulties. *Journal of Learning Disabilities, 39*(5), 447–466.
Denton, C. A., Fletcher, J. M., Taylor, W. P., Barth, A. E., & Vaughn, S. (2014). An experimental evaluation of guided reading and explicit interventions for primary-grade students at-risk for reading difficulties. *Journal of Research on Educational Effectiveness, 7*(3), 268–293.
Denton, C. A., & Mathes, P. G. (2003). Intervention for struggling readers: Possibilities and challenges. In B. R. Foorman (Ed.), *Preventing and remediating reading difficulties* (pp. 229–252). Baltimore:

York Press.
Denton, C. A., Nimon, K., Mathes, P. G., Swanson, E. A., Kethley, C., Kurtz, T. B., & Shih, M. (2010). Effectiveness of a supplemental early reading intervention scaled up in multiple schools. *Exceptional Children, 76*, 394–416.
Denton, C. A., Tolar, T. D., Fletcher, J. M., Barth, A. E., Vaughn, S., & Francis, D. J. (2013). Effects of Tier 3 intervention for students with persistent reading difficulties and characteristics of inadequate responders. *Journal of Educational Psychology, 105*, 633–648.
Denton, C. A., Vaughn, S., & Fletcher, J. M. (2003). Bringing research-based practice in reading intervention to scale. *Learning Disabilities Research and Practice, 18*, 201–211.
Devine, A., Soltész, F., Nobes, A., Goswami, U., & Szűcs, D. (2013). Gender differences in developmental dyscalculia depend on diagnostic criteria. *Learning and Instruction, 27*, 31–39.
Dexter, D. D., & Hughes, C. A. (2011). Graphic organizers and students with learning disabilities: A meta-analysis. *Learning Disability Quarterly, 26*, 204–213.
Di Filippo, G., Zoccolotti, P., & Ziegler, J. C. (2008). Rapid naming deficits in dyslexia: A stumbling block for the perceptual anchor theory of dyslexia. *Developmental Science, 11*, 40–47.
Dickman, E. (2017, March). Do we need a new definition of dyslexia? *The Examiner (International Dyslexia Association), 6*(1).
Doehring, D. G. (1978). The tangled web of behavioral research on developmental dyslexia. In A. L. Benton & D. Pearl (Eds.), *Dyslexia* (pp. 123–137). New York: Oxford University Press.
Dole, J., Hosp, J. L., Nelson, K. L., & Hosp, M. K. (2010). Second opinions on the Reading First Initiative: The view from Utah. *Journal of Literacy Research, 42*, 27–48.
Donnelly, P., Plomin, R., & Spencer, C. (2014). The correlation between reading and mathematics ability at age twelve has a substantial genetic component. *Nature Communications, 5*, 4204.
Donovan, M. S., & Cross, C. T. (2002). *Minority students in special and giftzed education*. Washington, DC: National Academies Press.
Doris, J. L. (1993). Defining learning disabilities: A history of the search for consensus. In G. R. Lyon, D. B Gray, J. F. Kavanagh, & N. A. Krasnegor (Eds.), *Better understanding learning disabilities: New views from research and their implications for education and public policies* (pp. 97–116). Baltimore: Brookes.
Duara, R., Kushch, A., Gross-Glenn, K., Barker, W. W., Jallad, B., Pascal, S., . . . Lubs, H. (1991). Neuroanatomic differences between dyslexic and normal readers on magnetic resonance imaging scans. *Archives of Neurology, 48*, 410–416.
Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLOS Genetics, 9*, e1003348.
Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry, 56*(8), 848–856.
Dunbar, S., & Welch, C. (2012). *Iowa Assessments, Form E*. New York: Houghton, Mifflin, Harcourt.
Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., . . . Sexton, H. (2007). School readiness and later achievement. *Developmental Psychology, 43*, 1428.
Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*, 4–58.
Dykman, R. A., Ackerman, P., Clements, S. D., & Peters, J. E. (1971). Specific learning disabilities: An attentional deficit syndrome. In H. R. Myklebust (Ed.), *Progress in learning disabilities* (Vol. 2, pp. 56–93). New York: Grune & Stratton.
Eckert, M. A., Berninger, V. W., Vaden, K. I., Gebregziabher, M., & Tsu, L. (2016, January). Gray matter features of reading disability: A combined meta-analytic and direct analysis approach. *eNeuro*,

ENEURO: 0103–2015.
Eckert, M. A., Leonard, C. M., Richards, T. L., Aylward, E. H., Thomson, J., & Berninger, V. W. (2003). Anatomical correlates of dyslexia: Frontal and cerebellar findings. *Brain, 126*, 482–494.
Eckert, M. A., Leonard, C. M., Wilke, M., Eckert, M., Richards, T., Richards, A., & Berninger, V. (2005). Anatomical signatures of dyslexia in children: Unique information from manual and voxel based morphometry brain measures. *Cortex, 41*, 304–315.
Eden, G. F., Jones, K. M., Cappell, K., Gareau, L., Wood, F. B., Zeffiro, T. A., . . . Flowers, D. L. (2004). Neural changes following remediation in adult developmental dyslexia. *Neuron, 44*, 411–422.
Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., . . . Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research, 79*, 262–300.
Eicher, J. D., Powers, N. R., Miller, L. L., Akshoomoff, N., Amaral, D. G., Bloss, C. S., . . . for the Pediatric Imaging, Neurocognition Genetics, Study. (2013). Genome-wide association study of shared components of reading disability and language impairment. *Genes, Brain and Behavior, 12*, 792–801.
Elbaum, B., & Vaughn, S. (2003). For which students with learning disabilities are self-concept interventions effective? *Journal of Learning Disabilities, 36*, 101–108.
Elbaum, B., Vaughn, S., Hughes, M. T., & Moody, S. W. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure?: A meta-analysis of the intervention research. *Journal of Educational Psychology, 92*, 605–619.
Elleman, A. M. (2017). Examining the impact of inference instruction on the literal and inferential comprehension of skilled and less skilled readers: A meta-analytic review. *Journal of Educational Psychology, 109*, 761–781.
Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness, 2*(1), 1–44.
Elliott, J. G., & Grigorenko, E. L. (2014). *The dyslexia debate*. New York: Cambridge University Press.
Elliott, J. G., & Resing, W. C. M. (2015). Can intelligence testing inform educational intervention for children with reading disability? *Journal of Intelligence, 3*, 137–157.
Ellis, A. W. (1984). The cognitive neuropsychology of developmental (and acquired) dyslexia: A critical survey. *Cognitive Neuropsychology, 2*, 169–205.
Elmore, R. F. (1996). Getting to scale with good educational practice. *Harvard Educational Review, 66*, 1–26.
Engelmann, S., Becker, W. C., Hanner, S., & Johnson, G. (1978). *Implementing the corrective reading series*. Eugene, OR: E-B Press.
Engelmann, S., Carnine, D. W., Engelmann, O., & Kelly, B. (1991). *Connecting math concepts*. Chicago: Science Research Associates.
Espin, C., Wallace, T., Lembke, E., Campbell, H., & Long, J. D. (2010). Creating a progress-monitoring system in reading for middle-school students: Tracking progress toward meeting high-stakes standards. *Learning Disability Research and Practice, 25*, 60–75.
Evans, B. J. W., & Joseph, F. (2002). The effect of coloured filters on the rate of reading in an adult student population. *Ophthalmic and Physiological Optics, 22*, 535–545.
Evans, T. M., Flowers, D. L., Luetje, M. M., Napoliello, E., & Eden, G. F. (2016). Functional neuroanatomy of arithmetic and word reading and its relationship to age. *NeuroImage, 143*, 304–315.
Evans, T. M., Flowers, D. L., Napoliello, E. M., & Eden, G. F. (2014). Sex-specific gray matter volume differences in females with developmental dyslexia. *Brain Structure and Function, 219*, 1041–1054.

Evans, T. M., Kochalka, J., Ngoon, T. J., Wu, S. S., Qin, S., Battista, C., & Menon, V. (2015). Brain structural integrity and intrinsic functional connectivity forecast 6 year longitudinal growth in children's numerical abilities. *Journal of Neuroscience, 35*(33), 11743–11750.

Facoetti, A., Corradi, N., Ruffino, M., Gori, S., & Zorzi, M. (2010). Visual spatial attention and speech segmentation are both impaired in preschoolers at familial risk for developmental dyslexia. *Dyslexia, 16*(3), 226–239.

Farris, E. A., Odegard, T. N., Miller, H. L., Ring, J., Allen, G., & Black, J. (2011). Functional connectivity between the left and right inferior frontal lobes in a small sample of children with and without reading difficulties. *Neurocase, 17*, 425–439.

Farris, E. A., Ring, J., Black, J., Lyon, G. R., & Odegard, T. N. (2016). Predicting growth in word level reading skills in children with developmental dyslexia using an object rhyming functional neuroimaging task. *Developmental Neuropsychology, 41*, 145–161.

Fernald, G. M. (1943). *Remedial techniques in basic school subjects*. New York: McGraw-Hill.

Fernandez, V. G., Juranek, J., Romanowska-Pawliczek, A., Stuebing, K., Williams, V. J., & Fletcher, J. M. (2016). White matter integrity of cerebellar–cortical tracts in reading impaired children: A probabilistic tractography study. *Brain and Language, 161*, 45–56.

Fernandez, V. G., Stuebing, K., Juranek, J., & Fletcher, J. M. (2013). Volumetric analysis of regional variability in the cerebellum of children with dyslexia. *Cerebellum, 12*, 906–915.

Ferstl, E. C., Neumann, J., Bogler, C., & von Cramon, D. Y. (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping, 29*, 581–593.

Ferstl, E. C., Rinck, M., & Von Cramon, D. Y. (2005). Emotional and temporal aspects of situation model processing during text comprehension: An event-related fMRI study. *Journal of Cognitive Neuroscience, 17*, 724–739.

Finch, A. J., Nicolson, R. I., & Fawcett, A. J. (2002). Evidence for a neuroanatomical difference within the olivo–cerebellar pathway of adults with dyslexia. *Cortex, 38*, 529–539.

Finn, E. S., Shen, X., Holahan, J. M., Scheinost, D., Lacadie, C., Papademetris, X., . . . Constable, R. T. (2014). Disruption of functional networks in dyslexia: A whole-brain, data-driven analysis of connectivity. *Biological Psychiatry, 76*, 397–404.

Fisher, S. E., & DeFries, J. C. (2002). Developmental dyslexia: Genetic dissection of a complex cognitive trait. *Nature Reviews Neuroscience, 3*, 767–780.

Flanagan, D. P., Ortiz, S. O., & Alphonso, V. C. (Eds.). (2007). *Essentials of cross-battery assessment*. New York: Wiley.

Fleishner, J. E. (1994). Diagnosis and assessment of mathematics learning disabilities. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 441–458). Baltimore: Brookes.

Fletcher, J. M. (2005). Predicting math outcomes: Reading predictors and comorbidity. *Journal of Learning Disabilities, 38*, 308–312.

Fletcher, J. M., Barth, A. E., & Stuebing, K. K. (2010). A response-to-intervention (RTI) approach to SLD identification. In D. P. Flanagan & V. C. Alfonso (Eds.), *Essentials of specific learning disability identification* (pp. 115–144). New York: Wiley.

Fletcher, J. M., & Currie, D. (2011, Winter). Vision efficiency interventions and reading disability. *Perspectives on Language and Literacy*, pp. 21–24.

Fletcher, J. M., Denton, C., & Francis, D. J. (2005a). Validity of alternative approaches for the identification of LD: Operationalizing unexpected underachievement. *Journal of Learning Disabilities, 38*, 545–552.

Fletcher, J. M., Foorman, B. R., Denton, C. A., & Vaughn, S. (2006). Scaling research on beginning reading: Consensus and conflict. In M. Constas & R. Sternberg (Eds.), *Translating education theory and research into practice* (pp. 53–75). Mahwah, NJ: Erlbaum.

Fletcher, J. M., Francis, D. J., Morris, R. D., & Lyon, G. R. (2005b). Evidence-based assessment of learning disabilities in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34*, 506–522.

Fletcher, J. M., Francis, D. J., Stuebing, K. K., Shaywitz, B. A., Shaywitz, S. E., Shankweiler, D. P., . . . Morris, R. (1996). Conceptual and methodological issues in construct definition. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive functions* (pp. 17–42). Baltimore: Brookes.

Fletcher, J. M., Lyon, G. R., Barnes, M., Stuebing, K. K., Francis, D. J., Olson, R., . . . Shaywitz, B. A. (2002). Classification of learning disabilities: An evidence-based evaluation. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 185–250). Mahwah, NJ: Erlbaum.

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

Fletcher, J. M., & Miciak, J. (2017). Comprehensive cognitive assessments are not necessary for the identification and treatment of learning disabilities. *Archives of Clinical Neuropsychology, 32*, 2–7.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I. Y., Stuebing, K. K., . . . Shaywitz, B. A. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. *Journal of Educational Psychology, 86*, 6–23.

Fletcher, J. M., Stuebing, K. K., Barth, A. E., Denton, C. A., Cirino, P. T., Francis, D. J., & Vaughn, S. (2011). Cognitive correlates of inadequate response to intervention. *School Psychology Review, 40*, 2–22.

Fletcher, J. M., Stuebing, K. K., Barth, A. E., Miciak, J., Francis, D. J., & Denton, C. A. (2014). Agreement and coverage of indicators of response to intervention: A multi-method comparison and simulation. *Topics in Language and Learning Disorders, 34*, 74–89.

Fletcher, J. M., Stuebing, K. K., Morris, R. D., & Lyon, G. R. (2013). Classification and definition of learning disabilities: A hybrid model. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 33–50). New York: Guilford Press.

Fletcher, J. M., Taylor, H. G., Morris, R., & Satz, P. (1982). Finger recognition skills and reading achievement: A developmental neuropsychological analysis. *Developmental Psychology, 18*, 124–132.

Fletcher, J. M., & Vaughn, S. (2009). Response to intervention: Preventing and remediating academic deficits. *Child Development Perspectives, 3*, 30–37.

Flexner. (1910). *Medical education in United States and Canada*. Washington, DC: Science and Health.

Florida Center for Reading Research. (2005). Spell Read P.A.T. Retrieved May 29, 2006, from *www.fcrr.org/FCRRReports/PDF/spell_read_pat.pdf*.

Flower, L., & Hayes, J. R. (1980). The cognition of discovery: Defining a rhetorical problem. *College Composition and Communication, 31*, 21–32.

Flowers, L., Meyer, M., Lovato, J., Wood, F., & Felton, R. (2001). Does third grade discrepancy status predict the course of reading development? *Annals of Dyslexia, 51*, 49–71.

Flynn, J. M., & Rahbar, M. H. (1994). Prevalence of reading failure in boys compared with girls. *Psychology in the Schools, 31*, 66–70.

Foley-Nicpon, M., Allmon, A., Sieck, B., & Stinson, R. D. (2011). Empirical investigation of twice-exceptionality: Where have we been and where are we going? *Gifted Child Quarterly, 55*, 3–17.

Foorman, B. R. (1994). The relevance of a connectionist model of reading for "the great debate. " *Educational Psychology Review, 16*, 25–47.

Foorman, B. R., Fletcher, J. M., & Francis, D. J. (2004). Early reading assessment. In W. Evans & H. J. Walberg (Eds.), *Student learning, evaluating teaching effectiveness* (pp. 81–125). Stanford, CA: Hoover Press.

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., Winikates, D., Mehta, P., Schatschneider, C., & Fletcher, J. M. (1997). Early interventions for children with reading disabilities. *Scientific Studies of Reading, 1*, 255–276.

Foorman, B. R., Herrera, S., Dombek, J., Schatschneider, C., & Petscher, Y. (2017). *The relative effectiveness of two approaches to early literacy intervention in grades K–2* (REL 2017-251). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, Regional Educational Laboratory Southeast. Retrieved from *http://ies.ed.gov/ncee/edlabs*.

Foorman, B. R., Petscher, Y., Lefsky, E. B., & Toste, J. R. (2010). Reading First in Florida: Five years of improvement. *Journal of Literacy Research, 42*, 71–93.

Foorman, B. R., Smith, K. G., & Kosanovich, M. L. (2017). *Rubric for evaluating reading/language arts instructional materials for kindergarten to grade 5* (REL 2017-219). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, Regional Educational Laboratory Southeast. Retrieved from *http://ies.ed.gov/ncee/edlabs*.

Fountas, I. C., & Pinnell, G. S. (1996). *Guided reading: Good first teaching for all children*. Portsmouth, NH: Heinemann.

Francis, D. J., Fletcher, J. M., Stuebing, K. K., Lyon, G. R., Shaywitz, B. A., & Shaywitz, S. E. (2005). Psychometric approaches to the identification of learning disabilities: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities, 38*, 98–110.

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 curves analysis. *Journal of Educational Psychology, 88*, 3–17.

Friend, A., DeFries, J. C., & Olson, R. K. (2008). Parental education moderates genetic influences on reading disability. *Psychological Science, 19*, 1124–1130.

Friend, A., DeFries, J., Olson, R., Pennington, B., Harlaar, N., Byrne, B., . . . Keenan, J. (2009). Heritability of high reading ability and its interaction with parental education. *Behavior Genetics, 39*, 427–436.

Fritschmann, N. S., Deshler, D. D., & Schumaker, J. B. (2007). The effects of instruction in an inference strategy on the reading comprehension skills of adolescents with disabilities. *Learning Disability Quarterly, 30*(4), 245–262.

Frith, U. (1980). *Cognitive processes in spelling*. New York: Academic Press.

Frye, R. E., Liederman J., Malmberg, B., McLean, J., Strickland, D., & Beauchamp, M. S. (2010). Surface area accounts for the relation of gray matter volume to reading-related skills and history of dyslexia. *Cerebral Cortex, 20*, 2625–2635.

Fuchs, D., Compton, D. L., & Changas, P. C. (2014a). Does calculation or word-problem instruction provide a stronger route to pre-algebraic knowledge? *Journal of Educational Psychology, 106*, 990–1006.

Fuchs, D., & Fuchs, L. S. (2005). Peer-assisted learning strategies: Promoting word recognition, fluency, and reading comprehension in young children. *Journal of Special Education, 39*, 34–44.

Fuchs, D., & Fuchs, L. S. (2017). Critique of the National Evaluation of Responsiveness-To-Intervention: A case for simpler frameworks. *Exceptional Children, 83*, 255–268.

Fuchs, D., Fuchs, L. S., Mathes, P., & Simmons, D. (1997). Peer-assisted learning strategies: Making classrooms more responsive to student diversity. *American Educational Research Journal, 34*, 174–206.

Fuchs, D., Fuchs, L. S., & Vaughn, S. (2014b). What is intensive intervention and why is it important? *Teaching Exceptional Children, 46*, 13–18.

Fuchs, D., McMaster, K. L., Fuchs, L. S., & Al Otaiba, S (2013). Data-based individualization as a means of providing intensive instruction to students with serious learning disorders. In H. L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 526–544.). New York: Guilford Press.

Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005a). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology, 97*, 493–513.

Fuchs, L. S., Compton, D. L., Fuchs, D., Powell, S. R., Schumacher, R. F., Hamlett, C. L., Namkung, J. M., & Vukovic, R. K. (2012). Contributions of domain-general cognitive resources and different forms of arithmetic development to pre-algebraic knowledge. *Developmental Psychology, 48*, 1315–1326.

Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: A simplifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research and Practice, 4*, 204–219.

Fuchs, L. S., & Fuchs, D. (2000). Building student capacity to work productively during peer-assisted reading activities. In B. Taylor, M. Graves, & P. van den Broek (Eds.), *Reading for meaning: Fostering comprehension in the middle grades* (pp. 95–115). New York: Teachers College Press.

Fuchs, L. S., Fuchs, D., & Bryant, J. D. (2005b). *Number Rockets*. Manual available at *www.peerassistedlearningstrategies.net*.

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2013a). Intervention effects for students with comorbid forms of learning disability: Understanding the needs of nonresponders. *Journal of Learning Disabilities, 46*, 534–548.

Fuchs, L. S., Fuchs, D., Compton, D. L., Hamlett, C. L., & Wang, A. Y. (2015a). Is word-problem-solving a form of text comprehension? *Scientific Studies of Reading, 19*, 204–223.

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., . . . Fletcher, J. M. (2006a). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word-problems. *Journal of Educational Psychology, 98*, 29–43.

Fuchs, L. S., Fuchs, D., Compton, D. L., Wehby, J., Schumacher, R. F., Gersten, R., & Jordan, N. C. (2015b). Inclusion versus specialized intervention for very low-performing students: What does *access* mean in an era of academic challenge? *Exceptional Children, 81*, 134–157.

Fuchs, L. S., Fuchs, D., Craddock, C., Hollenbeck, K. N., Hamlett, C. L., & Schatschneider, C. (2008a). Effects of small-group tutoring with and without validated classroom instruction on at-risk students' math problem-solving: Are two tiers of prevention better than one? *Journal of Educational Psychology, 100*, 491–509.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989). Effects of instrumental use of curriculum-based measurement to enhance instructional programs. *Remedial and Special Education, 10*(2), 43–52.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Lambert, W., Stuebing, K., & Fletcher, J. M. (2008b). Problem-solving and computational skill: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology, 100*, 30–47.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Phillips, N. B., Karns, K., & Dutka, S. (1997). Enhancing students' helping behavior during peer-mediated instruction with conceptual mathematical explanations. *Elementary School Journal, 97*, 223–250.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Powell, S. R., Capizzi, A. M., & Seethaler, P. M. (2006b). The effects of computer-assisted instruction on number combination skill in at-risk first graders. *Journal of Learning Disabilities, 39*, 467–475.

Fuchs, L. S., Fuchs, D., & Karns, K. (2001a). Enhancing kindergarten children's mathematical development: Effects of peer-assisted learning strategies. *Elementary School Journal, 101*, 495–510.

Fuchs, L. S., Fuchs, D., Karns, K., Yazdian, L., & Powell, S. (2001b). Creating a strong foundation for mathematics learning with kindergarten peer-assisted learning strategies. *Teaching Exceptional Children, 33*, 84–87.

Fuchs, L. S., Fuchs, D., & Prentice, K. (2004). Responsiveness to mathematical–problem-solving treatment among students with risk for mathematics disability, with and without risk for reading disability. *Journal of Learning Disabilities, 27*, 273–306.

Fuchs, L. S., Fuchs, D., Yazdian, L., & Powell, S. R. (2002). Enhancing first-grade children's mathematical development with peer-assisted learning strategies. *School Psychology Review, 31*, 569–584.

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., & Bryant, J. V. (2010a). The contributions of numerosity and domain-general abilities to school readiness. *Child Development, 81*, 1520–1533.

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., Seethaler, P. M., . . . Schatschneider, C. (2010b). Do different types of school mathematics development depend on different constellations of numerical and general cognitive abilities? *Developmental Psychology, 46*, 1731–1746.

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Schatschneider, C., Hamlett, C. L., . . . Changas, P. (2013b). Effects of first-grade number knowledge tutoring with contrasting forms of practice. *Journal of Educational Psychology, 105*, 58–77.

Fuchs, L. S., Geary, D. C., Fuchs, D., Compton, D. L., & Hamlett, C. L. (2016a). Pathways to third-grade calculation versus word-reading competence: Are they more alike or different? *Child Development, 87*, 558–567.

Fuchs, L. S., Gilbert, J. K., Powell, S. R., Cirino, P. T., Fuchs, D., Hamlett, C. L., . . . Tolar, T. D. (2016b). The role of cognitive processes, foundational math skill, and calculation accuracy and fluency in word-problem-solving versus prealgebraic knowledge. *Developmental Psychology, 52*, 2085–2096.

Fuchs, L. S., Powell, S. R., Cirino, P. T., Schumacher, R. F., Marrin, S., Hamlett, C. L., . . . Changas, P. C. (2014a). Does calculation or word-problem instruction provide a stronger route to pre-algebraic knowledge? *Journal of Educational Psychology, 106*, 990–1006.

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., & Hamlett, C. L. (2010c). The effects of strategic counting instruction, with and without deliberate practice, on number combination skill among students with mathematics difficulties. *Learning and Individual Differences, 20*, 89–100.

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., & Hamlett, C. L. (2011). The development of arithmetic and word-problem skill among students with mathematics difficulty. In J. M. Kauffman & D. P. Hallahan (Eds.), *Handbook of special education* (pp. 434–444). New York: Routledge.

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., . . . Zumeta, R. O. (2009). Remediating number combination and word-problem deficits among students with mathematics difficulties: A randomized control trial. *Journal of Educational Psychology, 101*, 561–576.

Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Hamlett, C. L., Cirino, P. T., . . . Changas, P. (2013c). Improving at-risk learners' understanding of fractions. *Journal of Educational Psychology, 105*, 683–700.

Fuchs, L. S., Schumacher, R. F., Malone, A. M., & Fuchs, D. (2015c). *Fraction challenge*. Manual available at *www.peerassistedlearningstrategies.net*.

Fuchs, L. S., Schumacher, R. F., Sterba, S. K., Long, J., Namkung, J., Malone, A., . . . Changas, P. (2014b). Does working memory moderate the effects of fraction intervention?: An aptitude–treatment interaction. *Journal of Educational Psychology, 106*, 499–514.

Fuchs, L. S., Sterba, S. K., Fuchs, D., & Malone, A. S. (2016c). Does evidence-based fractions intervention address the needs of very low-performing students? *Journal of Research on Educational Effectiveness, 9*, 662–677.

Fuchs, L. S., & Vaughn, S. (2012). Responsiveness-to-intervention a decade later. *Journal of Learning Disabilties, 45*, 195–204.

Fuhs, M. W., & Day, J. D. (2011). Verbal ability and executive functioning development in preschoolers at Head Start. *Developmental Psychology, 47*, 404–416.

Fuhs, M. W., & McNeil, N. M. (2013). ANS acuity and mathematics ability in preschoolers from low-income homes: Contributions of inhibitory control. *Developmental Science, 16*, 136–148.

Fuson, K. C., & Kwon, Y. (1992). Korean children's single-digit addition and subtraction: Numbers structured by ten. *Journal for Research in Mathematics Education, 23*, 148–165.

Gaab, N., Gabrieli, J. D. E., Deutsch, G. K., Tallal, P., & Temple, E. (2007). Neural correlates of rapid auditory processing are disrupted in children with developmental dyslexia and ameliorated with training: An fMRI study. *Restorative Neurology and Neuroscience, 25*(3–4), 295–310.

Galaburda, A. M. (Ed.). (1993). *Dyslexia and development: Neurobiological aspects of extra-ordinary brains*. Cambridge, MA: Harvard University Press.

Galaburda, A. M., Sherman, G. F., Rosen, G. D., Aboitiz, F., & Geschwind, N. (1985). Developmental dyslexia: Four consecutive patients with cortical anomalies. *Annals of Neurology, 18*, 222–233.

Gamse, B. C., Jacob, R. T., Horst, M., Boulay, B., & Unlu, F. (2008). *Reading First impact study final report* (NCEE 2009-4038). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research, 84*(1), 74–111.

Garzia, R. P., Borsting, E. J., Nicholson, S. B., Press, L. J., Scheiman, M. M., & Solan, H. A. (2008). *Optometric clinical practice guideline: Care of the patient with learning related vision problems*. St. Louis: American Optometric Association.

Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114*, 345–362.

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37*, 4–15.

Geary, D. C. (2005). Role of cognitive theory in the study of learning disability in mathematics. *Journal of Learning Disabilities, 38*, 305–307.

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47*, 1539–1552.

Geary, D. C. (2013). Early foundations for mathematics learning and their relations to learning disabilities. *Current Directions in Psychological Science, 22*, 23–27.

Geary, D. C., Boykin, A. W., Embretson, S., Reyna, V., Siegler, R., Berch, D. B., & Graban, J. (2008a). Report of the task group on learning processes. In *The Final Report of the National Mathematics Advisory Panel.* Washington, DC: U.S. Department of Education. Retrieved from *www.ed.gov/about/bdscomm/list/mathpanel/report/learning-processes.pdf*.

Geary, D. C., Hoard, M. K., & Bailey, D. H. (2012a). Fact retrieval deficits in low achieving children and children with mathematical learning disability. *Journal of Learning Disabilities, 45*, 291–307.

Geary, D. C., Hoard, M. K., Byrd-Craven, J., & DeSoto, M. C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology, 88*, 121–151.

Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent L., & Numtee C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development, 78*, 1343–1359.

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2012b). Mathematical cognition deficits in children with learning disabilities and persistent low achievement: A five-year prospective study. *Journal of Educational Psychology, 104*, 204–223.

Geary, D. C., Hoard, M. K., Nugent, L., & Byrd-Craven, J. (2008b). Development of number line representations in children with mathematical learning disability. *Developmental Neuropsychology, 33*, 277–299.

Gebauer, D., Enzinger, C., Kronbichler, M., Schurz, M., Reishofer, G., Koschutnig, K., . . . Fink, A. (2012a). Distinct patterns of brain function in children with isolated spelling impairment: New insights. *Neuropsychologia, 50*, 1353–1361.

Gebauer, D., Fink, A., Filippini, N., Johansen-Berg, H., Reishofer, G., Koschutnig, K., . . . Enzinger, C. (2012b). Differences in integrity of white matter and changes with training in spelling impaired children: A diffusion tensor imaging study. *Brain Structure and Function, 217*, 747–760.

Gebauer, D., Fink, A., Kargl, R., Reishofer, G., Koschutnig, K., Purgstaller, C., . . . Enzinger, C. (2012c). Differences in brain function and changes with intervention in children with poor spelling and reading abilities. *PLOS ONE*,*7*(5), e38201.

Gelman, R., & Butterworth, B. (2005). Number and language: How are they related? *Trends in Cognitive Sciences, 9*, 6–10.

Georgiou, G. K., & Parrila, R. (2013). Rapid automatized naming and reading: A review. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 169–185). New York: Guilford Press.

Georgiou, G. K., Parrila, R., & Liao, C. H. (2008). Rapid naming speed and reading across languages that vary in orthographic consistency. *Reading and Writing, 21*, 885–903.

Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Erlbaum.

Gernsbacher, M. A., & Kaschak, M. P. (2003). Neuroimaging studies of language production and comprehension. *Annual Reviews of Psychology, 54*, 91–114.

Gersten, R. M., Compton, D. L., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2009). *Assisting students struggling with reading: Response to intervention and multi-tier intervention in the primary grades*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from *http://ies.ed.gov/ncee/wwc/PracticeGuide.aspx?sid=3*.

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities, 38*, 293–304.

Gersten, R., Vaughn, S., Deshler, D., & Schiller, E. (1997). What we know about using research findings: Implications for improving special education practice. *Journal of Learning Disabilities, 30*, 466–476.

Geschwind, N., & Levitsky, W. (1968). Human brain: Left–right asymmetries in temporal speech region. *Science, 161*(3837), 186–187.

Ghassabian, A., Rescorla, L., Henrichs, J., Jaddoe, V. W., Verhulst, F. C., & Tiemeier, H. (2013). Early lexical development and risk of verbal and nonverbal cognitive delay at school age. *Acta Paediatrica, 103*(1), 70–80.

Gialluisi, A., Newbury, D. F., Wilcutt, E. G., Olson, R. K., DeFries, J. C., Brandler, W. M., . . . Fisher, S.

E. (2014). Genome-wide screening for DNA variants associated with reading and language traits. *Genes, Brain and Behavior, 13*, 686–701.
Gillam, R. B., Loeb, D. F., Hoffman, L. M., Bohman, T., Champlin, C. A., Thibodeau, L., . . . Friel-Patti, S. (2008). The efficacy of Fast ForWord language intervention in school-age children with language impairment: A randomized controlled trial. *Journal of Speech, Language, and Hearing Research, 51*, 97–119.
Gillingham, A., & Stillman, B. (1965). *Remedial training for children with specific disability in reading, spelling and penmanship* (7th ed.). Cambridge, MA: Educators.
Gilman, B. J., Lovecky, D. V., Kearney, K., Peters, D. B., Wasserman, J. D., Silverman, L. K., . . . Curry, P. H. (2013). Critical issues in the identification of gifted students with co-existing disabilities: The twice-exceptional. *Sage Open, 3*(3), 2158244013505855.
Gimenez, P., Bugescu, N., Black, J. M., Hancock, R., Pugh, K., Nagamine, M., . . . Hoeft, F. (2014). Neuroimaging correlates of handwriting quality as children learn to read and write. *Frontiers in Human Neuroscience, 8*, 155.
Glenberg, A. M. (1997). What memory is for: Creating meaning in the service of action. *Behavioral and Brain Sciences, 20*(01), 41–50.
Glod, M. (2008, May 2). Study questions "No Child" act's reading plan. *Washington Post*. Retrieved from *www.washingtonpost.com/wp-dyn/content/article/2008/05/01/AR2008050101399.html*.
Goldman, S. R., Pellegrino, J. W., & Mertz, D. L. (1988). Extended practice of addition facts: Strategy changes in learning-disabled students. *Cognition and Instruction, 5*, 223–265.
Goldstein, K. (1948). *Language and language disorders*. New York: Grune & Stratton.
Goodwin, A. P., & Ahn, S. (2010). A meta-analysis of morphological interventions: Effects on literacy achievement of children with literacy difficulties. *Annals of Dyslexia, 60*(2), 183–208.
Gori, S., & Facoetti, A. (2015). How the visual aspects can be crucial in reading acquisition: The intriguing case of crowding and developmental dyslexia. *Journal of Vision, 15*, 1–20.
Goswami, U. (2002). Phonology, reading development and dyslexia: A cross-linguistic perspective. *Annals of Dyslexia, 52*, 141–163.
Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Neuroscience, 15*, 3–10.
Goswami, U. (2014). The neural basis of dyslexia may originate in primary auditory cortex. *Brain, 137*, 3100–3102.
Goswami, U. (2015). Visual attention span deficits and assessing causality in developmental dyslexia. *Nature Reviews Neuroscience, 16*, 225–226.
Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., & Scott, S. K. (2002). Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proceedings of the National Academy of Sciences, 99*(16), 10911–10916.
Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading and reading disability. *Remedial and Special Education, 7*, 6–10.
Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology, 89*, 170–182.
Graham, S., Collins, A. A., & Rigby-Wills, H. (2017). Writing characteristics of students with learning disabilities and typically achieving peers: A meta-analysis. *Exceptional Children, 83*, 199–218.
Graham, S., Fishman, E. J., Reid, R., & Hebert, M. (2016). Writing characteristics of students with attention deficit hyperactive disorder: A meta-analysis. *Learning Disabilities Research and Practice, 31*(2), 75–89.
Graham, S., & Harris, K. R. (2003). Students with learning disabilities and the process of writing: A

meta-analysis of SRSD studies. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 323–344). New York: Guilford Press.
Graham, S., Harris, K. R., & Chorzempa, B. F. (2002). Contribution of spelling instruction to the spelling, writing, and reading of poor spellers. *Journal of Educational Psychology, 94*, 669–686.
Graham, S., Harris, K. R., & Fink, B. (2000). Is handwriting causally related to learning to write?: Treatment of handwriting problems in beginning writers. *Journal of Educational Psychology, 92*, 620–633.
Graham, S., Harris, K. R., & Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-regulated strategy development. *Contemporary Educational Psychology, 30*, 207–241.
Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review, 81*, 710–744.
Graham, S., McKeown, D., Kiuhara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology, 104*, 879–896.
Graham, S., & Perrin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools*. New York: Carnegie Corporation of New York.
Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers?: A meta-analytic review. *Reading and Writing, 27*, 1703–1743.
Graham, S., Weintraub, N., & Berninger, V. (2001). Which manuscript letters do primary grade children write legibly? *Journal of Educational Psychology, 93*, 488–497.
Gray, C., & Climie, E. A. (2016, July 5). Children with attention deficit/hyperactivity disorder and reading disability: A review of the efficacy of medication treatments. *Frontiers in Psychology, 7*, 988.
Greenberg, D., Wise, J. C., Morris, R., Fredrick, L. D., Rodrigo, V., Nanda, A. O., & Pae, H. K. (2011). A randomized control study of instructional approaches for struggling adult readers. *Journal of Research on Educational Effectiveness, 4*, 101–117.
Greven, C. U., Kovas, Y., Willcutt, E. G., Petrill, S. A., & Plomin, R. (2014). Evidence for shared genetic risk between ADHD symptoms and reduced mathematics ability: A twin study. *Journal of Child Psychology and Psychiatry, 55*, 39–48.
Griffiths, Y. M., & Snowling, M. J. (2002). Predictors of exception word and nonword-reading in dyslexic children: The severity hypothesis. *Journal of Educational Psychology, 94*, 34–43.
Grigorenko, E. L. (2005). A conservative meta-analysis of linkage and linkage-association studies of developmental dyslexia. *Scientific Studies of Reading, 9*, 285–316.
Grigorenko, E. L. (2007). Rethinking disorders of spoken and written language: Generating workable hypotheses. *Journal of Developmental and Behavioral Pediatrics, 28*(6), 478–486.
Grills, A. E., Fletcher, J. M., Vaughn, S. R., Denton, C. A., & Taylor, P. (2013). Anxiety and inattention as predictors of achievement in early elementary school children. *Anxiety, Stress and Coping, 26*, 391–410.
Groen, G., & Resnick, L. B. (1977). Can preschool children invent addition algorithms? *Journal of Educational Psychology, 69*, 645–652.
Gross-Tsur, V., Manor, O., & Shalev, R. S. (1996). Developmental dyscalculia: Prevalence and demographic features. *Developmental Medicine and Child Neurology, 38*, 25–33.
Hale, J., Alfonso, V., Berninger, V., Bracken, B., Christo, C., Clark, E., . . . Yalof, J. (2010). Critical issues in response-to-intervention, comprehensive evaluation, and specific learning disabilities identification and intervention: An expert white paper consensus. *Learning Disability Quarterly, 33*, 223–236.
Hale, J. B., & Fiorello, C. A. (2004). *School neuropsychology: A practitioner's handbook*. New York:

Guilford Press.
Hall, C. S. (2016). Inference instruction for struggling readers: A synthesis of intervention research. *Educational Psychology Review, 28*(1), 1–22.
Hallahan, D. P., Kauffman, J., & Lloyd, J. (1996). *Introduction to learning disabilities*. Needham Heights, MA: Allyn & Bacon.
Hammill, D. D., & Larsen, S. (2009). *Test of Written Language, Fourth Edition*. Austin, TX: PRO-ED.
Hampson, M., Tokoglu, F., Sun, Z., Schafer, R. J., Skudlarski, P., Gore, J. C., & Constable, R. T. (2006). Connectivity–behavior analysis reveals that functional connectivity between left BA39 and Broca's area varies with reading ability. *NeuroImage, 31*, 513–519.
Hamstra-Bletz, L., & Blöte, A. W. (1993). A longitudinal study on dysgraphic handwriting in primary school. *Journal of Learning Disabilities, 26*(10), 689–699.
Hancock, R., Gabrieli, J. D. E., & Hoeft, F. (2016). Shared temporoparietal dysfunction in dyslexia and typical readers with discrepantly high IQ. *Trends in Neuroscience and Education, 5*, 173–177.
Hanich, L. B., Jordan, N. C., Kaplan, D., & Dick, J. (2001). Performance across different areas of mathematical cognition in children with learning difficulties. *Journal of Educational Psychology, 93*, 615–626.
Hanley, J. R. (2005). Learning to read in Chinese. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 316–335). Oxford, UK: Blackwell.
Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1998). Does special education raise academic achievement for students with disabilities? (Working Paper No. 6690). Cambridge, MA: National Bureau of Economic Research. Retrieved on September 6, 2015, from *www.nber.org/papers/w6690*.
Harlaar, N., Cutting, L., Deater-Deckard, K., DeThorne, L. S., Justice, L. M., Schatschneider, C., . . . Petrill, S. A. (2010). Predicting individual differences in reading comprehension: A twin study. *Annals of Dyslexia, 60*(2), 265–288.
Harlaar, N., Kovas, Y., Dale, P. S., Petrill, S. A., & Plomin, R. (2012). Mathematics is differentially related to reading comprehension and word decoding: Evidence from a genetically sensitive design. *Journal of Educational Psychology, 104*(3), 622–635.
Harris, K. R., Graham, S., Mason, L. H., & Friedlander, B. (2008). *Powerful writing strategies for all students*. Baltimore: Brookes.
Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: Brookes.
Hart, S., Logan, J. A., Thompson, L. A., Kovas, Y., McLoughlin, G., & Petrill, S. (2016). A latent profile analysis of math achievement, numerosity, and math anxiety in twins. *Journal of Educational Psychology, 108*, 181–193.
Hart, S., Petrill, S. A., Thompson, L., & Plomin, R. (2009). The ABCs of math: A genetic analysis of mathematics and its links with reading ability and general cognitive ability. *Journal of Educational Psychology, 101*, 388–402.
Hasbrouck, J. E., Ihnot, C., & Rogers, G. (1999). Read Naturally: A strategy to increase oral reading fluency. *Reading Research and Instruction, 39*, 27–37.
Hawelka, S., Gagl, B., & Wimmer, H. (2010). A dual-route perspective on eye movements of dyslexic readers. *Cognition, 115*, 367–379.
Hawke, J. L., Wadsworth, S. J., Olson, R. K., & DeFries, J. C. (2007). Etiology of reading difficulties as a function of gender and severity. *Reading and Writing, 20*, 13–25.
Hayes, J. R., & Chenoweth, N. A. (2007). Working memory in an editing task. *Written Communication, 24*, 283–294.
Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of the writing process. In L. W. Gregg & E. R. Steinbery (Eds.), *Cognitive processes in writing* (pp. 3–30). Hillsdale, NJ: Erlbaum.

Hayiou-Thomas, M. E., Carroll, J. M., Leavett, R., Hulme, C., & Snowling, M. J. (2017). When does speech sound disorder matter for literacy?: The role of disordered speech errors, co-occurring language impairment and family risk of dyslexia. *Journal of Child Psychology and Psychiatry, 58*, 197–205.

He, Q., Xue, G., Chen, C., Chen, C., Lu, Z. L., & Dong, Q. (2013). Decoding the neuroanatomical basis of reading ability: A multivoxel morphometric study. *Journal of Neuroscience, 33*, 12835–12843.

Hebert, M., Bohaty, J. J., Nelson, J. R., & Brown, J. (2016). The effects of text structure instruction on expository reading comprehension: A meta-analysis. *Journal of Educational Psychology, 108*, 609–629.

Hecht, S. A., Close, L., & Santisi, M. (2003). Sources of individual differences in fraction skills. *Journal of Experimental Child Psychology, 86*, 277–302.

Hecht, S. A., & Vagi, K. J. (2010). Sources of group and individual differences in emerging fraction skills. *Journal of Educational Psychology, 102*, 843–859.

Henderson, L., Snowling, M., & Clarke, P. (2013). Accessing, integrating, and inhibiting word meaning in poor comprehenders. *Scientific Studies of Reading, 17*(3), 177–198.

Hess, F. M. (2008). *When education research matters* (Education Outlook No. 22641). Washington, DC: American Enterprise Institute for Policy Research. Retrieved from *www.aei.org/publication/when-education-research-matters*.

Hessler, G. L. (1987). Educational issues surrounding severe discrepancy. *Learning Disabilities Research, 3*, 43–49.

Hiebert, E. H. (1994). Reading Recovery in the United States: What difference does it make to an age cohort? *Educational Researcher, 23*(9), 15–25.

Hiebert, J., & Wearne, D. (1985). A model of students' decimal calculation procedures. *Cognition and Instruction, 2*, 175–205.

Hinshaw, S. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin, 111*, 127–155.

Hinshaw, S. P., & Scheffler, R. M. (2014). *The ADHD explosion*. New York: Oxford University Press.

Hinshelwood, J. (1917). Congenital word-blindness. *The Lancet, 190*(4922), 959–988.

Hodent, C., Bryant, P., & Houde, O. (2005). Language-specific effects on number computation in toddlers. *Developmental Science, 8*, 420–423.

Hoeft, F., McCandliss, B. D., Black, J. M., Gantman, A., Zakerani, N., Hulme, C., . . . Gabrieli, J. D. (2011). Neural systems predicting long-term outcome in dyslexia. *Proceedings of the National Academy of Sciences of the USA, 108*, 361–366.

Hoeft, F., Ueno, T., Reiss, A. L., Meyler, A., Whitfield-Gabrieli, S., Glover, G. H., . . . Just, M. A. (2007). Prediction of children's reading skills using behavioral, functional, and structural neuroimaging measures. *Behavioral Neuroscience, 121*, 602–613.

Holland, J., McIntosh, D., & Huffman, L. (2004). The role of phonological awareness, rapid automatized naming, and orthographic processing in word-reading. *Journal of Psychoeducational Assessment, 22*, 233–260.

Hood, P. (2003). *Scientific research and evidence-based practice*. San Franciso: WestEd.

Hooper, S. R., Montgomery, J., Swartz, C., Reed, M., Sandler, A., Levine, M., . . . Wasileski, T. (1994). Measurement of written language expression. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 375–418). Baltimore: Brookes.

Hooper, S. R., Swartz, C. W., Montgomery, J., Reed, M. S., Brown, T. T., Wasileski, T. J., . . . Thomas, J. (1993). Prevalence of writing problems across three middle school samples. *School Psychology*

Review, 22, 610–622.
Hooper, S. R., Swartz, C. W., Wakely, M. B., de Kruif, R. E. L., & Montgomery, J. W. (2002). Executive functions in elementary school children with and without problems in written expression. *Journal of Learning Disabilities, 35*, 57–68.
Hooper, S. R., Wakely, M. B., de Kruif, R. E. L., & Swartz, C. W. (2006). Aptitude–treatment interactions revisited: Effect of metacognitive intervention on subtypes of written expression in elementary school students. *Developmental Neuropsychology, 29*, 217–241.
Hooper, S. R., & Willis, W. G. (1989). *Learning disability subtyping: Neuropsychological foundations, conceptual models, and issues in clinical differentiation*. New York: Springer-Verlag.
Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2*, 127–160.
Horowitz-Kraus, T., Eaton, K., Farah, R., Hajinazarian, A., Vannest, J., & Holland, S. K. (2015). Predicting better performance on a college preparedness test from narrative comprehension at the age of 6 years: An fMRI study. *Brain Research, 1629*, 54–62.
Horowitz-Kraus, T., Vannest, J. J., & Holland, S. K. (2013). Overlapping neural circuitry for narrative comprehension and proficient reading in children and adolescents. *Neuropsychologia, 51*, 2651–2662.
Horwitz, B., Rumsey, J. M., & Donohue, B. C. (1998). Functional connectivity of the angular gyrus in normal reading and dyslexia. *Proceedings of the National Academy of Sciences of the USA, 95*, 8939–8944.
Hoskyn, M., & Swanson, H. L (2000). Cognitive processing of low achievers and children with reading disabilities: A selective meta-analytic review of the published literature. *School Psychology Review, 29*, 102–119.
Hosseini, S. H., Black, J. M., Soriano, T., Bugescu, N., Martinez, R., Raman, M. M., . . . Hoeft, F. (2013). Topological properties of large-scale structural brain networks in children with familial risk for reading difficulties. *NeuroImage, 71*, 260–274.
Hughes, C. A., & Dexter, D. D. (2008). Selecting a scientifically based core curriculum for Tier 1. *RTI Action Network e-newsletter*. Retrieved from *www.rtinetwork.org/learn/research/selectingcorecurriculum-tier1*.
Hughes, C. A., Ruhl, K. L., Schumaker, J. B., & Deshler, D. D. (2002). Effects of instruction in an assignment completion strategy on the homework performance of students with learning disabilities in general education classes. *Learning Disabilities Research, 17*, 1–18.
Hughes, D. C., Keeling, B., & Tuck, B. F. (1983). Effects of achievement expectations and handwriting quality on scoring essays. *Journal of Educational Measurement, 20*, 65–70.
Hulme, C. (1988). The implausibility of low-level visual deficits as a cause of children's reading difficulties. *Cognitive Neuropsychology, 5*, 369–374.
Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. J. (2012). The causal role of phoneme awareness and letter–sound knowledge in learning to read combining intervention studies with mediation analyses. *Psychological Science, 23*, 572–577.
Hulme, C., Nash, H. M., Gooch, D., Lervåg, A., & Snowling, M. J. (2015). The foundations of literacy development in children at familial risk of dyslexia. *Psychological Science, 26*, 1877–1886.
Hulme, C., & Snowling, M. J. (2009). *Developmental disorders of language learning and cognition*. Oxford, UK: Wiley-Blackwell.
Hulme, C., Snowling, M. J., 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*(4), 351–365.
Hunter, J. V., & Wang, Z. J. (2001). MR spectroscopy in pediatric neuroradiology. *MRI Clinics of North America, 9*, 165–189.

Hutchison, S. L., Ghuman, J. K., Ghuman, H. S., Karpov, I., & Schuster, J. M. (2016). Efficacy of atomoxetine in the treatment of attention-deficit hyperactivity disorder in patients with common comorbidities in children, adolescents and adults: A review. *Therapeutic Advances in Psychopharmacology, 6*, 317–334.

Hutton, J. S., Horowitz-Kraus, T., Mendelsohn, A. L., DeWitt, T., Holland, S. K., & C-MIND Authorship Consortium. (2015). Home reading environment and brain activation in preschool children listening to stories. *Pediatrics, 136*(3), 466–478.

Hyde, D. C., Khanum, S., & Spelke, E. S. (2014). Brief non-symbolic, approximate number practice enhances subsequent exact symbolic arithmetic in children. *Cognition, 131*, 92–107.

Hynd, G. W., Hall, J., Novey, E. S., Eliopulos, D., Black, K., Gonzalez, J. J., . . . Cohen, M. (1995). Dyslexia and corpus callosum morphology. *Archives of Neurology, 52*, 32–38.

Hynd, G. W., Semrud-Clikeman, M., Lorys, A. R., Novey, E. S., & Eliopulos, D. (1990). Brain morphology in developmental dyslexia and attention deficit disorder/hyperactivity. *Archives of Neurology, 47*(8), 919–926.

Ihnot, C., Mastoff, J., Gavin, J., & Hendrickson, L. (2001). *Read naturally*. St. Paul, MN: Read Naturally.

Im, K., Raschle, N. M., Smith, S. A., Grant, P. E., & Gaab, N. (2016). Atypical sulcal pattern in children with developmental dyslexia and at-risk kindergarteners. *Cerebral Cortex, 26*, 1138–1148.

Individuals with Disabilities Education Act of 2004, Public Law 108-466.

Iovino, I., Fletcher, J. M., Breitmeyer, B. G., & Foorman, B. R. (1999). Colored overlays for visual perceptual deficits in children with reading disability and attention deficit/hyperactivity disorder: Are they differentially effective? *Journal of Clinical and Experimental Neuropsychology, 20*, 791–806.

Irlen, H. (1994). Scotopic sensitivity/Irlen syndrome—Hypothesis and explanation of the syndrome. *Journal of Behavioral Optometry, 5*, 62–65.

Iuculano, T., Rosenberg-Lee, M., Richardson, J., Tenison, C., Fuchs, L., Supekar, K., & Menon, V. (2015). Cognitive tutoring induces widespread neuroplasticity and remediates brain function in children with mathematical learning disabilities. *Nature Communications, 6*, 8453.

Iversen, S., Tunmer, W. E., & Chapman, J. W. (2005). The effects of varying group size on the Reading Recovery approach to preventive early intervention. *Journal of Learning Disabilities, 38*, 456–472.

Jacob, R., & Parkinson, J. (2015). The potential for school-based interventions that target executive function to improve academic achievement: A review. *Review of Educational Research, 85*, 512–552.

Jaffe-Dax, S., Lieder, I., Biron, T., & Ahissar, M. (2016). Dyslexics' usage of visual priors is impaired. *Journal of Vision, 16*, 1–9.

James, K. H., & Engelhardt, L. (2012). The effects of handwriting experience on functional brain development in pre-literate children. *Trends in Neuroscience and Education, 1*(1), 32–42.

James-Burdumy, S., Deke, J., Gersten, R., Lugo-Gil, J., Newman-Gonchar, R., Dimino, J., . . . Liu, A. Y.-H. (2012). Effectiveness of four supplemental reading comprehension interventions. *Journal of Research on Educational Effectiveness, 5*(4), 345–383.

Jednorog, K., Marchewka, A., Altarelli, I., Monzalvo Lopez, A. K., van Ermingen-Marbach, M., Grande, M., . . . Ramus, F. (2015). How reliable are gray matter disruptions in specific reading disability across multiple countries and languages?: Insights from a large-scale voxel-based morphometry study. *Human Brain Mapping, 36*, 1741–1754.

Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003). Accuracy and fluency in list and context reading of skilled and RD groups: Absolute and relative performance levels. *Learning Disabilities Research and Practice, 18*, 237–245.

Jenkins, J. R., & O'Connor, R. E. (2003). Cooperative learning for students with learning disabilities:

Evidence from experiments, observations, and interviews. In H. L. Swanson, R. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 417–430). New York: Guilford Press.

Jiménez, J. E., de la Cadena, C. G., Siegel, L. S., O'Shanahan, I., García, E., & Rodríguez, C. (2011). Gender ratio and cognitive profiles in dyslexia: A cross-national study. *Reading and Writing, 24*, 729–747.

Jimerson, S. R., Burns, M., & VanDerHeyden, A. M. (Eds.). (2015). *The handbook of response to intervention: The science and practice of multi-tiered systems of support*. New York: Springer.

Jitendra, A. K., Edwards, L. L., Sacks, G., & Jacobson, L. A. (2004). What research says about vocabulary instruction for students with learning disabilities. *Exceptional Children, 70*, 299–322.

Joanisse, M. F., Manis, F. R., Keating, P., & Seidenberg, M. S. (2000). Language deficits in dyslexic children: Speech perception, phonology, and morphology. *Journal of Experimental Child Psychology, 77*, 30–60.

Johnson, D. J., & Blalock, J. (Eds.). (1987). *Adults with learning disabilities*. Orlando, FL: Grune & Stratton.

Johnson, D. J., & Myklebust, H. (1967). *Learning disabilities*. New York: Grune & Stratton.

Johnson, W., Bouchard, T. J., Jr., Segal, N. L., & Samuels, J. (2005). General intelligence and reading performance in adults: Is the genetic factor structure the same as for children? *Personality and Individual Differences, 38*, 1413–1428.

Joint Committee on Standards for Educational and Psychological Testing of the Educational Research Association, American Psychological Association, & National Council of Measurement in Education. (2014). *Standards for psychological and educational testing*. Washington DC: American Educational Research Association.

Jolles, D., Ashkenazi, S., Kochalka, J., Evans, T., Richardson, J., Rosenberg-Lee, M., . . . Menon, V. (2016). Parietal hyper-connectivity, aberrant brain organization, and circuit?based biomarkers in children with mathematical disabilities. *Developmental Science, 19*(4), 613–631.

Jones, D., & Christensen, C. (1999). The relationship between automaticity in handwriting and students' ability to generate written text. *Journal of Educational Psychology, 91*, 44–49.

Jones, J. N., Abbott, R. D., & Berninger, V. W. (2014). Predicting levels of reading and writing achievement in typically developing, English-speaking 2nd and 5th graders. *Learning and Individual Differences, 32*, 54–68.

Jordan, N. C., & Hanich, L. B. (2000). Mathematical thinking in second-grade children with different forms of LD. *Journal of Learning Disabilities, 33*, 567–578.

Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003). Arithmetic fact mastery in young children: A longitudinal investigation. *Journal of Experimental Child Psychology, 85*, 103–119.

Jordan, N. C., Hansen, N., Fuchs, L. S., Siegler, R. S., Gersten, R., & Micklos, D. (2013). Developmental predictors of fraction concepts and procedures. *Journal of Experimental Child Psychology, 116*, 45–58.

Joseph, L. M., & Hunter, A. D. (2001). Differential application of a cue card strategy for solving fraction problems: Exploring instructional utility of the cognitive assessment system. *Child Study Journal, 31*, 123–136.

Joshi, R. M., Treiman, R., Carreker, S., & Moats, L. C. (2008–2009). How words cast their spell. *American Educator, 32*, 6–16.

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78*, 243–255.

Justice, L., Mashburn, A., & Petscher, Y. (2013). Very early language skills of fifth-grade poor comprehenders. *Journal of Research in Reading, 36*, 172–185.

Kail, R., & Hall, L. K. (1994). Processing speed, naming speed, and reading. *Developmental Psychology,*

30, 949–954.
Kame'enui, E. J., Fien, H., & Korgesaar, J. (2013). Direct instruction as Eo nomine and contronym: Why the right words and details matter. In H. L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 489–506). New York: Guilford Press.
Kamhi, A. G. (2009). The case for the narrow view of reading. *Language, Speech, and Hearing Services in Schools, 40*, 174–177.
Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., & Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices* (IES Practice Guide, NCEE 2008-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Science, U.S. Department of Education.
Kamin, L. J. (1974). *The science and politics of I. Q*. Potomac, MD: Erlbaum.
Katusic, S. K., Colligan, R. C., Weaver, A. L., & Barbaresi, W. J. (2009). The forgotten learning disability: Epidemiology of written-language disorder in a population-based birth cohort (1976–1982), Rochester, Minnesota. *Pediatrics, 123*, 1306–1313.
Kaufman, A. S., & Kaufman, N. L. (2014). *Kaufman Test of Educational Achievement, Third Edition*. San Antonio, TX: Pearson.
Kaufmann, L., Wood, G., Rubinsten, O., & Henik, A. (2011). Meta-analyses of developmental fMRI studies investigating typical and atypical trajectories of number processing and calculation. *Developmental Neuropsychology, 36*(6), 763–787.
Kavale, K. A. (1988). Learning disability and cultural disadvantage: The case for a relationship. *Learning Disability Quarterly, 11*, 195–210.
Kavale, K. A., & Forness, S. R. (1984). A meta-analysis of the validity of Weschler Scale profiles and recategorizations: Patterns or parodies? *Learning Disability Quarterly, 7*, 136–156.
Kavale, K. A., & Forness, S. R. (2000). What definitions of learning disability say and don't say: A critical analysis. *Journal of Learning Disabilities, 33*, 239–256.
Kavale, K. A., & Mostert, M. P. (2004). Social skills interventions for individuals with learning disabilities. *Learning Disability Quarterly, 27*, 31–43.
Kavale, K. A., & Reese, J. H. (1992). The character of learning disabilities: An Iowa profile. *Learning Disability Quarterly, 15*(2), 74–94.
Kearns, D. M., & Fuchs, D. (2013). Does cognitively focused instruction improve the academic performance of low-achieving students? *Exceptional Children, 79*, 263–290.
Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading, 12*(3), 281–300.
Keenan, J. M., Betjemann, R. S., Wadsworth, S. J., DeFries, J. C., & Olson, R. K. (2006). Genetic and environmental influences on reading and listening comprehension. *Journal of Research in Reading, 29*, 75–91.
Keenan, J. M., Hua, A. N., Meenan, C. E., Olson, R. K., Pennington, B. F., & Willcutt, E. G. (2014). Issues in identifying poor comprehenders. *Topics in Cognitive Psychology, 114*, 753–777.
Keenan, J. M., & Meenan, C. E. (2014). Test differences in diagnosing reading comprehension deficits. *Journal of Learning Disabilities, 47*, 125–135.
Kellam, S. G., Rebok, G. W., Mayer, L. S., Ialongo, N., & Kalodner, C. R. (1994). Depressive symptoms over first grade and their response to a developmental epidemiologically based preventive trial aimed at improving achievement. *Development and Psychopathology, 6*, 463–481.
Keller-Margules, M. A. (2012). Fidelity of implementation framework: A critical need for response to intervention models. *Psychology in the Schools, 49*, 342–352.
Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The*

science of writing: Theories, methods, individual differences, and applications (pp. 57–71). Hillsdale, NJ: Erlbaum.

Kelly, B., Gersten, R., & Carnine, R. (1990). Student error patterns as a function of curriculum design: Teaching fractions to remedial high school students and high school students with learning disabilities. *Journal of Learning Disabilities, 23*, 23–29.

Kemple, J., Corrin, W., Nelson, E., Salinger, T., Herrmann, S., Drummond, K., & Strasberg, P. (2008). *The enhanced reading opportunities study: Early impact and implementation findings* (NCEE 2008-4015). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Kendeou, P., Bohn-Gettler, C., White, M. J., & van den Broek, P. (2008). Children's inference generation across different media. *Journal of Research in Reading, 31*(3), 259–272.

Kendeou, P., van den Broek, P. I., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology, 101*(4), 765–778.

Kent, S. C., & Wanzek, J. (2016). The relationship between component skills and writing quality and production across developmental levels: A meta-analysis of the last 25 years. *Review of Educational Research, 86*(2), 570–601.

Kibby, M. Y., Francher, J. B., Markanen, R., Lewandowski, A., & Hynd, G. W. (2003). A test of the cerebellar deficit hypothesis of dyslexia. *Journal of the International Neuropsychological Society, 9*, 219.

Kieren, T. E. (1993). Rational and fractional numbers: From quotient fields to recursive understanding. In T. P. Carpenter, E. Fennema, & T. A. Romberg (Eds.), *Rational numbers: An integration of research* (pp. 49–84). Hillsdale, NJ: Erlbaum.

Kim, Y.-S. (2016). Direct and mediated effects of language and cognitive skills on comprehension of oral narrative texts (listening comprehension) for children. *Journal of Experimental Child Psychology, 141*, 101–120.

Kim, Y.-S., Al Otaiba, S., Puranik, C., Folsom, J. S., Greulich, L., & Wagner, R. K. (2011). Componential skills of beginning writing: An exploratory study. *Learning and Individual Differences, 21*, 517–525.

Kim, Y.-S., Park, C., & Park, Y. (2015). Dimensions of discourse-level oral language skills and their relations to reading comprehension and written composition: An exploratory study. *Reading and Writing: An Interdisciplinary Journal, 28*, 633–654.

Kim, Y.-S. G., & Schatschneider, C. (2017). Expanding the developmental models of writing: A direct and indirect effects model of developmental writing (DIEW). *Journal of Educational Psychology, 109*, 35–50.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction–integration model. *Psychological Review, 95*(2), 163–182.

Kirby, J. R., & Savage, R. S. (2008). Can the simple view deal with the complexities of reading? *Literacy, 42*, 75–82.

Kirk, S. A. (1963). Behavioral diagnosis and remediation of learning disabilities. *Conference on Exploring Problems of the Perceptually Handicapped Child, 1*, 1–23.

Kirkby, J. A., Webster, L. A. D., Blythe, H. I., & Liversedge, S. P. (2008). Binocular coordination during reading and non-reading tasks. *Psychological Bulletin, 134*, 742–763.

Knapp, M. S. (1995). *Teaching for meaning in high-poverty classrooms*. New York: Teachers College Press.

Knopik, V. S., & DeFries, J. C. (1999). Etiology of covariation between reading and mathematics performance: A twin study. *Twin Research, 2*, 226–234.

Koponen, T., Aunola, K., Ahonen, T., & Nurmi, J. (2007). Cognitive predictors of single-digit and procedural calculation skills and their covariation with reading skill. *Journal of Experimental Child Psychology, 97*, 220–241.

Koponen, T., Salmi, P., Eklund, K., & Aro, T. (2013). Counting and RAN: Predictors of arithmetic calculation and reading fluency. *Journal of Educational Psychology, 105*, 162–175.

Kornilov, S. A., & Grigorenko, E. L. (2018). Genetics of specific reading disability: The state of affairs. In A. Galaburda, N. Gaab, F. Hoeft, & P. McCardle (Eds.), *Dyslexia and neuroscience: The Geschwind–Galaburda hypothesis 30 years later* (pp. 156–166). Baltimore, MD: Brookes.

Kovaleski, J. F., VanDerHeyden, A. M., & Shapiro, E. S. (2013). *The RTI approach to evaluating specific learning disabilities*. New York: Guilford Press.

Kovas, Y., Haworth, C. M., Petrill, S. A., & Plomin, R. (2007). Mathematical ability of 10-year-old boys and girls: Genetic and environmental etiology of typical and low performance. *Journal of Learning Disabilities, 40*, 554–567.

Koyama, M. S., Di Martino, A., Kelly, C., Jutagir, D. R., Sunshine, J., Schwartz, S. J., . . . Milham, M. P. (2013). Cortical signatures of dyslexia and remediation: An intrinsic functional connectivity approach. *PLOS ONE, 8*, e55454.

Koyama, M. S., Kelly, C., Shehzad, Z., Penesetti, D., Castellanos, F. X., & Milham, M. P. (2010). Reading networks at rest. *Cerebral Cortex, 20*, 2549–2559.

Krafnick, A. J., Flowers, D. L., Luetje, M. M., Napoliello, E. M., & Eden, G. F. (2014). An investigation into the origin of anatomical differences in dyslexia. *Journal of Neuroscience, 34*(3), 901–908.

Kraft, I., Cafiero, R., Schaadt, G., Brauer, J., Neef, N. E., Müller, B., . . . Skeide, M. A. (2015). Cortical differences in preliterate children at familiar risk of dyslexia are similar to those observed in dyslexic readers. *Brain, 138*, 378.

Kranzler, J. H., Floyd, R. G., Benson, N., Zaboski, B., & Thibodaux, L. (2016). Classification agreement analysis of cross-battery assessment in the identification of specific learning disorders in children and youth. *International Journal of School and Educational Psychology, 4*, 124–136.

Krashen, S. D. (2004). *The power of reading: Insights from the research* (2nd ed.). Portsmouth NH: Heinemann.

Kraus, N., & White-Schwoch, T. (2015). Unraveling the biology of auditory learning: A cognitive–sensorimotor–reward framework. *Trends in Cognitive Sciences, 19*, 642–654.

Kriss, I., & Evans, B. J. (2005). The relationship between dyslexia and Meares–Irlen syndrome. *Journal of Research in Reading, 28*, 350–364.

Kroesbergen, E. H., & Van Luit, J. E. H. (2003). Mathematics interventions for children with special needs: A meta-analysis. *Remedial and Special Education, 24*, 97–114.

Kucian, K., Ashkenazi, A. A., Hanggi, J., Rotzer, S., Jancke, L., Martin, E., & von Aster, M. (2014). Developmental dyscalculia: A disconnection syndrome? *Brain Structure and Function, 219*, 1721–1733.

Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schönmann, C., Plangger, F., . . . von Aster, M. (2011). Mental number line training in children with developmental dyscalculia. *NeuroImage, 57*(3), 782–795.

Kudo, M. F., Lussier, C. M., & Swanson, H. L. (2015). Reading disabilities in children: A selective meta-analysis of the cognitive literature. *Research in Developmental Disabilities, 40*, 51–62.

Kuhn, M. R., Schwanenflugel, P. J., Morris, R. D., Morrow, L. M., Woo, D. G., Meisinger, E. B., . . . Stahl, S. A. (2006). Teaching children to become fluent and automatic readers. *Journal of Literacy Research, 38*, 357–387.

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95*, 3–21.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information process in reading.

Cognitive Psychology, 6, 293–323.

LaForte, E. M., McGrew, K. S., & Schrank, F. A. (2014). *WJ IV technical abstract* (Woodcock–Johnson IV Assessment Service Bulletin No. 2). Rolling Meadows, IL: Riverside.

Lambe, E. K. (1999). Dyslexia, gender, and brain imaging. *Neuropsychologia, 37*, 521–536.

Landi, N., & Perfetti, C. A. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and Language, 102*, 30–45.

Langer, N., Benjamin, C., Minas, J., & Gaab, N. (2015). The neural correlates of reading fluency deficits in children. *Cerebral Cortex, 25*, 1441–1453.

Lawrence, J. F., Crosson, A. C., Paré-Blagoev, E. J., & Snow, C. E. (2015). Word generation randomized trial: Discussion mediates the impact of program treatment on academic word learning. *American Educational Research Journal, 52*(4), 750–786.

Leach, J. M., Scarborough, H. S., & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of Educational Psychology, 95*, 211–224.

Lee, K., Ng, S. F., Bull, R., Pe, M. L., & Ho, R. H. M. (2011). Are patterns important?: An investigation of the relationships between proficiency in patterns, calculation, executive functioning, and algebraic word-problems. *Journal of Educational Psychology, 103*, 269–281.

Lee, K., Ng, S. F., Ng, E. L., & Lim, Z. Y. (2004). Working memory and literacy as predictors of performance on algebraic word-problems. *Journal of Experimental Child Psychology, 89*, 140–158.

Lee, S. H., & Tsai, S.-F. (2017). Experimental intervention research in students with specific poor comprehension: A systematic review of treatment outcomes. *Reading and Writing, 30*, 917–943.

LeFevre, J.-A., DeStefano, D., Coleman, B., & Shanahan, T. (2005). Mathematical cognition and working memory. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 361–377). New York: Psychology Press.

LeFevre, J.-A., Fast, L., Skwarchuk, S.-L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., . . . Penne-Wilger, M. (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child Development, 81*, 1753–1767.

Lerner, J. W. (1989). Educational interventions in learning disabilities. *Journal of the American Academy of Child and Adolescent Psychiatry, 28*(3), 326–331.

Lervåg, A., & Hulme, C. (2009). Rapid automatized naming (RAN) taps a mechanism that places constraints on the development of early reading fluency. *Psychological Science, 20*, 1040–1048.

Lesaux, N. K., Kieffer, M. J., Kelley, J. G., & Harris, J. R. (2014). Effects of academic vocabulary instruction for linguistically diverse adolescents: Evidence from a randomized field trial. *American Educational Research Journal, 51*(6), 1159–1194.

Levy, B. A. (2001). Moving the bottom: Improving reading fluency. In M. Wolf (Ed.), *Dyslexia, fluency, and the brain* (pp. 357–382). Timonium, MD: York Press.

Lewis, C., Hitch, G. J., & Walker, P. (1994). The prevalence of specific arithmetic difficulties and specific reading difficulties in 9- to 10-year-old boys and girls. *Journal of Child Psychology and Psychiatry, 35*, 283–292.

Lewis, M., & Samuels, S. J. (2003). *Read more—read better?: A meta-analysis of the literature on the relationship between exposure to reading and reading achievement*. Unpublished paper, University of Minnesota. Retrieved January 18, 2016, from *www.tc.umn.edu/~samue001/publications.htm*.

Liberman, A. L. (1996). *Speech: A special code*. Cambridge, MA: MIT Press.

Liberman, I. Y. (1971). Basic research in speech and lateralization of language. *Bulletin of the Orton Society, 21*, 72–87.

Liberman, I. Y., & Shankweiler, D. (1991). Phonology and beginning reading: A tutorial. In L. Rieben & C. A. Perfetti (Eds.), *Learning to read: Basic research and its implications* (pp. 3–17). Hillsdale, NJ: Erlbaum.

Limpo, T., & Alves, R. A. (2013). Modeling writing development: Contribution of transcription and self-regulation to Portuguese students' text generation quality. *Journal of Educational Psychology, 105*, 401.

Lin, S. J., Monroe, B. W., & Troia, G. A. (2007). Development of writing knowledge in grades 2–8: A comparison of typically developing writers and their struggling peers. *Reading and Writing Quarterly, 23*, 207–230.

Lindamood, P., & Lindamood, P. D. (1998). *The Lindamood Phoneme Sequencing Program for Reading, Spelling, and Speech: The LiPS Program* (multimedia kit). Austin, TX: PRO-ED.

Linkersdorfer, J., Lonnemann, J., Lindberg, S., Hasselhorn, M., & Fiebach, C. J. (2012). Grey matter alterations co-localize with functional abnormalities in developmental dyslexia: An ALE meta-analysis. *PLOS ONE 7*(8), e43122.

Livingstone, M. S., Rosen, G. D., Drislane, F. W., & Galaburda, A. M. (1991). Physiological and anatomical evidence for a magnocellular defect in developmental dyslexia. *Proceedings of the National Academy of Sciences of the USA, 88*(18), 7943–7947.

Lloyd, J. W. (1980). Academic instruction and cognitive-behavior modification. *Exceptional Education Quarterly, 1*, 53–63.

Longcampe, M., Zerbata-Poudou, M.-T., & Velay, J.-L. (2005). The influence of writing practice on letter recognition in preschool children: A comparison between handwriting and typing. *Acta Psychologica, 119*, 67–79.

Lonigan, C. J., & Phillips, B. M. (2016). Response to instruction in preschool: Results of two randomized studies with children at significant risk of reading difficulties. *Journal of Educational Psychology, 108*(1), 114–129.

Lonigan, C. J., Phillips, B. M., Clancy, J. L., Landry, S. H., Swank, P. R., Assel, M., . . . Eisenberg, N. (2015). Impacts of a comprehensive school readiness curriculum for preschool children at risk for educational difficulties. *Child Development, 86*, 1773–1793.

Loras, H., Sigmundsson, H., Stensdotter, A.-K., & Talcott J. B. (2014). Postural control is not systematically related to reading skills: Implications for the assessment of balance as a risk factor for developmental dyslexia. *PLOS ONE 9*(6), e98224.

Lovegrove, W., Martin, F., & Slaghuis, W. (1986). A theoretical and experimental case for a visual deficit in specific reading disability. *Cognitive Neuropsychology, 3*, 225–267.

Lovett, B. J., & Lewandowski, L. J. (2006). Gifted students with learning disabilities: Who are they? *Journal of Learning Disabilities, 30*, 515–527.

Lovett, B. J., & Lewandowski, L. J. (2015). *Testing accommodations for students with disabilities: Research-based practice*. New York: American Psychological Association.

Lovett, M. W. (1987). A developmental approach to reading disability: Accuracy and speed criteria of normal and deficient reading skill. *Child Development, 58*, 234–260.

Lovett, M. W., Barron, R. W., & Frijters, J. C. (2013). Word identification difficulties in children and adolescents with reading disabilities: Intervention research findings. In H. L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 329–360). New York: Guilford Press.

Lovett, M. W., De Palma, M., & Frijters, J. (2012). Evaluating the efficacy of remediation for struggling readers in high school. *Journal of Learning Disabilities, 45*, 151–169.

Lovett, M. W., Frijters, J. C., Wolf, M. A., Steinbach, K. A., Sevcik, R. A., & Morris, R. D. (2017). Early intervention for children at risk for reading disabilities: The impact of grade at intervention and individual differences on intervention outcomes. *Journal of Educational Psychology, 109*(7), 889–914.

Lovett, M. W., Ransby, M. J., Hardwick, N., & Johns, M. S. (1989). Can dyslexia be treated?: Treatment-specific and generalized treatment effects in dyslexic children's response to remediation. *Brain and Language, 37*, 90–121.

Lovitt, T. C., & Curtiss, K. A. (1968). Effects of manipulating an antecedent event on mathematics response rate. *Journal of Applied Behavior Analysis, 1*, 329–333.

Luciano, M., Evans, D. M., Hansell, N. K., Medland, S. E., Montgomery, G. W., Martin, N. G., . . . Bates, T. C. (2013). A genome-wide association study for reading and language abilities in two population cohorts. *Genes, Brain and Behavior, 12*, 645–652.

Lukatela, G., & Turvey, M. T. (1998). Reading in two alphabets. *American Psychologist, 53*, 1057–1072.

Lukowski, S. L., Soden, B., Hart, S. A., Thompson, L. A., Kovas, Y., & Petrill, S. A. (2014). Etiological distinction of working memory components in relation to mathematics. *Intelligence, 47*, 54–62.

Lymberis, J., Christodoulou, J. A., O'Loughlin, P., Del Tufo, S., & Gabrieli, J. D. E. (2009). *Neural correlates of rapid automatized naming*. Paper delivered at the 39th annual meeting of the Society for Neuroscience, Chicago, IL. Retrieved from *http://web.mit.edu/tkp/www/Lymberis_etal_SfN-2009.pdf*.

Lyon, G. R., & Chhabra, V. (2004). The science of reading research. *Educational Leadership, 61*, 13–20.

Lyon, G. R., & Esterline, E. (2007). Advancing education through research: False starts, broken promises and light on the horizon. In M. Welsh-Ross & L. Fasig (Eds.), *Handbook on communicating and disseminating behavioral science* (pp. 317–340). Thousand Oaks, CA: SAGE.

Lyon, G. R., Fletcher, J. M., Shaywitz, S. E., Shaywitz, B. A., Torgesen, J. K., Wood, F. B., . . . Olson, R. (2001). Rethinking learning disabilities. In C. E. Finn, Jr., R. A. J. Rotherham, & C. R. Hokanson, Jr. (Eds.), *Rethinking special education for a new century* (pp. 259–287). Washington, DC: Thomas B. Fordham Foundation and Progressive Policy Institute.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia, 53*, 1–14.

Lyon, G. R., & Weiser, B. (2013). The state of the science in learning disabilities: Research impact on the field from 2001–2011. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 118–113). New York: Guilford Press.

Lyons, I. M., & Ansari, D. (2015). Numerical order processing in children: From reversing the distance-effect to predicting arithmetic. *Mind, Brain, and Education, 9*, 207–217.

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1–6. *Developmental Science, 17*, 714–726.

Ma, Y., Koyama, M. S., Milham, M. P., Castellanos, F. X., Quinn, B. T., Pardoe, H., . . . Blackmon, K. (2015). Cortical thickness abnormalities associated with dyslexia, independent of remediation status. *Neuroimage: Clinical, 7*, 177–186.

Macmann, G. M., & Barnett, D. W. (1985). Discrepancy score analysis: A computer simulation of classification stability. *Journal of Psychoeducational Assessment, 4*, 363–375.

Macmann, G. M., & Barnett, D. W. (1997). Myth of the master detective: Reliability of interpretations for Kaufman's "Intelligent Testing" approach to the WISC-III. *School Psychology Quarterly, 12*, 197–234.

Macmann, G. M., Barnett, D. W., Lombard, T. J., Belton-Kocher, E., & Sharpe, M. N. (1989). On the actuarial classification of children: Fundamental studies of classification agreement. *Journal of Special Education, 23*(2), 127–149.

Maisog, J. M., Einbinder, E. R., Flowers, D. L., Turkeltaub, P. E., & Eden, G. F. (2008). A meta-analysis of functional neuroimaging studies of dyslexia. *Annals of the New York Academy of Sciences, 1145*, 237–259.

Manis, F. R., Doi, L. M., & Bhadha, B. (2000). Naming speed, phonological awareness, and orthographic knowledge in second graders. *Journal of Learning Disabilities, 33*, 325–333.

Manis, F. R., Seidenberg, M. S., Doi, L. M., McBride-Chang, C., & Peterson, A. (1996). On the basis of two subtypes of developmental dyslexia. *Cognition, 58*, 157–195.

Mann, L. (1979). *On the trail of process*. New York: Grune & Stratton.

Manzo, K. K. (2008, May 14). Reading First doesn't help pupils "get it." *Education Week, 27*(36), 1.

Mapou, R. L. (2009). *Adult learning disabilities and ADHD: Research informed assessment*. New York: Oxford University Press.

Mapou, R. L. (2013). Process focused assessment of learning disabilities and ADHD in adults. In L. Ashen Dorf, R. Swenson, & D. J. Lisbon (Eds.), *The Boston Process approach to neuropsychological assessment: A practitioner's guide* (pp. 329–354). New York: Oxford University Press.

Maria, K. (1990). *Reading comprehension instruction: Issues and strategies*. Parkton, MD: York Press.

Marlow, A. J., Fisher, S. E., Richardson, A. J., Francks, C., Talcott, J. B., Monaco, A. P., . . . Cardon, L. R. (2001). Investigation of quantitative measures related to reading disability in a large sample of sib-pairs from the UK. *Behavior Genetics, 31*, 219–230.

Martin, A., Kronbichler, M., & Richlan, F. (2016). Dyslexic brain activation abnormalities in deep and shallow orthographies: A meta-analysis of 28 functional neuroimaging studies. *Human Brain Mapping, 37*, 2676–2699.

Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research, 80*(3), 300–335.

Mastropieri, M. A., & Scruggs, T. E. (1997). Best practices in promoting reading comprehension in students with learning disabilities: 1976 to 1996. *Remedial and Special Education, 18*, 197–214.

Matejko, A. A., & Ansari, D. (2015). Drawing connections between white matter and numerical and mathematical cognition: A literature review. *Neuroscience and Biobehavioral Reviews, 48*, 35–52.

Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., & Schatschneider, C. (2005). An evaluation of two reading interventions derived from diverse models. *Reading Research Quarterly, 40*, 148–183.

May, H., Gray, A., Sirinides, P., Goldsworthy, H., Armijo, M., Sam, C., & Tognatta, N. (2015). Year one results from the multisite randomized evaluation of the i3 scale-up of Reading Recovery. *American Educational Research Journal, 52*, 547–581.

May, H., Sirinides, P., Gray, A., & Goldsworthy, H. (2016). *Reading Recovery: An evaluation of the four-year i3 scale-up*. Philadelphia: Consortium for Policy Research in Education. Retrieved from *www.cpre.org/sites/default/files/reading_recovery_final_report.pdf.*

Mayes, S. D., Calhoun, S. L., & Crowell, E. W. (2000). Learning disabilities and ADHD overlapping spectrum disorders. *Journal of Learning Disabilities, 33*, 417–424.

Mazzocco, M. M., & Devlin, K. T. (2008). Parts and "holes": Gaps in rational number sense among children with vs. without mathematical learning disabilities. *Developmental Science, 11*, 681–691.

Mazzocco, M. M., & Myers, G. F. (2003). Complexities in identifying and defining mathematics learning disability in the primary school-age years. *Annals of Dyslexia, 53*, 218–253.

McDaniel, J., Sims, C., & Miskel, C. (2001). The national reading policy arena: Policy actors and perceived influence. *Educational Policy, 15*, 92–114.

McGrath, L. M., Pennington, B. F., Shanahan, A., Santerre-Lemmon, L. E., Barnard, H. D., Willcutt, E. G., . . . Olson, R. K. (2011). A multiple deficit model of reading disability and attention-deficit/hyperactivity disorder: Searching for shared cognitive deficits. *Journal of Child Psychology and Psychiatry, 52*(5), 547–557.

McGuiness, C., McGuiness, D., & McGuiness, G. (1996). Phono-Graphix: A new method for remediating reading difficulties. *Annals of Dyslexia, 46*, 73–96.

Medwell, J., & Wray, D. (2014). Handwriting automaticity: The search for performance thresholds. *Language and Education, 28*(1), 34–51.

Mehta, P., Foorman, B., Branum-Martin, L., & Taylor, W. (2005). Literacy as a unidimensional multilevel construct: Validation, sources of influence, and implications in a longitudinal study in grades 1 to 4. *Scientific Studies of Reading, 9*, 85–116.

Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin, 138*, 322–352.

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer": Evidence from a meta-analytic review. *Perspectives on Psychological Science, 11*(4), 512–534.

Menon, V. (2016). Working memory in children's math learning and its disruption in dyscalculia. *Current Opinion in Behavioral Sciences, 10*, 125–132.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Mahwah, NJ: Erlbaum.

Meyer, B. J. (1987). Following the author's top-level organization: An important skill for reading comprehension. In R. J. Tierney, P. R. Anders, & J. N. Mitchell (Eds.), *Understanding readers' understanding: Theory and practice* (pp. 59–76). New York: Routledge.

Meyer, M. S. (2002). Repeated reading: An old standard is revisited and renovated. *Perspectives, 28*, 15–18.

Meyler, A., Keller, T. A., Cherkassky, V. L., Gabrieli, J. D., & Just, M. A. (2008). Modifying the brain activation of poor readers during sentence comprehension with extended remedial instruction: A longitudinal study of neuroplasticity. *Neuropsychologia, 46*(10), 2580–2592.

Michaelsen, L. K., & Sweet, M. (2011). Team-based learning. *New Directions for Teaching and Learning, 128*, 41–51.

Miciak, J., Fletcher, J. M., & Stuebing, K. K. (2015a). Accuracy and validity of methods for identifying learning disabilities in a RTI service delivery framework. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *The handbook of response to intervention: The science and practice of multi-tiered systems of support* (2nd ed., pp. 421–440). New York: Springer.

Miciak, J., Fletcher, J. M., Stuebing, K. K., Vaughn, S., & Tolar, T. D. (2014a). Patterns of cognitive strengths and weaknesses: Identification rates, agreement, and validity for learning disabilities identification. *School Psychology Quarterly, 29*, 21–37.

Miciak, J., Stuebing, K. K., Vaughn, S., Roberts, G., Barth, A. E., & Fletcher, J. M. (2014b). Cognitive attributes of adequate and inadequate responders to reading intervention in middle school. *School Psychology Review, 43*, 407–427.

Miciak, J., Taylor, W. P., Denton, C. A., & Fletcher, J. M. (2015b). The effect of achievement test selection on identification of learning disabilities within patterns of strengths and weaknesses framework. *School Psychology Quarterly, 30*, 321–334.

Miciak, J., Williams, J. L., Taylor, W. P., Cirino, P. T., Fletcher, J. M., & Vaughn, S. (2016). Do processing patterns of strengths and weaknesses predict differential treatment response? *Journal of Educational Psychology, 108*, 898–1011.

Miles, T. R., & Haslum, M. N. (1986). Dyslexia: Anomaly or normal variation. *Annals of Dyslexia, 36*, 103–117.

Miller, A. C., Fuchs, D., Fuchs, L. S., Compton, D. L., Kearns, D., Zhang, W., . . . Kirchner, D. P. (2014). Behavioral attention: A longitudinal study of whether and how it influences the development of word-reading and reading comprehension among at-risk readers. *Journal of Research in Educational Effectiveness, 7*, 232–249.

Misra, M., Katzir, T., Wolf, M., & Poldrack, R. A. (2004). Neural systems for rapid automatized naming in skilled readers: Unraveling the RAN–reading relationship. *Scientific Studies of Reading, 8*, 241–256.

Moats, L. C. (2005). How spelling supports reading: And why it is more regular and predictable than you may think. *American Educator, 29*, 12–43.

Moats, L. C., & Farrell, M. L. (1999). Multisensory instruction. In J. Birsh (Ed.), *Multisensory teaching of basic language skills* (pp. 1–17). Baltimore: Brookes.

Mody, M., Studdert-Kennedy, M., & Brady, S. (1997). Speech perception deficits in poor readers: Auditory processing or phonological coding? *Journal of Experimental Child Psychology, 64*(2), 199–231.

Molfese, D. L. (2000). Predicting dyslexia at 8 years of age using neonatal brain responses. *Brain and Language, 72*, 238–245.

Molfese, P. J., Fletcher, J. M., & Denton, C. A. (2013). Adequate versus inadequate response to reading intervention: An event-related potentials assessment. *Developmental Neuropsychology, 38*, 534–549.

Moll, K., Gobel, S. M., & Snowling, M. J. (2015). Basic number processing in children with specific learning disabilities: Comorbidity of reading and mathematics disorders. *Child Neuropsychology, 21*, 399–417.

Moll, K., Kunze, S., Neuhoff, N., Bruder, J., & Schulte-Körne, G. (2014). Specific learning disorder: Prevalence and gender differences. *PLOS ONE, 9*(7), e103537.

Monzalvo, K., & Dehaene-Lambertz, G. (2013). How reading acquisition changes children's spoken language network. *Brain and Language, 127*, 356–363.

Monzalvo, K., Fluss, J., Billard, C., Dehaene, S., & Dehaene-Lambertz, G. (2012). Cortical networks for vision and language in dyslexic and normal children of variable socio-economic status. *NeuroImage, 61*, 258–274.

Morgan, P. L., Farkas, G., & Wu, Q. (2011). Kindergarten children's growth trajectories in reading and mathematics: Who falls increasingly behind? *Journal of Learning Disabilities, 44*, 472–488.

Morgan, W. P. (1896). A case of congenital word blindness. *British Medical Journal, 2*, 1378.

Morphy, P., & Graham, S. (2012). Word processing programs and weaker writers/readers: A meta-analysis of research findings. *Reading and Writing, 25*, 641–678.

Morris, R. D. (1988). Classification of learning disabilities: Old problems and new approaches. *Journal of Consulting and Clinical Psychology, 56*, 789–794.

Morris, R. D., Lovett, M. W., Wolf, M. A., Sevcik, R. A., Steinbach, K. A., Frijters, J. C., & Shapiro, M. (2012). Multiple-component remediation for developmental reading disabilities: IQ, socioeconomic status, and race as factors in remedial outcome. *Journal of Learning Disabilities, 45*, 99–127.

Morris, R. D., Stuebing, K. K., Fletcher, J. M., Shaywitz, S. E., Lyon, G. R., Shankweiler, D. P., . . . Shaywitz, B. A. (1998). Subtypes of reading disability: Variability around a phonological core. *Journal of Educational Psychology, 90*, 347–373.

Mullis, I., Campbell, J. R., & Farstrup, A. R. (1993). *NAEP 1992 reading report card for the nation and the states: Data from the national and trial state assessments.* Washington, DC: National Center for Education Statistics.

Murphy, L., & Pollatsek, A. (1994). Developmental dyslexia: Heterogeneity without discrete subgroups. *Annals of Dyslexia, 44*, 120–146.

Mussolin, C., Nys, J., Content, A., & Leybaert, J. (2014). Symbolic number abilities predict later approximate number system acuity in preschool children. *PLOS ONE, 9*(3), e91839.

Myers, C. A. (1978). Reviewing the literature on Fernald's technique of remedial reading. *Reading Teacher, 31*, 614–619.

Naglieri, J. A. (2010). The discrepancy–consistency approach to SLD identification using the PASS theory. In D. P. Flanagan & V. C. Alfonso (Eds.), *Essentials of specific learning disability identification* (pp. 145–172). New York: Wiley.

Naglieri, J. A., & Das, J. P. (1997). *Cognitive assessment system*. Chicago: Riverside.

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly, 20*, 233–253.

Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 269–284). New York: Routledge.

Nation, K. (2005). Children's reading comprehension difficulties. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 248–266). Oxford, UK: Blackwell.

Nation, K., Adams, J. W., Bowyer-Crane, A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology, 73*, 139–158.

Nation, K., Clarke, P., Marshall, C. M., & Durand, M. (2004). Hidden language impairments in children: Parallels between poor reading comprehension and specific language impairment? *Journal of Speech, Language, and Hearing Research, 47*, 199–211.

Nation, K., Cocksey, J., Taylor, J. S. H., & Bishop, D. V. M. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *Journal of Child Psychology and Psychiatry, 51*(9), 1031–1039.

Nation, K., & Hulme, C. (2011). Learning to read changes children's phonological skills: Evidence from a latent variable longitudinal study of reading and nonword repetition. *Developmental Science, 14*, 649–659.

Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition, 70*(1), B1–B13.

National Association of State Directors of Special Education. (2006). *Response to intervention: Policy considerations and implementation*. Alexandria, VA: Author.

National Center for Education Statistics. (2017). *Digest of education statistics*. Washington, DC: Institute for Educational Sciences.

National Center for Learning Disabilities. (2014). *The state of learning disabilities: Facts, trends, and emerging issues*. New York: Author.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: U.S. Department of Education.

National Council of Teachers of Mathematics. (2007). *The learning of mathematics*. Reston, VA: Author.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics*. Washington, DC: Author. Retrieved from *www.corestandards.org*.

National Institute of Child Health and Human Development (NICHD). (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.

National Joint Committee on Learning Disabilities (NJCLD). (1988). *Letter to NJCLD member organizations*. Author.

National Mathematics Advisory Panel. (2008). *The final report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.

Nelson, R. J., Benner, G. J., & Gonzalez, J. (2003). Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta-analytic review. *Learning Disabilities Research and Practice, 18*, 255–267.

Neuhaus, G., Foorman, B. R., Francis, D. J., & Carlson, C. D. (2001). Measures of information processing in rapid automatized naming (RAN) and their relation to reading. *Journal of Experimental Child Psychology, 78*, 359–373.

Ni, Y., & Zhou, Y. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist, 40*, 27–52.

Nicolson, R. I., & Fawcett, A. J. (2011). Dyslexia, dysgraphia, procedural learning and the cerebellum. *Cortex, 47*, 117–127.

Nicolson, R. I., Fawcett, A. J., & Dean, P. (2001). Developmental dyslexia: The cerebellar deficit hypothesis. *Trends in Neurosciences, 24*, 508–511.

Nigg, J. T. (2009). *What causes ADHD?: Understanding what goes wrong and why*. New York: Guilford Press.

Nikolopoulos, D., Goulandris, N., Hulme, C., & Snowling, M. J. (2006). The cognitive bases of learning to read and spell in Greek: Evidence from a longitudinal study. *Journal of Experimental Child Psychology, 94*, 1–17.

Noble, K. G., Wolmetz, M. G., Ochs, L. G., Farah, M. G., & McCandliss, B. D. (2006). Brain–behavior relationships in reading acquisition are modulated by socioeconomic factors. *Developmental Science, 9*, 632–654.

No Child Left Behind Act. (2001). P. L. 107-110, 20 U. S. C§310.

Noël, M. P. (2009). Counting on working memory when learning to count and to add: A preschool study. *Developmental Psychology, 45*(6), 1630–1643.

Norton, E. S., Black, J. M., Stanley, L. M., Tanaka, H., Gabrieli, J. D., Sawyer, C., & Hoeft, F. (2014). Functional neuroanatomical evidence for the double-deficit hypothesis of developmental dyslexia. *Neuropsychologia, 61*, 235–246.

Norton, E. S., & Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annual Review of Psychology, 63*, 427–452.

Nöthen, M. M., Schulte-Körne, G., Grimm, T., Cichon, S., Vogt, I. R., Müller-Myhsok, B., . . . Remschmidt, H. (1999). Genetic linkage analysis with dyslexia: Evidence for linkage of spelling disability to chromosome 15. *European Child and Adolescent Psychiatry, 8*, S56–S59.

Oakhill, J. (1993). Children's difficulties in reading comprehension. *Educational Psychology Review, 5*(3), 223–237.

Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading, 16*(2), 91–121.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word-reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes, 18*, 443–468.

Oakhill, J. V., & Kyle, F. (2000). The relation between phonological awareness and working memory. *Journal of Experimental Child Psychology, 75*(2), 152–164.

Oakland, T., Black, J. L., Stanford, G., Nussbaum, N. L., & Balise, R. R. (1998). An evaluation of the dyslexia training program: A multisensory method for promoting reading in students with reading disabilities. *Journal of Learning Disabilities, 31*, 140–147.

O'Connor, R. E., & Fuchs, L. S. (2013). Responsiveness to intervention in the elementary grades: Implications for early childhood education. In V. Buysse, E. Peisner-Feinberg, & J. Cantler (Eds.), *Handbook of response to intervention (RTI) in early childhood education* (pp. 41–56). Baltimore: Brookes.

O'Connor, R. E., & Jenkins, J. R. (2013). Cooperative learning for students with disabilities. In H. L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 489–506).

New York: Guilford Press.
O’Connor, R. E., White, A., & Swanson, H. L. (2007). Repeated reading versus continuous reading: Influences on reading fluency and comprehension. *Exceptional Children, 74*(1), 31–46.
Odegard, T. N., Ring, J., Smith, S., Biggan, J., & Black, J. (2008). Differentiating the neural response to intervention in children with developmental dyslexia. *Annals of Dyslexia, 58*, 1–14.
Ogle, J. W. (1867). Aphasia and agraphia. *Report of the Medical Research Council of Saint George’s Hospital, 2*, 83–122.
Olive, T. (2014). Toward a parallel and cascading model of the writing system: A review of research on writing processes coordination. *Journal of Writing Research, 6*, 141–171.
Olson, R. K. (2011, Winter). Evaluation of Fast ForWord Language® effects on language and reading. *Perspectives on Language and Literacy*, pp. 21–24.
Olson, R. K., Forsberg, H., Gayan, J., & DeFries, J. C. (1999). A behavioral–genetic analysis of reading disabilities and component processes. In R. M. Klein & P. A. McMullen (Eds.), *Converging methods for understanding reading and dyslexia* (pp. 133–153). Cambridge, MA: MIT Press.
Olson, R. K., Hulslander, J., Christopher, M., Keenan, J. M., Wadsworth, S. J., Willcutt, E. G., . . . DeFries, J. C. (2013). Genetic and environmental influences on writing and their relations to language and reading. *Annals of Dyslexia, 63*, 25–43.
Olson, R. K., Keenan, J. M., Byrne, B., & Samuelsson, S. (2014). Why do children differ in their development of reading and related skills? *Scientific Studies of Reading, 18*, 38–54.
Olulade, O. A., Napoliello, E. M., & Eden, G. F. (2013). Abnormal visual motion processing is not a cause of dyslexia. *Neuron, 79*(1), 180–190.
Orton, S. (1928). Specific reading disability—Strephosymbolia. *Journal of the American Medical Association, 90*, 1095–1099.
Owens, L. L. (2004). *The effects of the Handwriting Without Tears program on the handwriting of students in inclusion classrooms*. Master’s thesis, Virginia Commonwealth University, Richmond, VA.
Packard, J. L., Chen, X., Li, W., Wu, X., Gaffney, J. S., Li, H., & Anderson, R. C. (2006). Explicit instruction in orthographic structure and word morphology helps Chinese children learn to write characters. *Reading and Writing, 19*(5), 457–487.
Palinscar, A., & Brown, A. (1985). Reciprocal teaching: A means to a meaningful end. In J. Osborn, P. T. Wilson, & R. C. Anderson (Eds.), *Reading education: Foundations for a literate America* (pp. 66–87). Lexington, MA: Heath.
Papanicolaou, A. C. (Ed.). (2017). *The Oxford handbook of functional brain imaging in neuropsychology and cognitive neurosciences*. New York: Oxford University Press.
Park, J., & Brannon, E. M. (2013). Training the approximate number system improves math proficiency. *Psychological Science, 24*, 2013–2019.
Pastor, P. N., & Reuben, C. A. (2008). Diagnosed attention deficit hyperactivity disorder and learning disability: United States, 2004–2006. *Vital Health Statistics, 10*, 1–14.
Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2009). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest, 9*(3), 105–119.
Pastor, P. N., & Reuben, C. A. (2008). Diagnosed attention deficit hyperactivity disorder and learning disability: United States, 2004–2006. *Vital Health Statistics, 10*, 1–14.
Paulesu, E., Danelli, L., & Berlingeri, M. (2014). Reading the dyslexic brain: Multiple dysfunctional routes revealed by a new meta-analysis of PET and fMRI activation studies. *Frontiers in Human Neuroscience, 8*, 830.
Pearson. (2009). *Wechsler Individual Achievement Test—Third Edition*. San Antonio, TX: Author.
Pearson. (2010). *Stanford Achievement Test Series, Tenth Edition*. San Antonio, TX: Author.

Pearson, P. D. (2010). Reading First: Hard to live with—or without. *Journal of Literacy Research, 42*, 100–108.
Peng, P., Barnes, M. A., Wang, C., Wang, W., Swanson, H. L., Dardick, W., & Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin, 144*, 48–76.
Peng, P., & Fuchs, D. (2016). A meta-analysis of working memory deficits in children with learning difficulties: Is there a difference between verbal domain and numerical domain? *Journal of Learning Disabilities, 49*, 3–20.
Peng, P., Namkung, J., Barnes, M., & Sun, C. (2015). A meta-analysis of mathematics and working memory: Moderating effects of working memory domain, type of mathematics skill, and sample characteristics. *Journal of Educational Psychology, 108*, 455–473.
Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition, 101*, 385–413.
Pennington, B. F. (2009). *Diagnosing learning disorders: A neuropsychological framework* (2nd ed.). New York: Guilford Press.
Pennington, B. F. (2011, Winter). Controversial therapies for dyslexia. *Perspectives on Language and Literacy*, pp. 7–8.
Pennington, B. F., & Bishop, D. V. M. (2009). Relations among speech, language, and reading disorders. *Annual Review of Psychology, 60*, 283–306.
Pennington, B. F., Gilger, J. W., Olson, R. K., & DeFries, J. C. (1992). The external validity of age- versus IQ-discrepancy definitions of reading disability: Lessons from a twin study. *Journal of Learning Disabilities, 25*, 562–573.
Pennington, B. F., McGrath, L. M., Rosenberg, J., Barnard, H., Smith, S. D., Willcutt, E. G., . . . Olson, R. K. (2009). Gene $\frac{3}{2}$ environment interactions in reading disability and attention-deficit/hyperactivity disorder. *Developmental Psychology, 45*, 77–89.
Pennington, B. F., & Olson, R. K. (2005). Genetics of dyslexia. *The science of reading: A handbook* (pp. 453–472). Oxford, UK: Blackwell.
Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*, 357–383.
Perfetti, C., & Adlof, S. M. (2012). Reading comprehension: A conceptual framework from word meaning to text meaning. In J. Sabatini, E. Albro, & T. O'Reilly (Eds.), *Measuring up: Advances in how we assess reading ability* (pp. 3–20). Lanham, MD: Rowan & Littlefield. Education.
Perfetti, C. A., Landi, N., Oakhill, J., & Snowling, M. J. (2005). The acquisition of reading comprehension skill. In C. Hulme (Ed.), *The science of reading: A handbook* (pp. 227–247). Malden, MA: Blackwell.
Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18*, 22–37.
Perrachione, T. K., Del Tufo, S. N., Winter, R., Murtagh, J., Cyr, A., Chang, P., . . . Gabrieli, J. D. (2016). Dysfunction of rapid neural adaptation in dyslexia. *Neuron, 92*(6), 1383–1397.
Perrone-Bertolotti, M., Vidal, J. R., De Palma, L., Hamamé, C. M., Ossandon, T., Kahane, P., . . . Lachaux, J. P. (2014). Turning visual shapes into sounds: Early stages of reading acquisition revealed in the ventral occipitotemporal cortex. *NeuroImage, 90*, 298–307.
Peters, L., & De Smedt, B. (2017, May 17). Arithmetic in the developing brain: A review of brain imaging studies. *Developmental Cognitive Neuroscience*. [Epub ahead of print]
Peterson, R. L., Boada, R., McGrath, L. M., Willcutt, E. G., Olson, R. K., & Pennington, B. F. (2017). Cognitive prediction of reading, math, and attention: Shared and unique influences. *Journal of Learning Disabilities, 50*, 408–421.
Peterson, R. L., & Pennington, B. F. (2012). Developmental dyslexia. *The Lancet, 379*, 1997–2007.

Petrill, S. A., Deater-Deckard, K., Thompson, L. A., DeThorne, L. S., & Schatschneider, C. (2006). Reading skills in early readers: Genetic and shared environmental influences. *Journal of Learning Disabilities, 39*, 48–55.

Petrill, S. A., Deater-Deckard, K., Thompson, L. A., Schatschneider, C., DeThorne, L. S., & Vandenbergh, D. J. (2007). Longitudinal genetic analysis of early reading: The Western Reserve Reading Project. *Reading and Writing, 20*, 127–146.

Petrill, S. A., Hart, S. A., Harlaar, N., Logan, J., Justice, L. M., Schatschneider, C., . . . Cutting, L. (2010). Genetic and environmental influences on the growth of early reading skills. *Journal of Child Psychology and Psychiatry, 51*, 660–667.

Petrill, S. A., Kovas, Y., Hart, S. A., Thompson, L. A., & Plomin, R. (2009). The genetic and environmental etiology of high math performance in 10-year-old twins. *Behavior Genetics, 39*, 371–379.

Petrill, S., Logan, J., Hart, S., Vincent, P., Thompson, L., Kovas, Y., & Plomin, R. (2012). Math fluency is etiologically distinct from untimed math performance, decoding fluency, and untimed reading performance: Evidence from a twin study. *Journal of Learning Disabilities, 45*, 371–381.

Peverly, S. T. (2006). The importance of handwriting speed in adult writing. *Developmental Neuropsychology, 29*, 197–216.

Peverly, S. T., Garner, J. K., & Vekaria, P. C. (2014). Both handwriting speed and selective attention are important to lecture note-taking. *Reading and Writing, 27*, 1–30.

Phillips, B. M., & Lonigan, C. J. (2005). Social correlates of emergent literacy. In M. J. Snowling & C. Hulme (Eds.), *The science of reading handbook* (pp. 173–204). Oxford, UK: Blackwell.

Piazza, M., Pica, P., Izard, V., Spelke, E. S., & Dehaene, S. (2013). Education enhances the acuity of the nonverbal approximate number system. *Psychological Science, 24*, 1037–1043.

Pike, M., Swank, P., Taylor, H., Landry, S., & Barnes, M. A. (2013). Effect of preschool working memory, language, and narrative abilities on inferential comprehension at school-age in children with spina bifida myelomeningocele and typically developing children. *Journal of the International Neuropsychological Society, 19*(4), 390–399.

Pimperton, H., & Nation, K. (2010). Suppressing irrelevant information from working memory: Evidence for domain-specific deficits in poor comprehenders. *Journal of Memory and Language, 62*(4), 380–391.

Planton, S., Jucla, M., Roux, F. E., & Démonet, J. F. (2013). The "handwriting brain": A meta-analysis of neuroimaging studies of motor versus orthographic processes. *Cortex, 49*, 2772–2787.

Planton, S., Longcamp, M., Péran, P., Démonet, J. F., & Jucla, M. (2017). How specialized are writing-specific brain regions?: An fMRI study of writing, drawing and oral spelling. *Cortex, 88*, 66–80.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word-reading: Computational principles in quasi-regular domains. *Psychological Review, 103*, 56–115.

Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin, 131*, 592–617.

Plourde, V., Boivin, M., Forget-Dubois, N., Brendgen, M., Vitaro, F., Marino, C., . . . Dionne, G. (2015). Phenotypic and genetic associations between reading comprehension, decoding skills, and ADHD dimensions: Evidence from two population-based studies. *Journal of Child Psychology and Psychiatry, 56*, 1074–1082.

Pollack, C., Luk, G., & Christodoulou, J. A. (2015). A meta-analysis of functional reading systems in typically developing and struggling readers across different alphabetic languages. *Frontiers in Psychology, 6*, 191.

Powell, S. R., & Fuchs, L. S. (2015). Intensive intervention in mathematics. *Learning Disabilities*

Research and Practice, 30, 182–192.
Powell, S. R., & Stecker, P. M. (2014). Using data-based individualization to intensify mathematics intervention for students with disabilities. *Teaching Exceptional Children, 46*, 31–37.
Pressley, M. (2006). *Reading instruction that works* (3rd ed.). New York: Guilford Press.
Preston, J. L., Molfese, P. J., Frost, S. J., Mencl, W. E., Fulbright, R. K., Hoeft, F., . . . Pugh, K. R. (2016). Print–speech convergence predicts future reading outcomes in early readers. *Psychological Science, 27*, 75–84.
Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage, 62*, 815–827.
Price, G. R., Mazzocco, M. M., & Ansari, D. (2013). Why mental arithmetic counts: Brain activation during single digit arithmetic predicts high school math scores. *Journal of Neuroscience, 33*, 156–163.
Protopapas, A., Fakou, A., Drakopoulou, S., Skaloumbakas, C., & Mouzaki, A. (2013). What do spelling errors tell us?: Classification and analysis of errors made by Greek schoolchildren with and without dyslexia. *Reading and Writing, 26*, 615–646.
Protopapas, A., Simos, P. G., Sideridis, G. D., & Mouzaki, A. (2012). The components of the simple view of reading: A confirmatory factor analysis. *Reading Psychology, 33*, 217–240.
Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., . . . Shaywitz, B. A. (2000). Functional neuroimaging studies of reading and reading disability (developmental dyslexia). *Mental Retardation and Developmental Disabilities Research Reviews, 6*, 207–213.
Puranik, C. S., & Al Otaiba, S. (2012). Examining the contribution of handwriting and spelling to written expression in kindergarten children. *Reading and Writing, 25*, 1523–1546.
Qin, S., Cho, S., Chen, T., Rosenberg-Lee, M., Geary, D. C., & Menon, V. (2014). Hippocampal–neocortical functional reorganization underlies children's cognitive development. *Nature Neuroscience, 17*, 1263–1269.
Quinn, J. M., & Wagner, R. K. (2015). Gender differences in reading impairment and in the identification of impaired readers: Results from a large-scale study of at-risk readers. *Journal of Learning Disabilities, 48*(4), 433–445.
Quinn, J. M., Wagner, R. K., Petscher, Y., & Lopez, D. (2015). Developmental relations between vocabulary knowledge and reading comprehension: A latent change score modeling study. *Child Development, 86*, 159–175.
Raberger, T., & Wimmer, H. (2003). On the automaticity/cerebellar deficit hypothesis of dyslexia: Balancing and continuous rapid naming in dyslexic and ADHD children. *Neuropsychologia, 41*(11), 1493–1497.
Racsmány, M., Szőllősi, Á., & Bencze, D. (2017, June 29). Retrieval practice makes procedure from remembering: an automatization account of the testing effect. *Journal of Experimental Psychology. Learning, Memory, and Cognition*. [Epub ahead of print]
Raghubar, K., Barnes, M. A., & Hecht, S. (2010). Working memory and mathematics: A review of developmental, individual difference and cognitive approaches. *Learning and Individual Differences, 20*, 110–122.
Raghubar, K., Cirino, P., Barnes, M. A., Ewing-Cobbs, L., Fuchs, L., & Fletcher, J. M. (2009). Errors in multi-digit arithmetic and behavioral attention in children with math difficulties. *Journal of Learning Disabilities, 42*, 356–371.
Ramus, F. (2003). Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology, 13*, 212–218.
Ramus, F., Pidgeon, E., & Frith, U. (2003). The relationship between motor control and phonology in dyslexic children. *Journal of Child Psychology and Psychiatry, 44*, 712–722.

Raschle, N. M., Becker, B. L. C., Smith, S., Fehlbaum, L. V., Wang, Y., & Gaab, N. (2015). Investigating the influences of language delay and/or familial risk for dyslexia on brain structure in 5-year-olds. *Cerebral Cortex, 27*, 764–776.

Raschle, N. M., Chang, M., & Gaab, N. (2011). Structural brain alterations associated with dyslexia predate reading onset. *NeuroImage, 57*, 742–749.

Raschle, N. M., Zuk, J., & Gaab, N. (2012). Functional characteristics of developmental dyslexia in left-hemispheric posterior brain regions predate reading onset. *Proceedings of the National Academy of Sciences of the USA, 109*, 2156–2161.

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.

Raskind, W. H., Hsu, L., Berninger, V. W., Thomson, J. B., & Wijsman, E. M. (2000). Familial aggregation of dyslexia phenotypes. *Behavior Genetics, 30*, 385–396.

Raskind, W. H., Igo, R. P., Jr., Chapman, N. H., Berninger, V. W., Thomson, J. B., Matsushita, M., . . . Wijsman, E. M. (2005). A genome scan in multigenerational families with dyslexia: Identification of a novel locus on chromosome 2q that contributes to phonological decoding efficiency. *Molecular Psychiatry, 10*, 699–711.

Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. *Journal of Experimental Child Psychology, 91*, 137–157.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2002). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest, 2*, 31–74.

Rayner, K., Pollatsek, A., & Bilsky, A. B. (1995). Can a temporal processing deficit account for dyslexia? *Psychonomic Bulletin Review, 2*, 501–507.

Rayner, K., Schotter, E. R., Masson, M. E., Potter, M. C., & Treiman, R. (2016). So much to read, so little time: How do we read, and can speed reading help? *Psychological Science in the Public Interest, 17*, 4–34.

Reading First Advisory Committee. (2008). Response to the Reading First impact study. Retrieved January 5, 2010, from *www.cd.gov/programs/readingfirst/statement.pdf*.

Reed, M. A. (1989). Speech perception and the discrimination of brief auditory cues in reading disabled children. *Journal of Experimental Child Psychology, 48*(2), 270–292.

Reich, J., & Grigorenko, E. L. (2012). The effect of language and orthography on writing disabilities and the necessity for cross-linguistic research. In E. L. Grigorenko, E. Mambrino, & D. Preiss (Eds.), *Writing: A mosaic of new perspectives* (pp. 345–369). New York: Psychology Press.

Reynolds, C. R., & Shaywitz, S. E. (2009). Response to intervention: Ready or not? Or watch them fail. *School Psychology Quarterly, 24*, 130–145.

Reynolds, D., Nicolson, R. I., & Hambly, H. (2003). Evaluation of an exercise-based treatment for children with reading difficulties. *Dyslexia, 9*, 48–71.

Rezaie, R., Simos, P. G., Fletcher, J. M., Cirino, P. T., Vaughn, S., & Papanicolaou, A. C. (2011a). Engagement of temporal lobe regions predicts response to educational interventions in adolescent struggling readers. *Developmental Neuropsychology, 36*, 869–888.

Rezaie, R., Simos, P. G., Fletcher, J. M., Cirino, P. T., Vaughn, S., & Papanicolaou, A. C. (2011b). Temporo–parietal brain activity as a longitudinal predictor of response to educational interventions among middle school struggling readers. *Journal of the International Neuropsychological Society, 17*, 875–885.

Richards, T. L., Aylward, E. H., Berninger, V. W., Field, K. M., Grimme, A. C., Richards, A. L., & Nagy, W. (2006). Individual fMRI activation in orthographic mapping and morpheme mapping after orthographic or morphological spelling treatment in child dyslexics. *Journal of Neurolinguistics, 19*, 56–86.

Richards, T. L., & Berninger, V. W. (2008). Abnormal fMRI connectivity in children with dyslexia during a phoneme task: Before but not after treatment. *Journal of Neurolinguistics, 21*(4), 294–304.

Richards, T. L., Berninger, V. W., Stock, P., Altemeier, L., Trivedi, P., & Maravilla, K. R. (2009a). Functional magnetic resonance imaging sequential-finger movement activation differentiating good and poor writers. *Journal of Clinical and Experimental Neuropsychology, 31*(8), 967–983.

Richards, T., Berninger, V., & Fayol, M. (2009b). fMRI activation differences between 11-year-old good and poor spellers' access in working memory to temporary and long-term orthographic representations. *Journal of Neurolinguistics, 22*, 327–353.

Richards, T. L., Berninger, V., Nagy, W., Parsons, A., Field, K., & Richards, A. (2005). Brain activation during language task contrasts in children with and without dyslexia: Inferring mapping processes and assessing response to spelling instruction. *Educational and Child Psychology, 22*, 62–80.

Richards, T. L., Corina, D., Serafini, S., Steury, K., Echelard, D. R., Dager, S. R., . . . Berninger, V. W. (2000). Effects of a phonologically driven treatment for dyslexia on lactate levels as measured by proton MR spectroscopic imaging. *American Journal of Neuroradiology, 21*, 916–922.

Richlan, F., Kronbichler, M., & Wimmer, H. (2009). Functional abnormalities in the dyslexic brain: A quantitative meta-analysis of neuroimaging studies. *Human Brain Mapping, 30*, 3299–3308.

Richlan, F., Kronbichler, M., & Wimmer, H. (2011). Meta-analyzing brain dysfunctions in dyslexic children and adults. *NeuroImage, 56*, 1735–1742.

Richlan, F., Kronbichler, M., & Wimmer, H. (2013). Structural abnormalities in the dyslexic brain: A meta-analysis of voxel-based morphometry studies. *Human Brain Mapping, 34*(11), 3055–3065.

Rimrodt, S. L., Clements-Stephens, A. M., Pugh, K. R., Courtney, S. M., Gaur, P., Pekar, J. J., & Cutting, L. E. (2009). Functional MRI of sentence comprehension in children with dyslexia: Beyond word recognition. *Cerebral Cortex, 19*(2), 402–413.

Ritchey, K. D., & Coker, D. L. (2013). An investigation of the validity and utility of two curriculum-based measurement writing tasks. *Reading and Writing Quarterly, 29*, 89–119.

Ritchey, K. D., & Goeke, J. L. (2006). Orton–Gillingham and Orton–Gillingham-based reading instruction: A review of the literature. *Journal of Special Education, 40*, 171–183.

Ritchie, S. J., & Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological Science, 24*, 1301–1308.

Rivera, D., & Smith, D. D. (1988). Using a demonstration strategy to teach midschool students with learning disabilities how to compute long division. *Journal of Learning Disabilities, 21*, 66–69.

Rivera, S. M., Reiss, A. L., Eckert, M. A., & Menon, V. (2005). Developmental changes in mental arithmetic: Evidence for increased functional specialization in the left inferior parietal cortex. *Cerebral Cortex, 15*, 1779–1790.

Roberts, G., Rane, S., Fall, A. M., Denton, C. A., Fletcher, J. M., & Vaughn, S. (2015). The impact of intensive reading intervention on level of attention in middle school students. *Journal of Clinical Child and Adolescent Psychology, 44*, 942–953.

Robison, M. K., & Unsworth, N. (2017). Working memory capacity, strategic allocation of study time, and value-directed remembering. *Journal of Memory and Language, 93*, 231–244.

Rochelle, K. S., & Talcott, J. B. (2006). Impaired balance in developmental dyslexia?: A meta-analysis of the contending evidence. *Journal of Child Psychology and Psychiatry, 47*, 1159–1166.

Rodgers, B. (1983). The identification and prevalence of specific reading retardation. *British Journal of Educational Psychology, 53*, 369–373.

Roeske, D., Ludwig, K. U., Neuhoff, N., Becker, J., Bartling, J., Bruder, J., . . . Schulte-Körne, G. (2011). First genome-wide association scan on neurophysiological endophenotypes points to trans-regulation effects on SLC2A3 in dyslexic children. *Molecular Psychiatry, 16*, 97–107.

Rogosa, D. (1995). Myths and methods: Myths about longitudinal research. In J. M. Gottman (Ed.), *The* *analysis of change* (pp. 3–66). Mahwah, NJ: Erlbaum.

Romani, C., Olson, A., & Di Betta, A. M. (2005). Spelling disorders. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 431–448). Oxford, UK: Blackwell.

Rosenberg-Lee, M., Ashkenazi, S., Chen, T., Young, C. B., Geary, D. C., & Menon, V. (2015). Brain hyper-connectivity and operation-specific deficits during arithmetic problem-solving in children with developmental dyscalculia. *Developmental Science, 18*(3), 351–372.

Rosenblum, S., & Dror, G. (2017). Identifying developmental dysgraphia: Characteristics utilizing handwriting classification methods. *IEEE Transactions on Human–Machine Systems, 47*, 293–298.

Rourke, B. P. (1985). *Neuropsychology of learning disabilities: Essentials of subtype analysis*. New York: Guilford Press.

Rourke, B. P. (1989). *Nonverbal learning disabilities: The syndrome and the model*. New York: Guilford Press.

Rourke, B. P. (1993). Arithmetic disabilities specific and otherwise: A neuropsychological perspective. *Journal of Learning Disabilities, 26*, 214–226.

Rourke, B. P., & Finlayson, M. A. J. (1978). Neuropsychological significance of variations in patterns of academic performance: Verbal and visual–spatial abilities. *Journal of Pediatric Psychology, 3*, 62–66.

Rousselle, L., & Noël, M. P. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs. non-symbolic number magnitude. *Cognition, 102*, 361–395.

Rubenstein, K., Matsushita, M., Berninger, V. W., Raskind, W. H., & Wijsman, E. M. (2011). Genome scan for spelling deficits: Effects of verbal IQ on models of transmission and trait gene localization. *Behavior Genetics, 41*(1), 31–42.

Rumsey, J. M., Horwitz, B., Donohue, B. C., Nace, K., Maisog, J. M., & Andreason, P. (1997). Phonological and orthographic components of word recognition: A PET–rCBF study. *Brain, 120*(5), 739–759.

Rutter, M. (1982). Syndromes attributed to "minimal brain dysfunction" in childhood. *American Journal of Psychiatry, 139*, 21–33.

Rutter, M., Caspi, A., Fergusson, D., Horwood, J., Goodman, R., Maughn, B., . . . Carroll, J. (2004). Sex differences in developmental reading disability: New findings from 4 epidemiological studies. *Journal of the American Medical Association, 291*, 2007–2012.

Rutter, M., & Yule W. (1975). The concept of specific reading retardation. *Journal of Child Psychology and Psychiatry, 16*, 181–197.

Saenz, L., Fuchs, L. S., & Fuchs, D. (2005). Effects of peer-assisted learning strategies on English language learners: A randomized controlled study. *Exceptional Children, 71*, 231–247.

Sandmel, K., & Graham, S. (2011). The process writing approach: A meta-analysis. *Journal of Educational Research, 104*, 396–407.

Santangelo, T., & Graham, S. (2016). A comprehensive meta-analysis of handwriting instruction. *Educational Psychology Review, 28*, 225–265.

Savage, R. S., Frederickson, N., Goodwin, R., Patni, U., Smith, N., & Tuersley, L. (2005). Relationships among rapid digit naming, phonological processing, motor automaticity, and speech perception in poor, average, and good readers and spellers. *Journal of Learning Disabilities, 38*, 12–28.

Savage, R. S., Lavers, N., & Pillay, V. (2007). Working memory and reading difficulties: What we know and what we don't know about the relationship. *Educational Psychology Review, 19*, 185–221.

Saygin, Z. M., Norton, E. S., Osher, D. E., Beach, S. D., Cyr, A. B., Ozernov-Palchik, O., . . . Gabrieli, J. D. (2013). Tracking the roots of reading ability: White matter volume and integrity correlate with phonological awareness in prereading and early-reading kindergarten children. *Journal of*

Neuroscience, 33, 13251–13258.
Scammacca, N. K., Roberts, G. J., Cho, E., Williams, K. J., Roberts, G., Vaughn, S. R., & Carroll, M. (2016). A century of progress: Reading interventions for students in grades 4–12, 1914–2014. *Review of Educational Research, 86*, 756–800.
Scammacca, N., Roberts, G., Vaughn. S., Edmonds, M., Wexler, J., Reutebuch, C. K., & Torgesen, J. K. (2007), *Interventions for adolescent struggling readers: A meta-analysis with implications for practice*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.
Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4–12, 1980–2011. *Journal of Learning Disabilities, 48*, 369–390.
Scerri, T. S., & Schulte-Körne, G. (2010). Genetics of developmental dyslexia. *European Child and Adolescent Psychiatry, 19*, 179–197.
Schalock, R. L., Borthwick-Duffy, S. A., Bradley, V. J., Buntinx, W. H., Coulter, D. L., Craig, E. M., . . . Shogren, K. A. (2010). *Intellectual disability: Definition, classification, and systems of supports* (11th ed.). Washington, DC: American Association on Intellectual and Developmental Disabilities.
Schatschneider, C., Carlson, C. D., Francis, D. J., Foorman, B. R., & Fletcher, J. M. (2002). Relationships of rapid automatized naming and phonological awareness in early reading development: Implications for the double-deficit hypothesis. *Journal of Learning Disabilities, 35*, 245–256.
Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology, 96*, 265–282.
Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science, 20*, e12372.
Schneider, W. J., & Kaufman, A. (2017). Let's not do away with comprehensive cognitive assessments just yet. *Archives of Clinical Neuropsychology, 32*, 8–20.
Schneps, M. H., Brockmole, J. R., Sonnert, G., & Pomplun, M. (2012). History of reading struggles linked to enhanced learning in low spatial frequency scenes. *PLOS ONE, 7*(4), e35724.
Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock–Johnson IV Tests of Achievement*. Rolling Meadows, IL: Riverside.
Schulte-Körne, G. (2001). Genetics of reading and spelling disorder. *Journal of Child Psychology and Psychiatry, 42*, 985–997.
Schulte-Körne, G., Deimel, W., Müller, K., Gutenbrunner, C., & Remschmidt, H. (1996). Familial aggregation of spelling disability. *Journal of Child Psychology and Psychiatry, 37*, 817–822.
Schultz, R. T., Cho, N. K., Staib, L. H., Kier, L. E., Fletcher, J. M., Shaywitz, S. E., et al. (1994). Brain morphology in normal and dyslexic children: The influence of sex and age. *Annals of Neurology, 35*, 732–742.
Schumaker, J. B., Deshler, D. D., & McKnight, P. (2002). Ensuring success in the secondary general education curriculum through the use of teaching routines. In M. A. Shinn, H. M. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems: II. Preventive and remedial approaches* (pp. 791–823). Bethesda, MD: National Association of School Psychologists.
Schurz, M., Wimmer, H., Richlan, F., Ludersdorfer, P., Klackl, J., & Kronbichler, M. (2015). Resting-state and task-based functional brain connectivity in developmental dyslexia. *Cerebral Cortex, 25*, 3502–3514.
Scientific Learning Corporation. (1998). *Fast ForWord Language* [Computer software]. Berkeley, CA: Author.
Seabaugh, G. O., & Schumaker, J. B. (1993). The effects of self-regulation training on the academic productivity of secondary students with learning problems. *Journal of Behavioral Education, 4*,

109–133.
Seethaler, P. M., Fuchs, L. S., Star, J. R., & Bryant, J. (2011). The cognitive predictors of calculational skill with whole versus rational numbers: An exploratory study. *Learning and Individual Differences, 21*, 536–542.
Seidenberg, M. (2017). *Language at the speed of sight*. New York: Basic Books.
Shalev, R. S., Auerbach, J., Manor, O., & Gross-Tsur, V. (2000). Developmental dyscalculia: Prevalence and prognosis. *European Child and Adolescent Psychiatry, 9*, 58–64.
Shalev, R. S., Manor, O., Auerbach, J., & Gross-Tsur, V. (1998). Persistence of developmental dyscalculia: What counts? Results from a 3-year prospective follow-up study. *Journal of Pediatrics, 133*, 358–362.
Shalev, R. S., Manor, O., & Gross-Tsur, V. (2005). Developmental dyscalculia: A prospective six-year follow-up. *Developmental Medicine and Child Neurology, 47*, 121–125.
Shalev, R. S., Manor, O., Kerem, B., Ayali, M., Badichi, N., Friedlander, Y., & Gross-Tsur, V. (2001). Developmental dyscalculia is a familial learning disability. *Journal of Learning Disabilities, 34*, 59–65.
Shanahan, T., & Barr, R. (1995). Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk learners. *Reading Research Quarterly, 30*, 958–996.
Shanahan, T., & Lomax, R. G. (1986). An analysis and comparison of theoretical models of the reading–writing relationship. *Journal of Educational Psychology, 78*, 116–123.
Shanahan, T., & Lomax, R. G. (1988). A developmental comparison of three theoretical models of the reading–writing relationship. *Research in the Teaching of English, 22*, 196–212.
Shankweiler, D., & Crain, S. (1986). Language mechanisms and reading disorder: A modular approach. *Cognition, 24*(1), 139–168.
Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*(2), 151–218.
Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an “outlier” orthography. *Psychological Bulletin, 134*, 584615.
Share, D. L., McGee, R., & Silva, P. D. (1989). I.Q. and reading progress: A test of the capacity notion of I.Q. *Journal of the American Academy of Child and Adolescent Psychiatry, 28*, 97–100.
Share, D. L., & Stanovich, K. E. (1995). Cognitive processes in early reading development: Accommodating individual differences into a model of acquisition. *Issues in Education, 1*, 1–57.
Shavelson, R., & Towne, L. (Eds.). (2002). *Science and education*. Washington, DC: National Academy of Sciences.
Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., . . . Fletcher, J. M. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. *Biological Psychiatry, 55*, 926–933.
Shaywitz, S. E. (2004). *Overcoming dyslexia*. New York: Knopf.
Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine, 326*, 145–150.
Shaywitz, S. E., Fletcher, J. M., Holahan, J. M., Shneider, A. E., Marchione, K. E., Stuebing, K. K., . . . Shaywitz, B. A. (1999). Persistence of dyslexia: The Connecticut Longitudinal Study at adolescence. *Pediatrics, 104*, 1351–1359.
Shaywitz, S. E., Shaywitz, B. A., Fletcher, J. M., & Escobar, M. D. (1990). Prevalence of reading disability in boys and girls: Results of the Connecticut Longitudinal Study. *Journal of the American Medical Association, 264*, 998–1002.
Shaywitz, S. E., Shaywitz, B. A., Wietecha, L., Wigal, S., McBurnett, K., Williams, D., . . . Hooper, S. R.

(2017). Effect of atomoxetine treatment on reading and phonological skills in children with dyslexia or attention-deficit/hyperactivity disorder and comorbid dyslexia in a randomized, placebo-controlled trial. *Journal of Child and Adolescent Psychopharmacology, 27*, 19–28.

Shinn, M. (Ed.). (1989). *Curriculum-based measurement: Assessing special children*. New York: Guilford Press.

Sideridis, G. D., Simos, P., Mouzaki, A., & Stamovlasis, D. (2016). Efficient word-reading: Automaticity of print-related skills indexed by rapid automatized naming through cusp–catastrophe modeling. *Scientific Studies of Reading, 20*, 6–19.

Siegel, L. S. (1992). An evaluation of the discrepancy definition of dyslexia. *Journal of Learning Disabilities, 25*, 618–629.

Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled. *Child Development, 60*, 973–980.

Siegler, R. S., & Shrager, J. (1984). Strategy choice in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229–293). Hillsdale, NJ: Erlbaum.

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology, 62*, 273–296.

Silani, G., Frith, U., Demonet, J. F., Fazio, F., Perani, D., Price, C., Frith, C. D., . . . Paulesu, E. (2005). Brain abnormalities underlying altered activation in dyslexia: A voxel based morphometry study. *Brain, 128*, 453–461.

Silva, M., & Cain, K. (2015). The relations between lower and higher level comprehension skills and their role in prediction of early reading comprehension. *Journal of Educational Psychology, 107*(2), 321–331.

Silva, P. A., McGee, R., & Williams, S. (1985). Some characteristics of 9-year-old boys with general reading backwardness or specific reading retardation. *Journal of Child Psychology and Psychiatry, 26*, 407–421.

Simos, P. G., Fletcher, J. M., Bergman, E., Breier, J. I., Foorman, B. R., Castillo, E. M., . . . Papanicolaou, A. C. (2002a). Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology, 58*, 1203–1213.

Simos, P. G., Fletcher, J. M., Foorman, B. R., Francis, D. J., Castillo, E. M., Davis, R. N., . . . Papanicolaou, A. C. (2002b). Brain activation profiles during the early stages of reading acquisition. *Journal of Child Neurology, 17*, 159–163.

Simos, P. G., Fletcher, J. M., Sarkari, S., Billingsley, R. L., Denton, C., & Papanicolaou, A. C. (2007a). Altering the brain circuits for reading through intervention: A magnetic source imaging study. *Neuropsychology, 21*, 485–496.

Simos, P. G., Fletcher, J. M., Sarkari, S., Billingsley, R. L., Francis, D. J., Castillo, E. M., . . . Papanicolaou, A. C. (2005). Early development of neurophysiological processes involved in normal reading and reading disability: A magnetic source imaging study. *Neuropsychology, 19*, 787–798.

Simos, P. G., Fletcher, J. M., Sarkari, S., Billingsley-Marshall, R., Denton, C., & Papanicolaou, A. C. (2007b). Intensive instruction affects brain magnetic activity associated with reading fluency in children with persistent reading disabilities. *Journal of Learning Disabilities, 40*, 37–48.

Simos, P. G., Kanatsouli, K., Fletcher, J. M., Sarkari, S., Juranek, J., Cirino, P., Passaro, A., & Papanicolaou, A. C. (2008). Aberrant spatiotemporal activation profiles associated with math difficulties in children: A magnetic source imaging study. *Neuropsychology, 22*, 571–584.

Simos, P. G., Papanicolaou, A. C., Breier, J. I., Fletcher, J. M., Wheless, J. W., Maggio, W. W., . . . Kramer, L. (2000). Insights into brain function and neural plasticity using magnetic source imaging. *Journal of Clinical Neurophysiology, 17*, 143–162.

Simos, P. G., Pugh, K., Mencl, E., Frost, S., Fletcher, J. M., Sarkari, S., & Papanicolaou, A. C. (2009).

Temporal course of word recognition in skilled readers: A magnetoencephalography study. *Behavioral Brain Research, 197*, 45–54.

Simos, P. G., Rezaie, R., Fletcher, J. M., Juranek, J., & Papanicolaou, A. C. (2011a). Neural correlates of sentence reading in children with reading difficulties. *Neuroreport, 22*, 674–678.

Simos, P., Rezaie, R., Fletcher, J. M., Juranek, J., Passaro, A., Li, Z., Cirino, P. T., & Papanicolaou, A. C. (2011b). Functional disruption of the brain mechanism for reading: Effects of comorbidity and task difficulty among children with developmental learning problems. *Neuropsychology, 24*, 520–534.

Simos, P. G., Rezaie, R., Papanicolaou, A. C., & Fletcher, J. M. (2014). Does IQ affect the functional brain network involved in pseudoword-reading in students with reading disability: A magnetoencephalography study. *Frontiers in Human Neuroscience, 7*, 1–9.

Skiba, T., Landi, N., Wagner, R., & Grigorenko, E. L. (2011). In search for the perfect phenotype: An analysis of linkage and association studies of reading and reading-related processes. *Behavior Genetics, 41*, 6–30.

Skinner, H. (1981). Toward the integration of classification theory and methods. *Journal of Abnormal Psychology, 90*, 68–87.

Skottun, B. C., & Skoyles, J. R. (2008). Dyslexia and rapid visual processing: A commentary. *Journal of Clinical and Experimental Neuropsychology, 30*, 666–673.

Smith, T., Cobb, P., Farran, D. C., Cordray, D. S., & Munter, C. (2013). Evaluating Math Recovery: Assessing the causal impact of a diagnostic tutoring program on student achievement. *American Educational Research Journal, 50*, 397–428.

Smith, S. D., Grigorenko, E. L., Willcutt, E., Pennington, B., Olson, R. K., & DeFries, J. C. (2010). Etiologies and molecular mechanisms of communication disorders. *Journal of Developmental and Behavioral Pediatrics, 31*, 555–563.

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

Snowling, M., & Hulme, C. (1994). The development of phonological skills. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 346*(1315), 21–27.

Snowling, M., & Melby-Lervåg, M. (2016). Oral language deficits in familial dyslexia: A meta-analysis and review. *Psychological Bulletin, 142*, 498–545.

Soden, B., Christopher, M. E., Hulslander, J., Olson, R. K., Cutting, L., Keenan, J. M., . . . Petrill, S. (2015). Longitudinal stability in reading comprehension is largely heritable from grades 1 to 6. *PLOS ONE, 10*(1), e0113807.

Solan, H. A., & Richman, J. (1990). Irlen lenses: A critical appraisal. *Journal of the American Optometric Association, 61*, 789–796.

Song, M., Coggshall, J. C., & Miskel, C. (2004). Where does policy usually come from and why should we care? In P. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 445–462). Baltimore: Brookes.

Spear-Swerling, L. (2015). *The power of RTI and reading profiles: A blueprint for solving reading problems*. Baltimore: Brookes.

Speece, D. L., & Case, L. P. (2001). Classification in context: An alternative approach to identifying early reading disability. *Journal of Educational Psychology, 93*, 735–749.

Speer, N. K., Reynolds, J. R., Swallow, K. M., & Zacks, J. M. (2009). Reading stories activates neural representations of visual and motor experiences. *Psychological Science, 20*(8), 989–999.

Spelke, E. A. (2005). Sex differences in intrinsic aptitude for mathematics and science? *American Psychologist, 60*, 950–958.

Spelke, E. S., & Tsivkin, S. (2001). Initial knowledge and conceptual change: Space and number. In M.

Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 70–97). Cambridge, UK: Cambridge University Press.

Spencer, M., Quinn, J. M., & Wagner, R. K. (2014a). Specific reading comprehension disability: Major problem, myth, or misnomer? *Learning Disabilities Research and Practice, 29*, 3–9.

Spencer, M., Wagner, R. K., Schatschneider, C., Quinn, J. M., Lopez, D., & Petsher, Y. (2014b). Incorporating RTI in a hybrid model of reading disability. *Learning Disability Quarterly, 37*, 161–171.

Sperling, A. J., Lu, Z. L., Manis, F. R., & Seidenberg, M. S. (2006). Motion-perception deficits and reading impairment: It's the noise, not the motion. *Psychological Science, 17*(12), 1047–1053.

Spreen, O. (2011). Nonverbal learning disabilities: A critical review. *Child Neuropsychology, 17*(5), 418–443.

Sprenger-Charolles, L., Siegel, L. S., Jimenez, J. E., & Ziegler, J. C. (2011). Prevalence and reliability of phonological, surface, and mixed profiles in dyslexia: A review of studies conducted in languages varying in orthographic depth. *Scientific Studies of Reading, 15*, 498–521.

Spring, C., & French, L. (1990). Identifying reading-disabled children from listening and reading discrepancy scores. *Journal of Learning Disabilities, 23*, 53–58.

Stahl, S. A. (2004). What do we know about fluency?: Findings of the National Reading Panel. In P. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 187–212). Baltimore: Brookes.

Stahl, S. A., Heubach, K., & Cramond, B. (1997). *Fluency-oriented reading instruction*. Athens, GA/Washington, DC: National Reading Research Center/U.S. Department of Education, Office of Educational Research and Improvement, Educational Resources Information Center.

Stanberry, L. I., Richards, T. L., Berninger, V. W., Nandy, R. R., Aylward, E. H., Maravilla, K. R., . . . Cordes, D. (2006). Low-frequency signal changes reflect differences in functional connectivity between good readers and dyslexics during continuous phoneme mapping. *Magnetic Resonance Imaging, 24*, 217–229.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–407.

Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the garden-variety poor reader: The phonological–core variable-difference model. *Journal of Learning Disabilities, 21*, 590–604.

Stanovich, K. E. (1991). Discrepancy definitions of reading disability: Has intelligence led us astray? *Reading Research Quarterly, 26*, 7–29.

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

Stanovich, K. E., Siegel, L. S., & Gottardo, A. (1997). Converging evidence for phonological and surface subtypes of reading disability. *Journal of Educational Psychology, 89*, 114–127.

Stanovich, P. J., & Stanovich, K. E. (2003). *Using research and reason in education: How teachers can use scientifically based research to make curricular decisions*. Washington, DC: National Institute for Literacy.

Starkey, P., Spelke, E. S., & Gelman, R. (1991). Toward a comparative psychology of number. *Cognition, 39*, 171–172.

Starr, A., Libertus, M. E., & Brannon, E. M. (2013). Infants show ratio-dependent number discrimination regardless of set size. *Infancy, 18*, 927–941.

Stecker, P. M., Fuchs, L. S., & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of research. *Psychology in the Schools, 42*, 795–819.

Stein, J. (2001). The magnocellular theory of developmental dyslexia. *Dyslexia, 7*(1), 12–36.

Stein, J. (2014). Dyslexia: The role of vision and visual attention. *Current Developmental Disorders Reports, 1*, 267–280.

Stevenson, J., Graham, P., Fredman, G., & McLoughlin, V. (1987). A twin study of genetic influences on reading and spelling ability and disability. *Journal of Child Psychology and Psychiatry, 28*, 229–247.

Stockard, J., Wood, T. W., Coughlin, C., & Khoury, C. R. (in press). The effectiveness of DirectInstruction curricula: A meta-analysis of a half century of research. *Reading Research Quarterly*.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38*, 934–947.

Stothard, S. E., & Hulme, C. (1996). A comparison of reading comprehension and decoding difficulties in children. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 93–112). Mahwah, NJ: Erlbaum.

St. Pierre, E. A. (2000). The call for intelligibility in postmodern education research. *Educational Researcher, 29*, 25–28.

Strauss, A. A., & Lehtinen, L. E. (1947). *Psychopathology and education of the brain-injured child: Vol. 2. Progress in theory and clinic*. New York: Grune & Stratton.

Strong, G. K., Torgerson, C., Torgerson, D., & Hulme, C. (2011). A systematic meta-analytic review of evidence for the effectiveness of the "Fast ForWord" language intervention program. *Journal of Child Psychology and Psychiatry, 52*, 224–235.

Struiksma, C., van der Leij, A., & Stoel, R. (2009). Response to fluency-oriented intervention of Dutch poor readers. *Learning and Individual Differences, 19*, 541–548.

Stuebing, K. K., Barth, A. E., Cirino, P. T., Francis, D. J., & Fletcher, J. M. (2008). A response to recent reanalyses of the National Reading Panel Report: Effects of systematic phonics instruction are practically significant. *Journal of Educational Psychology, 100*, 123–134.

Stuebing, K. K., Barth, A. E., Molfese, P. J., Weiss, B., & Fletcher, J. M. (2009). IQ is *not* strongly related to response to reading instruction: A meta-analytic interpretation. *Exceptional Children, 76*, 31–51.

Stuebing, K. K., Barth, A. E., Trahan, L., Reddy, R., Miciak, J., & Fletcher, J. M. (2015). Are child characteristics strong predictors of response to intervention?: A meta-analysis. *Review of Educational Research, 85*, 395–429.

Stuebing, K. K., Fletcher, J. M., Branum-Martin, L., & Francis, D. J. (2012). Simulated comparisons of three methods for identifying specific learning disabilities based on cognitive discrepancies. *School Psychology Review, 4*, 3–22.

Stuebing, K. K., Fletcher, J. M., LeDoux, J. M., Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2002). Validity of IQ-discrepancy classifications of reading disabilities: A meta-analysis. *American Educational Research Journal, 39*, 469–518.

Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities, 49*(1), 77–96.

Supekar, K., Swigart, A. G., Tenison, C., Jolles, D. D., Rosenberg-Lee, M., Fuchs, L., & Menon, V. (2013). Neural predictors of individual differences in response to math tutoring in primary-grade school children. *Proceedings of the National Academy of Sciences of the USA, 110*, 8230–8235.

Swanborn, M. S., & De Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research, 69*, 261–285.

Swanson, E., Wanzek, J., Vaughn, S., Roberts, G., & Fall, A. M. (2015). Improving reading comprehension and social studies knowledge among middle school students with disabilities. *Exceptional Children, 81*, 426–442.

Swanson, H. L. (2013). Meta-analysis of research on children with learning disabilities. In H. L.

Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed.). New York: Guilford Press.

Swanson, H. L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem-solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 96*, 471–491.

Swanson, H. L., Harris, K. R., & Graham, S. (Eds.). (2013). *Handbook of learning disabilities* (2nd ed.). New York: Guilford Press.

Swanson, H. L., with Hoskyn, M., & Lee, C. (1999). *Interventions for students with learning disabilities: A meta-analysis of treatment outcome*. New York: Guilford Press.

Swanson, H. L., Howard, C. B., & Saez, L. (2006). Do different components of working memory underlie different subgroups of reading disabilities? *Journal of Learning Disabilities, 39*, 252–269.

Tager-Flusberg, H., & Cooper, J. (1999). Present and future possibilities for defining a phenotype for specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*, 1275–1278.

Talcott, J. B., Witton, C., McLean, M. F., Hansen, P. C., Rees, A., Green, G. G., & Stein, J. F. (2000). Dynamic sensory sensitivity and children's word decoding skills. *Proceedings of the National Academy of Sciences of the USA, 97*, 2952–2957.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language, 9*, 182–198.

Tallal, P. (2004). Improving language and literacy is a matter of time. *Nature Reviews Neuroscience, 5*, 721–728.

Tamm, L., Denton, C. A., Epstein, J. N., Schatschneider, C., Taylor, H., Arnold, L. E., . . . Maltinsky, J. (2017). Comparing treatments for children with ADHD and word-reading difficulties: A randomized clinical trial. *Journal of Consulting and Clinical Psychology, 85*, 434–451.

Tan, L. H., Spinks, J. A., Eden, G. F., Perfetti, C. A., & Siok, W. T. (2005). Reading depends on writing, in Chinese. *Proceedings of the National Academy of Sciences of the USA, 102*, 8781–8785.

Tanaka, H., Black, J., Hulme, C., Leanne, S., Kesler, S., Whitfield, G., . . . Hoeft, F. (2011). The brain basis of the phonological deficit in dyslexia is independent of IQ. *Psychological Science, 22*, 1442–1451.

Tannock, R. (2013). Rethinking ADHD and LD in DSM-5: Proposed changes in diagnostic criteria. *Journal of Learning Disabilities, 46*, 5–25.

Tannock, R., Martinussen, R., & Frijters, J. (2000). Naming speed performance and stimulant effects indicate effortful, semantic processing deficits in attention-deficit/hyperactivity disorder. *Journal of Abnormal Child Psychology, 28*, 237–252.

Taylor, J. S. H., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword-reading?: A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin, 139*, 766–791.

Taylor, J., Roehrig, A. D., Soden Hensler, B., Connor, C. M., & Schatschneider, C. (2010). Teacher quality moderates the genetic effects on early reading. *Science, 328*, 512–514.

Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., & Gabrieli, J. D. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. *Proceedings of the National Academy of Sciences of the USA, 100*(5), 2860–2865.

Thaler, V., Ebner, E. M., Wimmer, H., & Landerl, K. (2004). Training reading fluency in dysfluent readers with high reading accuracy: Word specific effects but low transfer to untrained words. *Annals of Dyslexia, 54*, 89–113.

Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading: A meta-analysis. *Remedial and Special Education, 4*, 252–261.

Thomas, L. (1983). *The youngest science: Notes of a medicine watcher*. New York: Penguin Books.

Tighe, E. L., & Schatschneider, C. (2014). A dominance analysis approach to determining predictor importance in third, seventh, and tenth grade reading comprehension skills. *Reading and Writing, 27*, 101–127.

Tighe, E. L., & Schatschneider, C. (2016). Examining the relationships of component reading skills to reading comprehension in struggling adult readers: A meta-analysis. *Journal of Learning Disabilities, 49*, 395–409.

Tiu, R. D., Jr., Wadsworth, S. J., Olson, R. K., & DeFries, J. C. (2004). Causal models of reading disability: A twin study. *Twin Research, 7*, 275–283.

Tolar, T. D., Barth, A. E., Fletcher, J. M., Francis, D. J., & Vaughn, S. (2014). Predicting reading outcomes with progress monitoring slopes among middle grade students. *Learning and Individual Differences, 30*, 46–57.

Tolar, T. D., Fuchs, L., Fletcher, J. M., Fuchs, D., & Hamlett, C. (2016). Cognitive profiles of mathematical problem-solving learning disability for different definitions of disability. *Journal of Learning Disabilities, 49*, 240–256.

Tolar, T. D., Lederberg, A. R., & Fletcher, J. M. (2009). A structural model of algebra achievement: Calculational fluency and spatial visualization as mediators of the effect of working memory on algebra achievement. *Educational Psychology, 29*, 239–266.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 40*(6), 1245–1260.

Tomblin, J. B., & Zhang, X. (1999). Language patterns and etiology in children with specific language impairment. In H. Tager-Flusberg (Ed.), *Neurodevelopmental disorders* (pp. 361–382). Cambridge, MA: MIT Press.

Torgerson, C., Porthouse, J., & Brooks, G. (2005). A systematic review of controlled trials evaluating interventions in adult literacy and numeracy. *Journal of Research in Reading, 28*(2), 87–107.

Torgesen, J. K. (2000). Individual responses in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research and Practice, 15*, 55–64.

Torgesen, J. K. (2002). Empirical and theoretical support for direct diagnosis of learning disabilities by assessment of intrinsic processing weaknesses. In R. Bradley, L. Danielson, & D. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 565–650). Mahwah, NJ: Erlbaum.

Torgesen, J. K. (2004). Lessons learned from research on interventions for students who have difficulty learning to read. In P. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 355–382). Baltimore: Brookes.

Torgesen, J. K. (2009). The Response to Intervention instructional model: Some outcomes from a large-scale implementation in Florida schools. *Child Development Perspectives, 3*, 38–40.

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., Wagner, R. K., & Rashotte, C. (2011). *Test of Word Reading Efficiency—Second Edition (TOWRE-2)*. Austin, TX: PRO-ED.

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

Tosto, M. G., Asbury, K., Mazzocco, M., Petrill, S., & Kovas, Y. (2016). From classroom environment to mathematics achievement: The mediating role of self-perceived ability and subject interest. *Learning and Individual Differences, 50*, 260–269.

Tosto, M. G., Petrill, S. A., Halberda, J., Trzaskowski, M., Tikhomirova, T. N., Bogdanova, O. Y., . . . Plomin, R. (2014). Why do we differ in number sense?: Evidence from a genetically sensitive investigation. *Intelligence, 43*, 35–46.

Treiman, R., & Kessler, B. (2005). Writing systems and spelling development. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 120–134). Oxford, UK: Blackwell.

Tunmer, W. E., Chapman, J. W., & Prochnow, J. E. (2003). Preventing negative Matthew effects in at-risk readers: A retrospective study. In B. R. Foorman (Ed.), *Preventing and remediating reading difficulties* (pp. 121–164). Baltimore: York Press.

Unsworth, N., & Spillers, G. J. (2010). Variation in working memory capacity and episodic recall: The contributions of strategic encoding and contextual retrieval. *Psychonomic Bulletin and Review, 17*(2), 200–205.

U.S. Department of Education. (1999). 34 CFR Parts 300 and 303: Assistance to the states for the education of children with disabilities and the early intervention program for infants and toddlers with disabilities. Final regulations. *Federal Register, 64*, 12406–12672.

U.S. Department of Education. (2004). *Individuals with Disabilities Education Improvement Act*, 20 U.S. C. § 1400. Washington, DC: Author.

U.S. Department of Education. (2006). 34 CFR Parts 300 and 301: Assistance to states for the education of children with disabilities and preschool grants for children with disabilities. Final rules. *Federal Register, 71*, 46540–46845.

U.S. Department of Education. (2008). Report of the Task Group on Learning Processes. Retrieved from *www.ed.gov/about/bdscomm/list/mathpanel/report/learning*.

U.S. Office of Education. (1968). *First annual report of the National Advisory Committee on Handicapped Children*. Washington, DC: U.S. Department of Health, Education and Welfare.

U.S. Office of Education. (1977). Assistance to states for education for handicapped children: Procedures for evaluating specific learning disabilities. *Federal Register, 42*, G1082–G1085.

Van Beek, L., Ghesquière, P., Lagae, L., & De Smedt, B. (2014). Left fronto–parietal white matter correlates with individual differences in children's ability to solve additions and multiplications: A tractography study. *NeuroImage, 90*, 117–127.

van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructional processes in accounts of reading comprehension. *Discourse Processes, 39*, 299–316.

van der Mark, S., Klaver, P., Bucher, K., Maurer, U., Schulz, E., Brem, S., . . . Brandeis, D. (2011). The left occipitotemporal system in reading: Disruption of focal fMRI connectivity to left inferior frontal and inferior parietal language areas in children with dyslexia. *NeuroImage, 54*, 2426–2436.

van der Wissell, A., & Zegers, F. E. (1985). Reading retardation revisited. *British Journal of Developmental Psychology, 3*, 3–9.

Vanbinst, K., Ansari, D., Ghesquière, P., & De Smedt, B. (2016). Symbolic numerical magnitude processing is as important to arithmetic as phonological awareness is to reading. *PLOS ONE, 11*(3), e0151045.

VanDerHeyden, A. M., Witt, J. C., & Gilbertson, D. (2007). A multi-year evaluation of the effects of a Response to Intervention (RTI) model on identification of children for special education. *Journal of School Psychology, 45*, 225–256.

Vandermosten, M., Boets, B., Wouters, J., & Ghesquiere, P. (2012). A qualitative and quantitative review of diffusion tensor imaging studies in reading and dyslexia. *Neuroscience and Biobehavioral Reviews, 36*, 1532– 1552.

Vandermosten, M., Hoeft, F., & Norton, E. S. (2016). Integrating MRI brain imaging studies of pre-reading children with current theories of developmental dyslexia: A review and quantitative meta-analysis. *Current Opinion in Behavioral Sciences, 10*, 155–161.

Vandermosten, M., Vanderauwera, J., Theys, C., De Vos, A., Vanvooren, S., Sunaert, S., . . . Ghesquière, P. (2015). A DTI tractography study in pre-readers at risk for dyslexia. *Developmental Cognitive Neuroscience, 14*, 8–15.

Vaughn, S. (1989). Gifted learning disabilities: Is it such a bright idea? *Learning Disabilities Focus, 4*, 123–128.

Vaughn, S., Cirino, P. T., Wanzek, J., Wexler, J., Fletcher, J. M., Denton, C. A., . . . Francis, D. J. (2010a). Response to intervention for middle school students with reading difficulties: Effects of a primary and secondary intervention. *School Psychology Review, 39*, 3–21.

Vaughn, S., & Dammann, J. E. (2001). Science and sanity in special education. *Behavioral Disorders, 27*, 21–29.

Vaughn, S., & Fletcher, J. M. (2012). Response to intervention with secondary school students with reading difficulties. *Journal of Learning Disabilities, 45*, 244–256.

Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research and Practice, 18*, 137–146.

Vaughn, S., Klingner, J. K., & Bryant, D. P. (2001). Collaborative strategic reading as a means to enhance peer-mediated instruction for reading comprehension and content-area learning. *Remedial and Special Education, 22*, 66–74.

Vaughn, S., Klingner, J. K., Swanson, E. A., Boardman, A. G., Roberts, G., Mohammed, S. S., & Stillman, S. J. (2011a). Efficacy of collaborative strategic reading with middle school students. *American Educational Research Journal, 48*, 938–964.

Vaughn, S., Linan-Thompson, S., Kouzekanani, K., Bryant, D. P., Dickson, S., & Blozis, S. A. (2003). Reading instruction grouping for students with reading difficulties. *Remedial and Special Education, 24*, 301–315.

Vaughn, S. R., Moody, S. W., & Schumm, J. S. (1998). Broken promises: Reading instruction in the resource room. *Exceptional Children, 64*, 211–225.

Vaughn, S., Roberts, G., Swanson, E. A., Wanzek, J., Fall, A. M., & Stillman-Spisak, S. J. (2015). Improving middle-school students' knowledge and comprehension in social studies: A replication. *Educational Psychology Review, 27*(1), 31–50.

Vaughn, S., Wanzek, J., & Fletcher, J. M. (2007). Multiple tiers of intervention: A framework for prevention and identification of students with reading/learning disabilities. In B. M. Taylor & J. E. Ysseldyke (Eds.), *Effective instruction for struggling readers, K–6* (pp. 173–195). New York: Teachers College Press.

Vaughn, S., Wanzek, J., Wexler, J., Barth, A., Cirino, P. T., Fletcher, J. M., . . . Francis, D. J. (2010b). The relative effects of group size on reading progress of older students with reading difficulties. *Reading and Writing, 23*, 931–956.

Vaughn, S. R., Wanzek, J., Woodruff, A. L., & Linan-Thompson, S. (2006). A three-tier model for preventing reading difficulties and early identification of students with reading disabilities. In D. H. Haager, S. Vaughn, & J. K. Klingner (Eds.), *Validated reading practices for three tiers of intervention* (pp. 11–28). Baltimore: Brookes.

Vaughn, S., Wexler, J., Leroux, A., Roberts, G., Denton, C., Barth, A., & Fletcher, J. (2012). Effects of intensive reading intervention for eighth-grade students with persistently inadequate response to intervention. *Journal of Learning Disabilities, 45*, 515–525.

Vaughn, S., Wexler, J., Roberts, G., Barth, A. A., Cirino, P. T., Romain, M. A., . . . Denton, C. A. (2011b). Effects of individualized and standardized interventions on middle school students with reading disabilities. *Exceptional Children, 77*, 391–407.

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

Vellutino, F. R., Fletcher, J. M., Scanlon, D. M., & Snowling, M. J. (2004). Specific reading disability

(dyslexia): What have we learned in the past four decades? *Journal of Child Psychiatry and Psychology, 45*, 2–40.
Vellutino, F. R., Scanlon, D. M., & Jaccard, J. (2003). Toward distinguishing between cognitive and experiential deficits as primary sources of difficulty in learning to read: A two year follow-up of difficult to remediate and readily remediated poor readers. In B. R. Foorman (Ed.), *Preventing and Remediating reading difficulties: Bringing science to scale* (pp. 73–120). Baltimore: York Press.
Vellutino, F. R., Scanlon, D. M., & Lyon, G. R. (2000). Differentiating between difficult to remediate and readily remediated poor readers: More evidence against the IQ Achievement discrepancy definition of reading disability. *Journal of Learning Disabilities, 33*, 223–238.
Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology, 88*, 601–618.
Vellutino, F. R., Scanlon, D. M., Small, S., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *Journal of Learning Disabilities, 39*, 157–169.
Vellutino, F. R., Scanlon, D. M., & Tanzman, M. S. (1994). Components of reading ability: Issues and problems in operationalizing word identification, phonological coding, and orthographic coding. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 279–329). Baltimore: Brookes.
Verhoeven, L., van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading, 15*, 8–25.
Vidyasagar, T. R., & Pammer, K. (2010). Dyslexia: A deficit in visuo-spatial attention, not in phonological processing. *Trends in Cognitive Sciences, 14*, 57–63.
Vogel, A. C., Miezin, F. M., Petersen, S. E., & Schlaggar, B. L. (2012). The putative visual word form area is functionally connected to the dorsal attention network. *Cerebral Cortex, 22*, 537–549.
Vogel, A. C., Petersen, S. E., & Schlaggar, B. L. (2014). The VWFA: It's not just for words anymore. *Frontiers in Human Neuroscience, 8*(88), 1–10
Vukovic, R. K. (2012). Mathematics difficulty with and without reading difficulty: Findings and implications from a four-year longitudinal study. *Exceptional Children, 78*, 280–300.
Vukovic, R. K., Fuchs, L. S., Geary, D. C., Jordan, N. C., Gersten, R., & Siegler, R. S. (2014). Sources of individual differences in children's understanding of fractions. *Child Development, 85*, 1461–1476.
Vukovic, R. K., & Siegel, L. S. (2006). The double-deficit hypothesis: A comprehensive analysis of the evidence. *Journal of Learning Disabilities, 39*, 25–47.
Waber, D. (2010). *Rethinking learning disabilities: Understanding children who struggle in school*. New York: Guilford Press.
Waber, D. P., Forbes, P. W., Wolff, P. H., & Weiler, M. D. (2004). Neurodevelopmental characteristics of children with learning impairments classified according to the double-deficit hypothesis. *Journal of Learning Disabilities, 37*, 451–461.
Waber, D. P., Weiler, M. D., Wolff, P. H., Bellinger, D., Marcus, D. J., Ariel, R., . . . Wypij, D. (2001). Processing of rapid auditory stimuli in school-age children referred for evaluation of learning disorders. *Child Development, 72*, 37–49.
Waber, D. P., Wolff, P. H., Forbes, P. W., & Weiler, M. D. (2000). Rapid automatized naming in children referred for evaluation of heterogeneous learning problems: How specific are naming speed deficits to reading disability? *Child Neuropsychology, 6*, 251–261.
Wadsworth, S. J., DeFries, J. C., Willcutt, E. G., Pennington, B. F., & Olson, R. K. (2015). The Colorado Longitudinal Twin Study of Reading Difficulties and ADHD: Etiologies of comorbidity and

stability. *Twin Research and Human Genetics, 18*, 755–761.
Wadsworth, S. J., Olson, R. K., Pennington, B. F., & DeFries, J. C. (2000). Differential genetic etiology of reading disability as a function of IQ. *Journal of Learning Disabilities, 33*, 192–199.
Wagner, R. K., Puranik, C. S., Foorman, B., Foster, E., Tschinkel, E., & Kantor, P. T. (2011). Modeling the development of written language. *Reading and Writing, 24*, 203–220.
Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). The development of reading-related phonological processing abilities: New evidence of bi-directional causality from a latent variable longitudinal study. *Developmental Psychology, 30*, 73–87.
Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Hecht, S. A. (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.
Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2010). *Test of Silent Reading Efficiency and Comprehension*. Austin TX: PRO-ED.
Wang, Y., Mauer, M. V., Raney, T., Peysakhovich, B., Becker, B. L., Sliva, D. D., . . . Gaab, N. (2016). Development of tract-specific white matter pathways during early reading development in at-risk children and typical controls. *Cerebral Cortex, 27*, 2469–2485.
Wanzek, J., & Vaughn, S. (2007). Research-based implications from extensive early reading interventions. *School Psychology Review, 36*, 541–561.
Wanzek, J., Vaughn, S., Scammacca, N., Gatlin, B., Walker, M. A., & Capin, P. (2016). Meta-analyses of the effects of Tier 2 type reading interventions in grades K–3. *Educational Psychology Review, 28*(3), 551–576.
Wanzek, J., Vaughn, S., Scammacca, N., Metz, K., Murray, C., Roberts, G., & Danielson, L. (2013). Extensive reading interventions for students with reading difficulties after grade 3. *Review of Educational Research, 83*, 163–195.
Watkins, M. W., & Canivez, G. L. (2004). Temporal stability of WISC-III subtest composite: Strengths and weaknesses. *Psychological Assessment, 16*, 133–138.
Wechsler, D. (2001). *Wechsler Individual Achievement Test* (2nd ed.). San Antonio, TX: Psychological Corporation.
Weiser, B., & Mathes, P. (2011). Using encoding instruction to improve the reading and spelling performances of elementary students at risk for literacy difficulties: A best-evidence synthesis. *Review of Educational Research, 81*, 170–200.
Welch-Ross, M. K., & Fasig, L. G (Eds.). (2007). *Handbook on communicating and disseminating behavioral science*. Thousand Oaks, CA: SAGE.
Welcome, S. E., Chiarello, C., Thompson, P. M., & Sowell, E. R. (2011). Reading skill is related to individual differences in brain structure in college students. *Human Brain Mapping, 32*, 1194–1205.
Wexler, J., Reed, D. K., Pyle, N., Mitchell, M., & Barton, E. E. (2015). A synthesis of peer mediated interventions for secondary struggling learners. *Journal of Learning Disabilities, 48*, 451–470.
What Works Clearinghouse. (2012, May). Peer-assisted learning/literacy strategies. Retrieved from *https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_pals_050112.pdf.*
What Works Clearinghouse. (2013). Reading intervention report: Reading Recovery. Retrieved March 1, 2018, from *https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_readrecovery_071613.pdf.*
White-Schwoch, T., Carr, K. W., Thompson, E. C., Anderson, S., Nicol, T., Bradlow, A. R., . . . Kraus, N. (2015). Auditory processing in noise: A preschool biomarker for literacy. *PLOS Biology, 13*(7), e1002196.
Wiig, E. H., Neilsen, N. P., Minthon, L., McPeek, D., Said, K., & Warkentin, S. (2002). Parietal lobe activation in rapid, automatized naming by adults. *Perceptual and Motor Skills, 94*, 1230–1244.

Wilder, A. A., & Williams, J. P. (2001). Students with severe learning disabilities can learn higher-order comprehension skills. *Journal of Educational Psychology, 93*, 268–278.
Wilkinson, G. S, & Robertson, G. R. (2006). *Wide Range Achievement Test–4*. Lutz, FL: Psychological Assessment Resources.
Willburger, E., & Landerl, K. (2010). Anchoring the deficit of the anchor deficit: Dyslexia or attention? *Dyslexia, 16*, 175–182.
Willcutt, E. G. (2014). Behavioral genetic approaches to understand the etiology of comorbidity. In S. H. Rhee & A. Ronald (Eds.), *Behavior genetics of psychopathology* (pp. 231–252). New York: Springer.
Willcutt, E. G., Betjemann, R. S., McGrath, L. M., Chhabildas, N. A., Olson, R. K., DeFries J. C., & Pennington, B. F. (2010a). Etiology and neuropsychology of comorbidity between RD and ADHD: The case for multiple-deficit models. *Cortex, 46*, 1345–1361.
Willcutt, E. G., Betjemann, R., Pennington, B. F., Olson, R. K., DeFries, J. C., & Wadsworth, S. J. (2007). Longitudinal study of reading disability and attention-deficit/hyperactivity disorder: Implications for education. *Mind, Brain, and Education, 1*, 181–192.
Willcutt, E. G., Pennington, B. F., Duncan, L., Smith, S. D., Keenan, J. M., Wadsworth, S. J., & DeFries, J. C. (2010b). Understanding the complex etiology of developmental disorders: Behavioral and molecular genetic approaches. *Journal of Developmental and Behavioral Pediatrics, 31*, 533–544.
Willcutt, E. G., Petrill, S. A., Wu, S., Boada, R., DeFries, J. C., Olson, R. K., & Pennington, B. F. (2013). Comorbidity between reading disability and math disability: Concurrent psychopathology, functional impairment, and neuropsychological functioning. *Journal of Learning Disabilities, 46*, 500–516.
Williams, J. P. (2005). Instruction in reading comprehension for primary-grade students: A focus on text structure. *Journal of Special Education, 39*(1), 6–18.
Williams, J. P., Hall, K. M., Lauer, K. D., Stafford, B., DeSisto, L. A., & deCani, J. S. (2005). Expository text comprehension in the primary grade classroom. *Journal of Educational Psychology, 97*, 538–550.
Williams, J. P., Lauer, K. D., Hall, K. M., Lord, K. M., Gugga, S. S., Bak, S. J., . . . deCani, J. S. (2002). Teaching elementary school students to identify story themes. *Journal of Educational Psychology, 94*, 235–248.
Williams, J. P., Pollini, S., Nubla-Kung, A. M., Snyder, A. E., Garcia, A., Ordynans, J. G., & Atkins, J. G. (2014). An intervention to improve comprehension of cause/effect through expository text structure instruction. *Journal of Educational Psychology, 106*, 1–17.
Williams, K. J., Walker, M. A., Vaughn, S., & Wanzek, J. (2017). A synthesis of reading and spelling interventions and their effects on spelling outcomes for students with learning disabilities. *Journal of Learning Disabilities, 50*(3), 286–297.
Williams, V. J., Juranek, J., Cirino, P. T., & Fletcher, J. M. (2018). Cortical thickness and local gyrification in children with developmental dyslexia. *Cerebral Cortex, 28*, 963–973.
Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Applied Psycholinguistics, 14*, 1–33.
Wimmer, H. (2006). Don't neglect reading fluency! *Developmental Science, 9*(5), 447–448.
Wimmer, H., & Mayringer, H. (2002). Dysfluent reading in the absence of spelling difficulties: A specific disability in regular orthographies. *Journal of Educational Psychology, 94*, 272–277.
Wimmer, H., Mayringer, H., & Landerl, K. (2000). The double-deficit hypothesis and difficulties in learning to read a regular orthography. *Journal of Educational Psychology, 92*, 668–680.
Wimmer, H., Mayringer, H., & Raberger, T. (1999). Reading and dual-task balancing: Evidence against the automatization deficit explanation of developmental dyslexia. *Journal of Learning Disabilities,*

32(5), 473–478.

Wimmer, H., & Schurz, M. (2010). Dyslexia in regular orthographies: Manifestation and causation. *Dyslexia, 16*(4), 283–299.

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.

Witt, J. C., Daly, E. M., & Noell, G. (2000). *Functional assessments: A step by step guide to solving academic and behavior problems*. Longmont, CO: Sopris West.

Wolf, M. A. (2007). *Proust and the squid: The story and science of the reading brain*. New York: Harper Collins.

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91*, 415–438.

Wolf, M., Miller, L., & Donnelly, K. (2002). Retrieval, Automaticity, Vocabulary Elaboration, Orthography (RAVE-O): A comprehensive, fluency-based reading intervention program. *Journal of Learning Disabilities, 33*, 375–386.

Wolf, M., & Obregon, M. (1992). Early naming deficits, developmental dyslexia, and a specific deficit hypothesis. *Brain and Language, 42*, 19–47.

Wolf, M., O'Brien, B., Adams, K. D., Joffe, T., Jeffrey, J., Lovett, M., & Morris, R. (2003). Working for time: Reflections on naming speed, reading fluency, and intervention. In B. R. Foorman (Ed.), *Preventing and remediating reading difficulties* (pp. 355–380). Baltimore: York Press.

Wolff, P. (1993). Impaired temporal resolution in developmental dyslexia: Temporal information processing in the nervous system. In P. Tallal, A. Galaburda, R. Llinas, & C. von Euler (Eds.), *Annals of the New York Academy of Sciences, 682*, 87–103.

Wong, B. Y. L. (1991). The relevance of metacognition to learning disabilities. In B. Y. L. Wong (Ed.), *Learning about learning disabilities* (pp. 231–258). San Diego, CA: Academic Press.

Wood, F. B., & Felton, R. H. (1994). Separate linguistic and attentional factors in the development of reading. *Topics in Language Disorders, 14*, 42–57.

Wood, F. B., Felton, R. H., Flowers, L., & Naylor, C. (1991). Neurobehavioral definition of dyslexia. In D. D. Duane & D. B. Gray (Eds.), *The reading brain: The biological basis of dyslexia* (pp. 1–26). Parkton, MD: York Press.

Wood, F. B., & Grigorenko, E. L. (2001). Emerging issues in the genetics of dyslexia: A methodological preview. *Journal of Learning Disabilities, 34*, 503–512.

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

Woodcock, R. W., Munoz-Sandoval, A. F., McGrew, K. M., & Mather, N. (2007). *Batería III Woodcock–Muñoz*. New York: Houghton, Mifflin, Harcourt.

World Health Organization. (2013). *The ICD-10 classification of mental and behavioral disorders: Clinical descriptions and diagnostic guidelines*. Geneva, Switzerland: Author.

Wynn, K. (1998). Psychological foundations of number: Numerical competence in human infants. *Trends in Cognitive Sciences, 2*, 296–303.

Xia, Z., Hoeft, F., Zhang, L., & Shu, H. (2016). Neuroanatomical anomalies of dyslexia: Disambiguating the effects of disorder, performance, and maturation. *Neuropsychologia, 81*, 68–78.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*(1), B1–B11.

Yamada, Y., Stevens, C., Dow, M., Harn, B. A., Chard, D. J., & Neville, H. J. (2011). Emergence of the neural network for reading in five-year-old beginning readers of different levels of pre-literacy abilities: An fMRI study. *NeuroImage, 57*, 704–713.

Yeo, D. J., Wilkey, E. D., & Price, G. R. (2017). The search for the number form area: A functional

neuroimaging meta-analysis. *Neuroscience and Biobehavioral Reviews, 78*, 145–160.
Yoshimasu, K., Barbaresi, W. J., Colligan, R. C., Killian, J. M., Voigt, R. G., Weaver, A. L., & Katusic, S. K. (2011). Written-language disorder among children with and without ADHD in a population-based birth cohort. *Pediatrics, 128*(3), e605–e612.
Yuill, N., & Oakhill, J. (1991). *Children's problems in text comprehension: An experimental investigation*. Cambridge, UK: Cambridge University Press.
Zeleke, S. (2004). Self-concepts of students with learning disabilities and their normally achieving peers: A review. *European Journal of Special Needs Education, 19*, 145–170.
Ziegler, J. C., Castel, C., Pech-Georgel, C., George, F., Alario, F. X., & Perry, C. (2008). Developmental dyslexia and the dual route model of reading: Simulating individual differences and subtypes. *Cognition, 107*, 151–178.
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.
Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science, 12*, 732–745.
Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Körne, G. (2003). Developmental dyslexia in different languages: Language-specific or universal? *Journal of Experimental Child Psychology, 86*, 169–193.
Zigmond, N., & Baker, J. M. (1996, Winter). Full inclusion for students with learning disabilities: Too much of a good thing. *Theory into Practice, 35*(1), 26–34.
Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162–185.

Index

The pagination of this electronic edition does not match the edition from which it was created. To locate a specific passage, please use the search feature of your e-book reader or select a page number link below.

Note. *f* or *t* following a page number indicates a figure or a table.

B

C

D

E

F

G

H

I

K

L

M

N

O

P

S

T

U

V

W

About Guilford Press

www.guilford.com

Founded in 1973, Guilford Publications, Inc., has built an international reputation as a publisher of books, periodicals, software, and DVDs in mental health, education, geography, and research methods. We pride ourselves on teaming up with authors who are recognized experts, and who translate their knowledge into vital, needed resources for practitioners, academics, and general readers. Our dedicated editorial professionals work closely on each title to produce high-quality content that readers can rely on. The firm is owned by its founding partners, President Bob Matloff and Editor-in-Chief Seymour Weingarten, and many staff members have been at Guilford for more than 20 years.