



TEACHING STAMINA & SILENT READING IN THE DIGITAL-GLOBAL AGE

Edited by Elfrieda H. Hiebert



Teaching Stamina and Silent Reading in the Digital-Global Age

EDITED BY

Elfrieda H. Hiebert

TextProject

University of California, Santa Cruz

Reading Essentials Original Series

June 2015

TextProject, Inc.

SANTA CRUZ, CALIFORNIA



textproject.org

ISBN: 978-1-937889-04-3

© 2015 Elfrieda H. Hiebert. Some rights reserved.



This work is licensed under the Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 United States License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/us/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

“TextProject” and the TextProject logo are trademarks of TextProject, Inc.

Cover photo (top) © istockphoto.com/CEFutcher. All rights reserved. Used under license.

Cover photo (bottom) © istockphoto.com/CEFutcher. All rights reserved. Used under license.

Contents

FORWARD iii

Striking the Right Balance: Why Silent and Extended Reading of
Challenging Materials Matters

Timothy Rasinski

Preface vii

Teaching Stamina and Silent Reading in the Digital-Global Age

Elfrieda H. Hiebert

I. UNDERSTANDING THE PROBLEM & THE CONSTRUCT

CHAPTER 1 16

The Forgotten Reading Proficiency: Stamina In Silent Reading

Elfrieda H. Hiebert

CHAPTER 2 32

Eye Movements Make Reading Possible

S. Jay Samuels, Elfrieda H. Hiebert, & Timothy Rasinski

CHAPTER 3 58

Are Students Really Reading in Independent Reading Contexts?

An Examination of Comprehension-Based Silent Reading Rate

Elfrieda H. Hiebert, Kathleen M. Wilson & Guy Trainin

II. INSTRUCTION AND PRACTICES THAT FOSTER STAMINA AND SILENT READING

CHAPTER 4 79

Stretching Elementary Students in Complex Text:

Why? How? When?

Heidi Anne E. Mesmer

CHAPTER 5 **100**

The Relationship Between a Silent Reading Fluency Instructional Protocol on Students’ Reading Comprehension and Achievement in an Urban School Setting

Timothy V. Rasinski, S. Jay Samuels, Elfrieda H. Hiebert, Yaacov Petscher, & Karen Feller

CHAPTER 6 **121**

Exploring the Added Value of a Guided Silent Reading Intervention: Effects on Struggling Third-Grade Readers’ Achievement

D. Ray Reutzel, Yaacov Petscher, & Alexandra N. Spichtig

III: LOOKING AHEAD

CHAPTER 7 **147**

Comprehension-Based Silent Reading Rates: What Do We Know? What Do We Need To Know?

Elfrieda H. Hiebert, S. Jay Samuels, & Timothy V. Rasinski

CHAPTER 8 **169**

Revisiting Silent Reading in 2020 and Beyond

Elfrieda H. Hiebert, & D. Ray Reutzel

FORWARD

Striking the Right Balance: Why Silent and Extended Reading of Challenging Materials Matters

Timothy Rasinski
Kent State University

Over the past dozen years or so, reading assessment and reading instruction itself increasingly have come to be defined primarily by oral reading, often for speed, and for very short periods of time. This evolution has been due to a number of factors. First, the curriculum-based measurement (CBM) approach to reading assessment reduced reading assessment to measures of reading rate over one-minute periods (Deno, 1985). Second, although the National Reading Panel (National Institute of Child Health and Human Development, 2000) identified reading fluency as a critical variable for proficient reading, the panel restricted fluency to oral reading. These developments, as well as others, have affected how reading is being taught. In many primary and intermediate classrooms around the country, oral reading is the predominant form of reading. Time is allocated to daily fluency instruction where students read a short passage repeatedly for the primary purpose of reading it faster. This practice is accompanied by regular assessments of students' reading (as often as weekly in some classrooms) on the number of words students can read correctly in a minute on an instructional-level passage.

These developments in reading instruction are based on a solid foundation of research and indeed I feel there is a legitimate place for them in the classroom. Reading rate as determined by CBMs and other similar one-minute readings of short passages is a good measure of word-recognition automaticity, a critical factor in comprehension. Oral reading experiences, especially authentic oral experiences, such as the recitation of poetry or the performance of a script, have been shown to improve reading fluency and overall reading proficiency. However, these are not the only instructional factors that must be considered for effective reading and reading instruction.

Most reading done by adults is silent reading. As such, it is not

unreasonable to expect students to receive instruction and support in silent reading in their classrooms. Further, a fair amount of adult reading consists of lengthy texts. Again, therefore, it is not unreasonable to provide students with support and opportunities to read such texts on their own, primarily silently. Additionally, as adults we are occasionally called on to closely read material that we might consider difficult or challenging—for example, technical texts related to our profession or courses we may be taking, or even texts we read for our own pleasure and entertainment purposes that may be more challenging in nature. Certainly, giving students similar opportunities to read challenging material—with appropriate support—needs to be part of our reading instruction. Therefore, we may safely conclude that issues of silent reading fluency, stamina, and close reading of complex texts are foundational for proficiency in reading and success in various academic and technical fields. These issues are challenging for literacy scholars and educators alike. Each element is critical in its own right. Additionally, these elements interact with one another and other critical variables in the reading process. Despite their importance, however, silent reading, stamina, and complex texts are issues that, until recently, have not received sufficient scholarly attention.

That is what this book is about. Dr. Elfrieda Hiebert is one of the few literacy scholars who has extensively studied issues of silent reading, stamina, and text complexity, and in this volume, she has assembled a collection of original and previously published papers that explore these vital issues in depth. This volume offers readers the opportunity to explore the conceptual nature of these issues and discover how exactly we may begin to go about providing students with the relevant instruction that will help them achieve success in these areas. The first four chapters explore just what stamina and silent reading mean. The notion of text complexity is embedded within these chapters, as readers must engage in silent reading with stamina in order to negotiate such texts successfully and efficiently. Here we are confronted with the reality of what happens in school-based reading instruction. To develop proficiency in reading, students need to practice reading.

As Hiebert stresses in her opening chapter, many students simply do not spend enough time reading in school. Increasing the amount that students read silently in schools is one of the solutions explored in the second half of the volume. Other ways that students can be supported in developing silent reading stamina are explored in the final four chapters. These applications range from a computer-based instructional protocol

that requires minimal teacher input to an approach that relies on the teacher to provide scaffolding and support for silent reading. These chapters may provide foundational principles that educators and scholars can use in order to develop their own approaches to instruction that develop students' silent reading stamina.

Success in real world reading is not measured by how fast a person can orally read a short text. Rather, reading success is more likely to be an outcome of how well a person can engage in meaningful, close, silent readings of lengthy, complex, and challenging material for extended periods of time. This volume is a significant step in moving the literacy field, scholars, curriculum developers, and practitioners toward a deeper consideration and understanding of these critical issues.

References

Deno, S.L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52(3), 219–232.

National Institute of Child Health and Human Development. (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. 00-4769). Washington, DC: U.S. Government Printing Office.

Preface

Teaching Stamina and Silent Reading in the Digital-Global Age

Elfrieda H. Hiebert

TextProject & University of California, Santa Cruz

The 21st century demands that individuals have a high level of literacy to successfully participate in the tasks of colleges, communities, and jobs. Nonetheless, many students in the United States are not attaining the necessary levels of literacy, according to the National Assessment of Educational Progress (National Center for Education Statistics, 2013). The assessment shows that approximately two-thirds of a grade cohort fail to attain the proficient level. Such poor performances are often traced to a lack of word recognition skills, and solutions have been designated to ameliorate this perceived gap (see California Board of Education, 2014). But evidence is strong that all but a small percentage of American students—approximately 2% of a grade cohort—can recognize the majority of words in a grade-level text by the end of the primary grades (Bielinski, Daniel, & Hiebert, 2015). Most students can read, but they don't have rigorous independent reading habits. What many students lack is stamina—the ability to persevere in reading texts on their own.

In the contexts of most real-life reading and as reflected in assessments, individuals read texts on their own. That setting is significantly different than the oral reading context, which, in recent times, has dominated reading instruction in American elementary classrooms. In silent reading, students need to monitor their own comprehension. If they are reading too quickly, they need to accommodate their rate of reading to match their comprehension. They need to revisit a word if they couldn't figure it out the first time. By contrast, a teacher, tutor, or peer is present when students are reading orally. It's hard to stop and scan a page when someone is expecting an oral rendition of the text.

Without sufficient experience or strategies for silent reading, many students read slowly. In lieu of a teacher or tutor to monitor student performance, many students soon engage in less than efficacious behaviors in the silent reading context. Some eventually engage in counter-productive behaviors such as skimming. Mandates and interpretations

of mandates during the No Child Left Behind era may have exacerbated the situation, as more emphasis was placed on oral reading in assessment and instruction. But assessments of oral reading fluency and accuracy do not describe student performances in a silent reading context (Trainin, Hiebert, & Wilson, 2015).

The nature of silent reading and the manner in which it can be guided has been relegated to secondary status relative to oral reading. But ultimately, it is silent reading that is most important. In the tasks of colleges, workplaces, and communities—tasks such as voting, reviewing documents to purchase large-ticket items such as cars or houses, and seeking employment—individuals read silently, not orally. Silent reading patterns and examinations of instruction that supports efficient habits have received short shrift in both the research and pedagogical literature. The handful of programs of research that have been devoted to the topic are represented in this volume. These papers indicate that silent reading stamina can be improved through intentional instruction and teachers' design of tasks.

An Overview of the Volume

Three critical aspects of the topic of stamina in silent reading are addressed by the chapters in this volume: (a) describing the problem and the construct of stamina, (b) describing evidence that stamina can be increased through intentional instruction, and (c) describing the next steps in the design of instruction and research.

What Is Silent Reading Stamina and Why Is It Important?

The first section of this volume consists of three chapters that describe the construct of stamina and its importance in proficient reading. In the first chapter, I attend to three themes. The first of those themes is that stamina is a major challenge for many students in the U.S. This theme describes the problem that the research in the volume addresses. The second theme pertains to the association between silent reading proficiency and extensive reading opportunities. The third theme describes the need for appropriate instructional applications to ensure silent reading stamina, especially for students who are vulnerable or struggling readers.

To understand silent reading and to design activities that support its proficient development requires an understanding of the ocular-motor processes involved in reading. The second chapter by Samuels, Hiebert, and Rasinski provides information on how efficient eye movements are

foundational in proficient reading. They describe the processes in which the eyes engage during reading: saccades, fixations, and regressions. Most students, Samuels et al. note, master the complexities of eye movements on their own. But some students do not. For those students, opportunities need to be crafted that develop eye movement efficiency.

In addition to understanding basic processes such as eye movements, educators' beliefs about students' abilities to read silently are critical in designing instruction that supports stamina in silent reading. Often, teachers are reluctant to have students read independently because of the perception that many students do not engage in meaningful reading in independent contexts. The third chapter by Hiebert, Wilson, and Trainin provides information on the veracity of this perception. Specifically, they present data from a study of comprehension-based silent reading rates (CBSRRs) that addresses the rates at which students read silently with comprehension. Their findings indicate that the majority of students employ fairly consistent rates of reading with comprehension in independent contexts. A portion of students, at least at the fourth-grade level, do not perform reliably in silent reading contexts. Hiebert et al. suggest that this minority should not mean that all students are denied opportunities to read silently in classrooms. After all, students read and learn considerably more in silent reading than in oral reading because the former is typically faster than the latter. In addition, students with poor silent reading patterns require instruction that guides them in setting goals and monitoring their comprehension and reading rates in independent contexts. The longer the reading task is completely scaffolded by others (as is the case with oral reading), the longer it will take students to develop the habits of proficient reading.

What Types of Instruction Support Silent Reading Stamina?

The next section of the volume deals with instructional responses aimed at solving problems of poor silent reading habits among students. To read in a sustained fashion requires that readers have sufficient automaticity in recognizing words so that their attention is directed to the meaning of the message. Heidi Anne Mesmer's chapter highlights a particular feature of next-generation standards that increases the demands on stamina in silent reading: the mandate that students at all grade levels be able to read texts of higher complexity than was the case in previous standards. Mesmer uses the terms "stretch text" and "stretching" students' capacity in reading to describe the perspective evident in the staircase

of text complexity within the CCSS (National Governors Association, Center for Best Practices, & Council of Chief State School Officers, 2010). Mesmer's chapter sheds considerable light on demands for stamina in reading complex texts by (a) examining what is meant by complex text, (b) explaining how increased demands on text complexity could influence reader-text interactions, (c) explaining the rationale for stretch text, and (d) noting factors that may contribute or hinder students' increased capacity when texts stretch their reading capacity.

The next two chapters in this section describe a response to the problems related to stamina in silent reading. The response is an intervention in which the amount of reading is considerable. Even students whose silent reading habits are not efficacious read more in the intervention when students' reading is scaffolded digitally. Texts are matched to students' reading levels with daily assignments based on their ongoing comprehension performances. In addition, the digital context makes it possible to vary length of reading segments, number of comprehension questions, use of repeated readings, and assignment of prereading techniques. The results reported in these two chapters indicate that increasing silent reading of texts of appropriate levels influences students at all levels—from the end of the primary grades through high school.

In the first project that is described in Chapter 5, Rasinski, Samuels, Hiebert, Petscher, and Feller describe the results of the intervention with students from grades 4 through 10. Students who participated in the program for a minimum of 40 lessons (20 hours of instruction) over approximately 6 months made significantly greater gains on both criterion-referenced and norm-referenced tests than students who participated in alternative interventions. The gains were found generally in all grade levels studied and in all subpopulations, except for English learners.

Chapter 6 by Reutzel, Petscher, and Spichtig describes the efficacy of the same intervention with third graders. In this study, the authors compared the efficacy of the intervention to three other interventions. The students in the Reading Plus intervention demonstrated significantly superior performance on the state's reading assessment than students in the other interventions. Reutzel et al. conclude that the silent reading intervention afforded struggling third-grade students with appropriately challenging and varied reading genres to be both motivating and within their reach. These two reports, then, offer evidence that even struggling readers, when provided with scaffolded support, can develop stronger patterns of stamina in silent reading.

How Can Stamina Become a Focus of English/Language Arts Instruction?

Stamina in silent reading poses a substantial problem for the success of many students in attaining world-level literacy standards. The chapters in this section consider the steps that need to be taken by educators and researchers to make stamina in silent reading an integral part of students' school experiences.

In Chapter 7, Hiebert, Samuels, and Rasinski illustrate the efficacy of interventions that emphasize silent reading stamina at three developmental levels: primary, intermediate to high school, and young adult. The authors conclude there is sufficient support for initiating policies and practices in classrooms on all levels aimed at increasing silent reading stamina. They also conclude that the process of developing silent reading stamina extends through the elementary grades and into middle and high school as students encounter new genres and content. At least for the students who depend on schools to become literate, good silent reading habits require that they participate in structured silent reading experiences that model efficient reading.

The final chapter of the volume makes a final plea for attending to the critical proficiency of silent reading stamina. The chapter ends with a note of optimism. In particular, Ray Reutzel and I conclude that, while the digital age increases demands for literacy, it also offers increased opportunities. Digital contexts can support students who are especially vulnerable when they enter school or who have not been successful in typical learning contexts. Digital contexts can provide consistency of exposure to ensure that students are reading at appropriate levels and are staying on task. These opportunities can support students in successful participation in other literacy contexts including the large and small group and independent classroom contexts. Efficacious silent reading patterns depend on thoughtful and strategic actions that are part of interventions (such as the ones provided digitally) and typical instructional contexts.

The Importance of This Book for Teachers, Teacher Leaders, and Teacher Educators

The aim of this volume is to bring the topic of silent reading stamina into the mainstream of ELA instruction and research. New assessments that are considerably more rigorous and complex than previous assessments are already being implemented in many states. The

outcomes, which promise to be lackluster (Ujifusa, 2012, 2013), are likely to result in a great deal of hand-wringing among educators. The release of the assessment results will likely be accompanied by explanations and accusations on the part of pundits as they attempt to interpret results.

One missing element is an explanation of what the less than propitious results of students in many states on the new-generation assessments have to do with attention to the typical tasks of instruction and the manner in which they support students' ability to read silently for extended periods. Stamina, as the chapters in this volume illustrate, is critical for ensuring that students are ready for the tasks of college, communities, and the workplace as well as the new generation of assessments.

As the first volume to address the topic of silent reading stamina, *Teaching Stamina and Silent Reading in the Digital-Global Age* will be a useful guide for many constituencies. Among those who will benefit from this volume are teacher educators and professional developmental leaders who interact with teachers in courses and workshops. The volume is especially pertinent to supervisors and curriculum leaders in districts, states, and agencies such as regional laboratories who work in the translation of policies to practices. Further, graduate students and professors who study the efficacy of practice in supporting proficient student reading will find the volume useful in the design of research, especially regarding instructional interventions. The conclusions and suggestions offered in the chapters in this volume are intended to serve as grist for study groups of teachers, graduate and undergraduate courses, professional development sessions, and conversations among colleagues.

A Note of Gratitude

This book would not be possible without the generosity of a number of publishers and authors who gave permission to reprint several of the chapters in this volume. Readers who are unfamiliar with the legalities of academic book and journal publishers may be unaware that scholars retain the rights to their work until a manuscript begins the copyediting phase of publishing as a journal article or a book chapter. Colleagues at Taylor Francis were generous in acknowledging this policy and confirming that the papers that appear as Chapters 5, 6, and 7 could be used in the present volume.

A special note of gratitude is owed to both Tim Rasinski and Ray Reutzel. Tim generously wrote the foreword and, additionally, provided the

manuscript that was accepted for publication in *Reading Psychology*, which later became Chapter 5 in this volume. Ray was most generous in agreeing to permit the republication of chapters from *Revisiting Silent Reading* after the International Reading Association reverted the rights of the volume to Ray and me as coeditors. This generosity made possible the publication in this volume of Chapters 2, 3, and 8. In addition, he generously provided the accepted version of the manuscript provided to the *Journal of Educational Research* (Chapter 6 in this volume).

A skillful editor is truly a gift, and I thank Stacy Sharp who meticulously edited all of the prepublication manuscripts and the chapters commissioned for this volume. An individual who can produce an e-book is also priceless and, for serving that role for this volume, I thank Alice Folkins. She has produced many of the products of TextProject, but creating an e-book is a new venture that she bravely—and successfully—took on.

I conclude with thanks to Charley Fisher who handles the many logistics that make TextProject possible. Without Charley's generosity and unfailing willingness to attend to the details, this project—and many others at TextProject—would only be a dream. I hope this volume will help the many students who come to our schools every day with dreams of success see those dreams come true.

EHH
Santa Cruz, CA
June 2015

References

- Bielinski, J., Daniel, M., & Hiebert, E. H. (February 19, 2015). *Patterns of silent reading fluency and accuracy: What they mean for instruction and intervention*. Presentation at the annual meeting of the National Association of School Psychologists, Orlando, FL. Retrieved from: https://www.academia.edu/11191138/Patterns_of_Silent_Reading_Fluency_and_Accuracy_What_They_Mean_for_Instruction_and_Intervention
- California Board of Education. (2014). *English language arts/English language development framework*. Sacramento, CA: California Department of Education. Retrieved from <http://www.cde.ca.gov/ci/r/cf/elaeldfrmwrksbeadopted.asp>
- National Center for Education Statistics. (2013). *The nation's report card: A first look: 2013 mathematics and reading* (NCES 2014-451). Washington, DC: Institute of Education Sciences, U.S. Department of Education.
- National Governors Association, Center for Best Practices, & Council of Chief State School Officers. (2010). *Common Core state standards for English language arts and literacy in history/social studies, science, and technical subjects: Appendix A*. Washington, DC: Authors. Retrieved from http://www.corestandards.org/assets/Appendix_A.pdf
- Trainin, G., Hiebert, E. H., & Wilson, K. (2015). A comparison of reading rates, comprehension, and stamina in oral and silent reading of fourth-grade students. *Reading Psychology*. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/02702711.2014.966183>
- Ujifusa, A. (2012). Scores drop on KY's Common Core-aligned tests. Education Week Online. Retrieved from <http://www.edweek.org/ew/articles/2012/11/02/11standards.h32.html>
- Ujifusa, A. (2013). Tests aligned to Common Core in New York State trigger score drops. Education Week Online. Retrieved from http://blogs.edweek.org/edweek/state_edwatch/2013/08/one_interesting_aspect_of.html

I. UNDERSTANDING THE PROBLEM & THE CONSTRUCT

CHAPTER 1

The Forgotten Reading Proficiency: Stamina In Silent Reading

Elfrieda H. Hiebert

TextProject & University of California, Santa Cruz

The new assessments developed by the Smarter Balanced Assessment Consortium and the Partnership for Assessment of Readiness for College and Careers (PARCC) to align with the Common Core State Standards (CCSS; National Governors Association Center for Best Practices [NGA Center] & Council of Chief State School Officers [CCSSO], 2010) require all but the most severely disabled students to read and respond to texts in a digital context. Beginning at third grade, students are expected to read and respond to texts silently over extensive periods of time (see Table 1.1). And, unlike typical classroom reading tasks, students will have no access to teachers to present a first read or to help them by scaffolding a section of text, monitoring their reading, or advising them when it is time to start answering questions or writing responses.

Of course, extended silent reading is not a requirement limited to the new CCSS-related assessments. For the tasks of college, citizenry, and the workplace, we most often conduct reading tasks silently on our own for sustained periods of time. But it is highly likely that many students will not be prepared for the challenge of the silent reading tasks posed by the new assessments. The reason for this challenge is not—as pundits and observers of education frequently suggest—that American students cannot read. Indeed, most American students can read. What many students cannot do is independently maintain reading focus over long periods of time. The proficiency they lack is stamina—the ability to sustain mental effort without the scaffolds or adult supports.

In this chapter, I provide an overview for three themes that are echoed in the chapters of this book: (a) stamina is a major challenge for many American students, (b) silent reading proficiency depends on extensive reading opportunities, and (c) appropriate instructional applications can increase students' silent reading proficiency. First, however, I identify and define the constructs that are the foci of this

book—silent reading, comprehension-based silent reading rate, and the role of oral reading (including oral reading of instructional texts by teachers).

Table 1.1: Administration Times and Number of Sessions: CCSS Assessment Consortia¹

Grade	PARCC	SBAC
3	EOY ² : 60 min. x 2 sessions Perf: 40-60 min. per task TOTAL: Approximately 4.5 hours	CAT ³ : 1 hr. 45 min. Perf: 35 min. (stimulus + research Qs; 70 min. writing prompt) TOTAL: Approximately 3.5 hours
4-5	EOY: 70 min. x 2 sessions Perf: 50-80 min. per task TOTAL: Approximately 5 hrs. 50 min.	CAT: 1 hr. 45 min. Perf: 35 min. (stimulus + research Qs; 70 min. writing prompt) TOTAL: Approximately 3.5 hrs.
6-8	EOY: 70 min. x 2 sessions Perf: 50-85 min. per task TOTAL: Approximately 5 hrs. 55 min.	CAT: 1 hr. 45 min. Perf: 35 min. (stimulus + research Qs; 70 min. writing prompt) TOTAL: Approximately 3.5 hrs.
9-11	EOY: 70 min. x 2 sessions Perf: 50-85 min. per task TOTAL: Approximately 5 hrs. 55 min.	CAT: 2 hrs. Perf: 35 min. (stimulus + research Qs; 70 min. writing prompt) TOTAL: Approximately 4 hrs.

¹From Wixson (2013).

²EOY: End-of-Year

³CAT: Computer Adaptive Technology

Definitions and Distinctions

Oral reading assessments are a critical method for gaining insights into an individual’s mental processing capacities. Oral reading rate is an indicator of automaticity with words, so it’s not surprising that it has been shown to be a strong predictor of students’ comprehension. However, during the No Child Left Behind (NCLB) era, reading assessments often stopped with oral reading measures, ignoring a crucial fact: Ultimately, it is the silent reading performance of students that is most critical to their comprehension. After all, in the real world—whether we are college students, newly minted college graduates who are beginning their first jobs, or seasoned professionals—we are generally not asked to read articles or manuals orally. Silent reading, not oral reading, dominates.

Further, it is not the rate at which we read articles or manuals that is most valuable. What is critical is how well we understand, use, remember, and apply the content of what we read. Yet the rate at which this silent reading occurs can also be important. If readers read too slowly, it can create problems for both comprehension and memory. Consequently,

Hiebert, Wilson, and Trainin (2010) have introduced the construct of comprehension-based silent reading rate. Initially, we gave the construct the acronym CBSRR but, over time, we have shortened this to CSR, which stands for comprehension-silent reading rate. As this term implies, the emphasis of CSR is on establishing the rate at which students read silently with comprehension.

Stamina: A Challenge for Many American Students

Continuing a persistent trend, the reading scores for the most recent National Assessment of Educational Progress (NAEP) show that approximately one-third of our fourth graders score below the basic level and another one-third at the basic level (National Center for Education Statistics [NCES], 2014). Often, this pattern is interpreted to mean that our students cannot read, and the “solution” provided is to immerse them in more word-recognition instruction. Often the intervention programs chosen for use with struggling readers emphasize English grapheme-phoneme relationships, including with middle- and high-school students. But is the problem really that students cannot recognize words?

In the early 1990s, a group of scholars asked precisely this question. In response, the NAEP commissioned a special study in 1995 (Pinnell, Pikulski, Wixson, Campbell, Gough, & Beatty, 1995) and a follow-up study a decade later (NCES, 2005). Both studies involved a sample of fourth graders reading orally a portion of a text that had been part of the silent reading comprehension assessment. Students’ accuracy, fluency (i.e., prosody), and rate were assessed.

The two student samples in these special NAEP studies (Pinnell et al., 1995; NCES, 2005) did not read the same texts, but the texts are similar in their levels of complexity (around the end of the third-grade level). The two studies also were not precisely the same in terms of procedures (e.g., students in the 2002 sample read beyond one minute, while students in the 1992 sample read for a minute). However, the studies were similar enough to conclude that, within a representative sample of American fourth graders, the percentages of students who were reading with insufficient accuracy was relatively low. Clay (1985) deemed a 90% accuracy level as sufficient for determining whether students were reading a text adequately. The percentages of students within the two samples who were performing below 90% accuracy were similar at both assessment periods, as is evident in the information presented in Table 1.2: 2% of a cohort. The students who were reading with accuracy levels of 90 to 97% read more slowly

than students who were reading at the frequently cited independent level of 98% or higher (Betts, 1946). This pattern would suggest that students lack automaticity, not the fundamental ability to recognize words as is frequently assumed in policies and mandates. For example, the current California textbook requires required (California State Board of Education, 2014) that intervention programs for students in grades 4 through 8 contain decodable readers for each of the 43 phonemes and their graphemes.

Table 1.2: Accuracy Levels for Words Read without Meaning Change (Percentages) for Students Within Two NAEP Studies

	100-98%	97-95%	94-90%	<90%
1992 ¹	41	51	5	2
2002 ²	76	15	5	2

¹Pinnell et al., 1995

²National Center for Education Statistics, 2005

There is also evidence from DIBELS assessments (Dewey, Kaminski, & Good, 2013) that students (at least fourth graders) maintain a high level of word-recognition accuracy, even with texts that are more complex than the below grade-level texts in the two NAEP studies (Pinnell et al., 1995; NCES, 2005). As shown in Table 1.3, benchmark assessment passages for the end of the fourth grade on the DIBELS have approximately 2 more rare words per 100 than the NAEP passages. Unlike the passages in the two aforementioned NAEP oral fluency studies, which fall within the band for grades 2 through 3 on the CCSS staircase of text complexity (NGA Center, 2010, 2012), the DIBELS passages fall within the band for grades 4 through 5.

Table 1.3: Features of Fourth-Grade Passages (NAEP, DIBELS, and CSR Studies)

Text (Fourth-Grade Designation)	Lexile	Mean Sentence Length	Mean Log Word Frequency	Core Vocabulary
<i>Hungry Spider</i> (1992 NAEP)	660	10.70	3.69	96%
<i>Box in a Barn</i> (2002 NAEP)	620	10.29	3.72	96.5%
<i>The Youngest Rider</i> (end-of-year fourth grade DIBELS)	810	12.55	3.61	96%
<i>Temporary Homes</i> (Hiebert et al., 2011)	890	13.3	3.50	94%
<i>Theseus</i> (Hiebert et al., 2011)	800	12.5	3.60	90%

The DIBELS norms are based on approximately 167,000 students in kindergarten through grade 12 representing every census region in the U.S. (Dewey et al., 2013)—approximately 24,000 students per grade level. Table 1.4 provides accuracy, rate, and comprehension data for fourth graders. These data support the NAEP data, as even students at the 10th percentile display reasonable accuracy—95%. Their rate, however, is approximately 60% of the oral reading rate of typical grade-level readers. DIBELS developers have added a retelling measure to the assessment. Differing considerably from the comprehension measures typical of the NAEP and of the new CCSS-aligned assessments, this measure indicates that students’ challenges lie not in their ability to recognize individual words but in their ability to think about text.

Table 1.4: Fourth Graders’ Rate, Accuracy, and Comprehension on DIBELS (2011 to 2012 Cohort)

Percentile	Rate	Accuracy	Comprehension
10	80	95	21
20	98	97	27
30	109	97	32
40	118	98	36
50	128	98	41
60	138	99	45
70	147	99	50
80	160	99	57
90	176	100	67
99	212	100	94

A third source of information about American students’ word-recognition capabilities is a line of CSR studies that examined students’ reading rates and comprehension scores in silent reading contexts (Hiebert et al., 2010; Hiebert, Trainin, & Wilson, 2011). A similar context has been used in these two studies, one that replicates many norm-referenced reading assessments. Students read a set of short texts (each 200 to 250 words) about the same topic, and after reading a passage, they respond to multiple-choice comprehension questions.

In the first study (Hiebert et al., 2010), fourth graders read comparable texts in two different contexts: (a) digital and (b) print. For reading comprehension, no significant differences emerged across the two contexts. But for silent reading rate, differences did show up, with students reading significantly faster in the digital rather than the print context.

The Hiebert et al. (2010) study also considered differences in reading rate and comprehension across quartile groups. Rates for different

comprehension quartiles differed as a function of performance level and part of text. Students in the two lower quartiles started out at a reasonable rate, but their rates changed dramatically over the sections of the assessment (but not with increases in comprehension). The readers from the lowest quartile increased their speed after one passage but with lower levels of comprehension. The second-lowest quartile followed a similar pattern (i.e., increase in rate, decrease in comprehension) as the lowest quartile, but only after the first two sections of the assessment. The students in the top two quartiles had a stable rate and comprehension performance that changed very little across sections of the text.

In a subsequent study (Hiebert et al., 2011), fourth graders' performances on narrative and informational texts were compared. CSR was computed for the reading of each of four 250-word passages, and correct responses to four comprehension questions. For both the narrative and informational texts, percentages of students who attained the CSR level dropped steadily from the first text to the third. Whereas 85% of the students comprehended the first text, only 66% (narrative) and 56% (informational) attained the CSR criterion on the last texts.

An examination of the data by Hiebert et al. (2011) also identified the following six stamina patterns among students:

1. **Nonstarters** (i.e., students who did not attain the CSR criterion for any passage)
2. **Quitters after passage 2** (students who attained the CSR criterion on the two passages but engaged in rapid reading with insufficient comprehension on the first two subsequent passages)
3. **Quitters after passage 3** (students who attained CSR criterion on three passages but engaged in “fake reading” on the final passage)
4. **Monitors** (students who engaged in fake reading after failing to comprehend at least one text)
5. **Persisters** (students who, at best, attained a minimal level of comprehension on two texts but continued to engage with the same rate on other texts)
6. **Comprehenders** (students who attained the criterion on all passages)

The number of nonstarters was low (3%), but approximately 27% of the students fell into the quitters group and another 6% were classified as monitors. Of the remaining students, 56% were comprehenders and 8% were persisters.

This review of research leads to the conclusion that the vast majority of American students in an age cohort can recognize words—yet word recognition is the focus of most reading interventions. Although a

lack of automaticity in word recognition does appear to be an issue for the students in the bottom 5% or even 7% of a cohort, most students can recognize the core vocabulary. However, when they are asked to sustain their attention in silent reading, these students appear not to have the stamina that is required to interact with texts in a meaningful manner.

Silent Reading: Proficiency Depends on Reading Opportunities

For any given activity, whether it is highly demanding (e.g., performing brain surgery or playing a Rachmaninoff piano concerto) or prosaic (e.g., riding a bike or using a computer keyboard), it is absurd to think that we can become proficient without participating extensively in the activity. When it comes to teaching students to read, however, attention typically focuses on the nature of instruction rather than on the quality or quantity of deliberate practice time for students. For example, in the NCLB era, the five pillars of proficient reading identified by the National Reading Panel (NRP; NICHD, 2000)—phonemic awareness, phonics, fluency, vocabulary, and comprehension—became the focus of instruction. In the era of the CCSS, ensuring that students are engaging in close-reading strategies has taken center stage in discussions of pedagogy and implementation. Instruction about critical reading strategies and content is important, but instruction does not necessarily ensure that students have the opportunities they need to become proficient independent readers. For this to happen, students also need to have an abundance of occasions that allow them to take responsibility for getting meaning from a text, or as Guthrie, Schafer, and Huang (2001) have described it, students need opportunities to read. It is especially the case that students require opportunities to read silently in classrooms.

The research on the nature and effects of students' opportunities to read in classrooms is sparse at best. In the late 1970s, several research groups (e.g., Fisher, Berliner, Filby, Marliave, Cahen, & Dishaw, 1980; Leinhardt, Zigmond, & Cooley, 1981) examined the relationship of classroom time spent in silent reading to students' reading achievement. They found that students in classrooms where more time was devoted to reading practice and instruction attained higher levels of reading achievement.

More recently, an observation study of more than 1,000 first and second graders and their teachers (Foorman, Schatschneider, Eakin, Fletcher, Moats, & Francis, 2006) showed that of 20 time allocation

variables, it was only when time was allocated for text reading in classrooms that significant gains were found on any post-test measures (including word reading, decoding, and passage comprehension). No other time factors, including time spent on word recognition, alphabetic knowledge, or phonemic awareness instruction, independently contributed to reading growth. In another study, Kuhn and Schwanenflugel (2009) reported that the distinguishing feature in a large scale-up of an intervention was not in the results demonstrated by the intervention but rather the success of students in relation to the amount of time that they spent reading. Students in the seven most successful classes read seven minutes more each day than did the students in the seven least successful classrooms, regardless of whether classrooms were part of the intervention.

Observational studies over the decades have shown, however, that the percentage of school time students reading texts in many of classrooms is limited. Leinhardt et al. (1981) found that the amount of time that students spent reading was approximately 15% of the time allocated to reading instruction. Taylor, Frye, and Maruyama (1990) found that students spent an average of 15.8 minutes a day in either assigned reading or sustained silent reading (SSR).

All evidence points to the fact that, although the amount of time devoted to reading instruction increased during and following the NCLB era (Dorph et al., 2007), the amount of time that students actually spend reading has not increased substantially. Brenner, Hiebert, and Tompkins (2009) observed the amount—and kinds—of reading in which third graders participated in a sample of classrooms that were participating in a state's Reading First program. On average, across the 64 classrooms, teachers reported that they were devoting twice as much time to English language arts instruction than they had prior to the implementation of Reading First, but their students were involved with text less than 20% of the time, spending an average of 18 minutes a day reading text. This amount of reading practice is less than those amounts proposed by Allington (2001) and Fisher and Ivey (2006) but it was greater than the national average of 12 minutes a day reported by the NCES (1999). Even so, nearly a quarter of students did not read at all during the observed reading periods in the classrooms in Brenner et al.'s sample.

In the classrooms that Brenner et al. (2009) observed, less than 10% of total reading instructional time was allocated to unassisted reading, where students are responsible for reading texts on their own without teacher assistance or immediate monitoring. The small amount of time that students read on their own can be tied to interpretations that were

prominent as a result of the report of the NRP (i.e., NICHD, 2000), which concluded that there was insufficient evidence to support independent reading in classroom time. Teachers in the study had been informed of this finding as part of Reading First trainings and they appeared to follow this advice, even though the teacher's guides in their mandated core reading programs included in-school independent reading.

During the NCLB era, many educators extended the NRP's conclusion on independent reading to silent (or unassisted reading, as Brenner et al. called it) reading as part of instructional sessions (Allington, Billen, & McCuiston, 2015). This interpretation of this conclusion to independent, silent reading did not accurately reflect the studies on which the NRP based their conclusions—studies of SSR where students read texts of their own choosing and without teacher monitoring or scaffolding. The popular interpretation of this finding among educators, however, was understandable in that the NRP did not provide a highly nuanced description of the findings and also failed to include descriptive studies in their database such as the Manning and Manning (1984) study that showed that SSR was more effective when it included peer discussion or teacher conferencing.

Following the NRP report, Lewis (2002) analyzed a broader group of independent reading studies, many pertaining to students' silent reading. Out of more than 100 separate student samples that Lewis examined, the majority showed positive results for silent reading. The samples in most of the studies that reported no effects or negative growth from silent reading experiences consisted of students in fourth grade or above. Lewis speculated that because older students already have some reading proficiency, 10- to 15-minute silent reading periods—as was typical in these studies—may have been insufficient to significantly influence these students' performance. For students who were less-proficient readers (e.g., beginning readers, learning disabled, second- language learners), even such short periods typically produced benefits. Specifically, the studies suggest that when there is some form of scaffolding, students' silent reading proficiencies improve as a result of increased opportunities to read (Nunnery, Ross, & McDonald, 2006). Scaffolding may need to take numerous forms, including support for selecting appropriate texts (Mervar & Hiebert, 1989).

On the 1998 NAEP (NCES, 1999), fourth graders were asked to report the number of pages that they read daily in school. Even though a measure of self-reported reading is a rather simple tool (and not necessarily the most accurate), this measure predicted students' performances on

the NAEP. A follow-up study that focused specifically on the students within the state of Maryland confirmed that, after parental education was statistically controlled, the amount of engaged reading significantly predicted reading achievement on the NAEP (Guthrie, Schaffer, & Huang, 2001).

The survey used in the Guthrie et al. study used number of pages read to determine the amount of reading. In Table 1.5, I have converted pages read to number of words likely read by a hypothetical student in each of three proficiency groups on the NAEP, using the average number of words per page in a set of 100 fourth-grade texts. It is highly unlikely that all three hypothetical students, representing different proficiency groups on the NAEP, read at a similar rate (Pinnell et al., 1995; NCES, 2005), making the disparities in amount of text read daily in school by less-proficient and more-proficient students likely greater than the amounts shown in Table 1.5. But as Table 1.5 illustrates, even when a similar reading rate is used across proficiency levels, differences in amount of time spent reading in school mean that the poor readers keep getting poorer and the proficient readers keep getting better (Stanovich, 1986).

Table 1.5: Typical Reading Volume: Reading Levels of Three Hypothetical Students

	Alex	Alice	Abby
Daily reading in school (in minutes)	7.2	11	15
Daily # of words read (yearly total words)	715 (127,700)	1,100 (198,000)	1,485 (267,300)
Projected new words (with morphological family members)	290 (1,160)	446 (1,784)	601 (2,406)
Performance on NAEP	Below-basic	Basic	Proficient

^aSame reading rate used for all students: 100 wpm

Instructional Applications: Appropriate Opportunities Can Increase Students’ Reading Proficiency

Especially for students whose reading experiences occur primarily in school settings, a strong silent reading habit (of which stamina is a part) greatly depends on the experiences that their teachers provide them. The development of a habit like silent reading does not occur over the course of only a single grade. How children start out is incredibly important, but a habit is formed over an extended period of time—grade after grade in school. If students haven’t had the kind of support that develops solid silent reading habits over several grades, it is highly likely that changing direction and developing appropriate habits may require instructional

programs that are particularly well designed—often referred to as interventions.

One such instructional program that was carefully designed to increase silent reading proficiency for students who were still developing as readers was the project of Reutzel, Fawson, and Smith (2008). Reutzel et al. reconfigured SSR (where students read independently without substantial teacher monitoring or guidance) into Scaffolded Silent Reading (ScSR), in which students read widely in independent-level texts covering a range of genres but with periodic teacher monitoring and accountability. ScSR was compared to Guided Repeated Oral Reading (GROR), the approach that the NRP (NICHD, 2000) identified as effective. In GROR, third graders orally read a single text repeatedly, typically at grade level or instructional level, while receiving feedback from a teacher or other students. At the end of the yearlong study, Reutzel et al. concluded that the two forms of reading did not produce significant differences in students' fluency and comprehension. What this study showed is that, when students are guided in what and when to read silently, students' achievement is as good as that of students reading orally. In that silent reading is the proficiency that typifies most reading done by adolescents and adults, such scaffolded opportunities to read silently lay the foundation for subsequent tasks in a way that a heavy diet of oral reading in the primary grades does not.

One context in which consistency and adaptive solutions can be part of reading lessons is the digital environment. Online contexts give structure to learning experiences, which may be particularly valuable for struggling readers who have spent three or four years in classrooms where appropriate scaffolding has not been provided (Hiebert, Menon, Martin, & Bach, 2009). In a digital environment, there are ways to monitor students' involvement—which, of course, is a difficult thing to do with 25 or more students in a classroom. When considered relative to the approximately 1,200 hours most students spend in school annually, even a small amount of consistent support in an online context leads to considerable improvement in the CSRs of struggling readers in grades 3 and beyond.

Rasinski, Samuels, Hiebert, Petscher, and Feller (2011) found that as little as 20 hours of participation in a digital context over a school year resulted in improved performances on both a norm-referenced test (NRT) and a criterion-referenced test (CRT). Reutzel, Petscher, and Spichtig (2012) found that a similar digital intervention of additional reading was also efficacious in increasing the reading proficiency of struggling third-grade readers.

In a recent assessment of CSR completed by 350,000 students

from grades 2 through 12 (Hiebert, Spichtig, Bender, 2013), over 14% of the students could not comprehend a first-grade text. What is surprising is what these students gained from consistent reading—on computers—over a two-month period following the assessment. After only 10 hours of instruction that consisted of reading extended texts and answering comprehension and vocabulary questions, these students had moved from a 58% to 79% (on average) level of comprehension, moved to one grade level higher of text, and were reading an average of nine words faster (Hiebert et al., 2013). These students had sufficient word recognition—even the lowest scoring ones—to increase substantially in their comprehension on a first-grade passage. And this growth happened after students had read approximately 40,000 words over the course of 40 lessons. Even a relatively small increase in reading apparently can mean substantial increases in students' proficiency.

These reports (Rasinski et al., 2011; Reutzel et al., 2012; Hiebert et al., 2013) all indicate that there are instructional mechanisms that can support students in developing the reading habits that are needed for the 21st century—and that build on the research on cognitive and linguistic processes. But most teachers don't have access to digital technology such as that I have discussed, nor am I advocating that digital technology or a particular program is the solution to all reading problems. Instead, it is critical to consider the important components of various kinds of successful programs. Using knowledge about research, theory, and practice, I have generated seven actions that teachers can take to support increased stamina in silent reading. The actions are listed below.

1. Give students responsibility for the first read of texts.
2. Be explicit about the degree of challenge.
3. Have students make explicit goals for increased stamina and reading.
4. Increase the amount that students are reading.
5. Increase students' engagement in reading through connected homework reading and magazine articles.
6. Increase students' responses to texts through writing and discussions.
7. Have monthly "on your own" sessions using available sample assessments.

Individual teachers can implement these actions over the course of a school year with a cohort of students. Getting support in one year may make a difference (as was the case in the Rasinski et al., 2011 and Reutzel et al., 2012 studies). As the Hiebert and colleagues (2013) project indicates,

students can benefit even from several months of consistent and deliberate opportunities of increased silent reading. But for students who have developed poor reading habits in the early grades, the effort of creating strong silent reading patterns, including stamina, will likely require the involvement of teachers over several years of students' school careers. Opportunities need to be consistent and aimed at acquiring knowledge. The texts can't be vacuous—otherwise students won't be engaged in reading—but neither should the texts be far out of the realm of students' knowledge or their vocabulary expertise.

Conclusions

The need for efficient silent reading habits for success in the digital-global age is unarguable. There is emerging evidence that these habits can be enhanced through scaffolding, both on the part of teachers and from digital supports. These supports look quite different than the SSR that Hunt (1970) advocated in favor of. This structuring can begin when students are in the early stages of reading (Reutzel et al., 2008). Further, it is highly likely that the process is an ongoing endeavor, extending through the elementary grades and into middle and high schools as students encounter new genres and content. At least for the students who depend on schools to become literate, good silent reading does not just happen as a result of an emphasis on oral-reading fluency training. For many students, good silent reading habits require that they participate in structured silent reading experiences that model efficient reading. The target activities can be summarized as a succinct mantra (Hiebert, 2013) that provides the meanings for increasing stamina in silent reading: Read often. Mostly silently. Focus on knowledge.

References

- Allington, R. (2001). *What really matters for struggling readers: Designing research-based programs*. New York, NY: Addison-Wesley.
- Allington, R., Billen, M.T., & McCuiston, K. (2015). The potential impact of the Common Core State Standards on reading volume. In P.D. Pearson & E.H. Hiebert (Eds.), *Research Based Practices for Teaching Common Core Literacy*. (pp.161-178) New York, NY: Teachers College Press.
- Betts, E.A. (1946). *Foundations of reading instruction, with emphasis on differentiated guidance*. New York, NY: American Book Company.
- Brenner, D., Hiebert, E.H., & Tompkins, R. (2009). How much and what are third graders reading? In E.H. Hiebert (Ed.), *Reading more, reading better* (pp. 118–140). New York, NY: Guilford.
- California State Board of Education (July 9, 2014). *The ELA/ELD Framework*. Sacramento, CA: Author.
- Clay, M.M. (1985). *The early detection of reading difficulties* (3rd ed.). Portsmouth, NH: Heinemann.
- Dewey, E.N., Kaminski, R.A., & Good, R.H., III. (2013). *2011-2012 DIBELS net system-wide percentile ranks for DIBELS Next*. Eugene, OR: Dynamic Measurement Group.
- Dorph, R., Goldstein, D., Lee, S., Lepori, K., Schneider, S., & Venkatesan, S. (2007). *The status of science education in the Bay Area: Research brief*. Berkeley, CA: Lawrence Hall of Science, University of California, Berkeley.
- Fisher, D. & Ivey, G. (2006). Evaluating the interventions for struggling adolescent readers. *Journal of Adolescent & Adult Literacy*, 50(3), 180—189.
- Fisher, C.W., Berliner, D.C., Filby, N.N., Marliave, R., Cahen, L.S., & Dishaw, M.M. (1980). Teaching behaviors, academic learning time, and student achievement: An overview. In C. Denham & A. Lieberman (Eds.), *Time to learn: A review of the beginning teacher evaluation* (pp. 7–32). Washington, DC: U.S. Department of Education, National Institute of Education.
- Foorman, B. R., Schatschneider, C., Eakin, M. N., Fletcher, J. M., Moats, L. C., & Francis, D. J. (2006). The impact of instructional practices in grades 1 and 2 on reading and spelling achievement in high poverty schools. *Contemporary Educational Psychology*, 31(1), 1-29.
- Guthrie, J.T., Schafer, W.D., & Huang, C.W. (2001). Benefits of opportunity to read and balanced instruction on the NAEP. *The Journal of Educational Research*, 94(3), 145–162.
- Hiebert, E.H. (August 30, 2013). *Reading rules for becoming proficient with complex texts* [Frankly Freddy]. Santa Cruz, CA: TextProject. Retrieved from <http://www.textproject.org/library/frankly-freddy/reading-rules-for-becoming-proficient-with-complex-texts/>
- Hiebert, E.H., Menon, S., Martin, L.A., & Bach, K.E. (2009). *Online scaffolds that support adolescents' comprehension* (Research Brief). Seattle, WA: Apex Learning.
- Hiebert, E.H., Spichtig, A., & Bender, R. (2013). *Building capacity in low-performing readers: Results of two months of Reading Plus practice* (Research Brief 2.1). Winooski, VT: Reading Plus. Retrieved from <http://www.readingplus.com/results/research-briefs/>.
- Hiebert, E.H., Trainin, G., & Wilson, K. (July 15, 2011). *Comprehension and reading rates across extended grade-appropriate texts*. Paper presented at the annual conference of the Society for the Scientific Study of Reading, St. Petersburg, FL.
- Hiebert, E.H., Wilson, K.M., & Trainin, G. (2010). Are students really reading in independent reading contexts? An examination of comprehension-based silent reading rate. In E.H. Hiebert & D.R.

Reutzel (Eds.), *Revisiting silent reading: New directions for teachers and researchers* (pp. 151–167). Newark, DE: International Reading Association.

Hunt, L.C., Jr. (1970). The effect of self-selection, interest, and motivation upon independent, instructional, and frustration levels. *The Reading Teacher*, 24(2), 146–151, 158.

Kuhn, M.R., & Schwanenflugel, P.J. (2009). Time, engagement, and support: Lessons from a 4-year fluency intervention (pp. 141–161). In E.H. Hiebert (Ed.), *Reading more, reading better*. New York, NY: Guilford.

Leinhardt, G., Zigmond, N., & Cooley, W.W. (1981). Reading instruction and its effects. *American Educational Research Journal*, 18(3), 343–361.

Lewis, M. (2002). *Read more—read better? A meta-analysis of the literature on the relationship between exposure to reading and reading achievement*. Unpublished doctoral dissertation, University of Minnesota, Twin Cities.

Manning, G.L., & Manning, M. (1984). What models of recreational reading make a difference? *Reading World*, 23(4), 375–380.

Mervar, K., & Hiebert, E.H. (1989). Literature selection strategies and amount of reading in two literacy approaches. In S. McCormick & J. Zutell (Eds.), *Cognitive and social perspectives for literacy research and instruction* (38th Yearbook of the National Reading Conference; pp. 529–535). Chicago, IL: National Reading Conference.

National Center for Education Statistics. (1999). *NAEP 1998 reading report card for the nation and states* (NCES 1999-500). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics. (2005). *The nation's report card. Fourth-grade students reading aloud: NAEP 2002 special study of oral reading* (NCES 2006-469). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics. (2006). *The nation's report card. Reading 2005* (NCES 2006-451). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics. (2014). *The nation's report card. A first look: 2013 mathematics and reading. National Assessment of Educational Progress at grades 4 and 8* (NCES 2014-451). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects with Appendices A-C*. Washington, DC: Authors.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2012). *Supplemental information for Appendix A of the Common Core State Standards for English language arts and literacy: New research on text complexity*. Washington, DC: Author. Retrieved from <http://www.corestandards.org/resources>.

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. 00-4769). Washington, DC: U.S. Government Printing Office.

Nunnery, J.A., Ross, S.M., & McDonald, A. (2006). A randomized experimental evaluation of the impact of Accelerated Reader/Reading Renaissance implementation on reading achievement in grades 3 to 6. *Journal of Education for Students Placed at Risk*, 11(1), 1-18.

Pinnell, G.S., Pikulski, J.J., Wixson, K.K., Campbell, J.R., Gough, P.B. & Beatty, A.S. (1995). *Listening to children read aloud: Data from NAEP's integrated reading performance record (IRPR) at*

grade 4 (Rep. No. 23-FR-04). Washington, DC: Office of Educational Research Improvement, U.S. Department of Education.

Rasinski, T., Samuels, S.J., Hiebert, E., Petscher, Y., & Feller, K. (2011). The relationship between a silent reading fluency instructional protocol on students' reading comprehension and achievement in an urban school setting. *Reading Psychology*, 32(1), 75–97.

Reutzel, D.R., Fawson, P.C., & Smith, J.A. (2008). Reconsidering silent sustained reading: An exploratory study of scaffolded silent reading. *The Journal of Educational Research*, 102(1), 37–50.

Reutzel, D.R., Petscher, Y., & Spichtig, A.N. (2012). Exploring the value added of a guided, silent reading intervention: Effects on struggling third-grade readers' achievement. *The Journal of Educational Research*, 105(6), 404–415.

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360–407.

Taylor, B.M., Frye, B.J., & Maruyama, G.M. (1990). Time spent reading and reading growth. *American Educational Research Journal*, 27(2), 351–362.

Wixson, K.K. (April 24, 2013). *Key shifts in assessment and instruction related to CCSS-ELA* [webinar]. TextProject. Retrieved from <http://www.youtube.com/watch?v=IHYcJAX0AO8>

CHAPTER 2

Eye Movements Make Reading Possible¹

S. Jay Samuels

University of Minnesota-Twin Cities

Elfrieda H. Hiebert

TextProject & University of California, Santa Cruz

Timothy Rasinski

Kent State University

The ability to read and understand printed words represents a remarkable human accomplishment. Although the ability to communicate through the spoken word seems to have been genetically hardwired into our species over the eons of time it has taken our species to develop (an estimated 5-8 million years), the skill of reading has been with us only for about 7,000 years. Because of the huge time differences between the development of language by ear versus language by eye, there appears to be some design flaws in the human eye that must be overcome before reading can occur. In essence, as remarkable an instrument as is the human eye, it is not ideally constructed for reading. An argument that we make in this chapter is that without eye movements, reading alphabetic texts would not be possible.

A century ago, the study of eye movements was one of the hottest topics in reading psychology. In the classic volume *The Psychology and Pedagogy of Reading*, Huey (1908/1968) devotes two chapters to eye movements. Despite this auspicious start, it is not the hot topic in reading that it once was, as evidenced in Cassidy and Cassidy's (2009) "What's Hot for 2009" survey in the United States. In this list, ocular-motor eye movement is not listed as a topic to be rated by the experts. Because of the critical role that eye movements play in the reading process, the topic should be of interest to educational leaders at all levels who desire to see improvements in reading achievement. This chapter on eye movements in

¹ This chapter was previously published in *Revisiting Silent Reading: New Directions for Teachers and Researchers*. The definitive publisher-authenticated version published in 2010 and in 2014 is available online at: <http://www.reading.org/general/Publications/Books.aspx> & <http://textproject.org/library/books/revisiting-silent-reading/>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

reading should prove to be useful to educators in understanding how some reading problems that beginning readers encounter can be traced to faulty but correctable eye movements. When students' eye movements become more accurate and effective, schools can anticipate gains in reading achievement (Gelzer & Santore, 1968).

When reading scholars were asked to explain why eye movements are not hot today, responses indicated a perception of eye movements as purely mechanical, unrelated to cognitive or social processes. It is true that eye movements seem purely mechanical when the reading process is going well. Eye movements, however, are influenced by cognitive factors, such as the need to locate information of personal interest or to reread a portion of the text to do a comprehension check (Just & Carpenter, 1980). These cognitive factors influence the duration of eye fixations and where the eye searches for information in texts.

Kaakinen and Hyönä (2008) examine how eye movements mirror the ongoing cognitive processing in which readers engage. In their study, half of the readers were told to read a passage about a house from the point of view of a burglar, while the other half were told to take the perspective of an interior designer. Both groups wore eye-tracking apparatus that indicated which parts of the text received attention and the duration of eye fixations. Those who read from a burglar's point of view had greater gaze duration time on words that dealt with how one might burglarize a house, while those who read from an interior designer's point of view spent more time on words that dealt with what made the house attractive.

In a similar study (Sipel & van den Broek, 2009) where the emphasis was on text comprehension, college students read a text that contained recently learned rare words, common words, and unknown words. Eye tracking revealed that students spent more time fixating on unlearned rare words than on recently learned rare words. The extra duration of eye fixations on the rare, unlearned words may also reflect greater cognitive emphasis on decoding and meaning generation of the rare, unlearned words. These studies strongly suggest that eye movements and gaze duration, rather than being purely mechanical and immutable, seem to be under the cognitive control of readers and influenced by factors such as personal interests and purpose for reading.

Numerous scholars have attempted to explain how reading is made possible. Given the number of different models of the reading process (see Ruddell & Unrau, 2004), one may rightfully wonder if we need yet another description of the reading process. The answer to this question may be found in the brilliant poem entitled "The Blind Men and the Elephant"

(Saxe, 1873). In the poem, each of six blind men described the elephant from his perspective. Saxe claims that each blind man was partly in the right, though all were in the wrong. Like the blind men, each model of the reading process in the research literature on reading describes only a part of this complicated miracle called reading, and more information is needed to fill in the missing parts in the reading mosaic.

As remarkable as the human information processing system is, there are areas that can be considered to be design flaws insofar as reading is concerned. In essence, the eye was not designed for reading. In this chapter, we explain how eye movements and selective attention represent ways in which human beings overcame the information processing design flaws in the human eye. Furthermore, as we reviewed a considerable body of eye movement research, we became aware that most of it has been done using a convenience sample of adults. Keith Rayner (personal communication, May 10, 2009), one of the leading researchers in the field of eye movements, agrees that it is important that researchers learn more about the eye movements of beginning readers, especially young children. When, in the past, information that was derived from adults was used either to understand beginning reading or to justify the methods used in beginning reading instruction, it led to serious and regrettable consequences. Keeping this admonition in mind, we describe how ocular-motor eye movements make reading possible and, in fact, overcome some of the bottlenecks that are part of our human information processing system. To accomplish this goal, we begin by indicating how Javal's (1879) discovery set the stage for research on eye movements. Then we describe the physiology of the eye and the eye movements that enable reading to occur. In the final section, we suggest implications that can lead to enhanced reading achievement in classrooms.

Discovery of Eye Movements

More than a century ago, it was a commonly held belief that the eye uninterruptedly and smoothly took in information as it swept along a line of print or when looking at a scene outdoors. Contradicting this common belief, Javal (1879) found that the eye seemed to jump from spot to spot and then paused during reading. He concluded correctly that the eye took in information only when it paused. He called these ocular-motor eye movements saccades. Dodge (1900) supported Javal's conclusions indicating that when the eye movements were unbroken the observer was unable to tell what had been exposed. In fact, before an eye movement

occurs, vision is suppressed to prevent the reader from seeing the blur that occurs during a saccade (Latour, 1962).

Whether the content is print or the view out of a window, eye movements occur in what Hochberg (1970) describes as installments. There are three types of ocular-motor eye movements that occur during reading:

1. Fixations—when the eye pauses momentarily on a line of print to take in information or integrate information across fixation pauses.
2. Forward saccades—when reading English script the eye seems to jump from left to right on a line of print to bring the eye to the next fixation pause.
3. Regressions and rereadings—where eye movements occur backward from right to left.

Generally, regressions go back about one word, whereas rereading allows the eye to reexamine a previously fixated portion of the text. If the saccade extends back several words, we identify this as a rereading saccade, not a regression. In rereading, a student moves quite a few words back to a prior section of a line and then proceeds in a usual manner to reread from that point forward as a comprehension check.

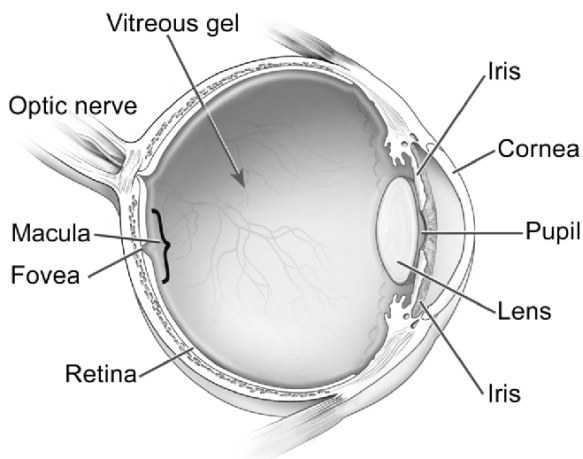
Human Eye Physiology

As depicted in Figure 2.1, the human eye contains three major parts. The first part is the cornea, located in the front of the eye. It acts like a window and allows light waves reflected from visual images on the printed page to pass through so they can settle upon the retina, which is located at the back of the eye. The second part is the retina, which primarily consists of two kinds of receptor cells: rods and cones. Some of these are sensitive to letter and word shape. The third part consists of a collection of communication wires called the optic nerve that carry information from the retina to the visual perceptions areas of the human brain.

The cornea of the eye contains a hole called the pupil through which visual information from the page passes through on its way to the retina. Surrounding the pupil is the colored portion called the iris, which contains muscles that alter the size of the opening of the pupil. Under dim light, the opening of the pupil is larger to admit more light and, under bright light, the opening is smaller to admit less light. Located directly behind the pupil is the lens that has the function of focusing the visual images from the page as sharply as possible on the retina. The retina

contains cells that function like the film in a camera. These retinal cells are sensitive to particular bands of light wavelengths (i.e., red, blue, green) and fire only when the right wavelength causes the rhodopsin—a pigment in the retina—to react. The mosaic, or pixel-like, electrochemical impulses are sent to the brain via the optic nerve, where they are reconstructed to make an image. Some specialized parts of the brain, in turn, control the ocular-motor eye movements that we discuss shortly.

Figure 2.1. The Human Eye



Graphic from National Eye Institute, National Institutes of Health.

A key idea in this chapter is that the human eye is not ideally designed for reading. Although the eyes are designed to move to perceive things, the typical perception pattern of a visual image differs from that of a line of print. Consequently, the eyes need to learn to make particular kinds of movements if proficient reading is to occur. Imagine that you are trying to identify the person who is standing in front of you. As you look at this person, all that is in focus is the person's nose and eyes. The rest is fuzzy, but you can detect shape. You rapidly shift your points of focus to other parts, so that in time the various parts of the individual are in focus. The difficulty in determining the identity of this person is somewhat similar to the problem of recognizing words when reading a text.

The problem with the eye when reading is that at any given moment only a tiny amount of printed material from a page is in enough focus to enable easy reading. Consequently, rapid eye movements are required to bring different parts of a text onto that tiny area on the retina that can see the letters and words clearly—the fovea. The retina of the eye

contains two kinds of cells, rods and cones. Both kinds are important and have different reading functions. Cone cells provide the visual acuity that enables readers to see letters and words clearly. A major design flaw of the eye in regards to reading is that the cone cells that enable the reader to see letters most clearly are not evenly distributed across the retina but are concentrated in a tiny area called the fovea. There are about 10 million cone cells packed into the fovea of each eye where vision is most acute—where the reader can identify letters and words with precision. Because of the fovea's location in the middle of the retina, print has to be front and center and not off to the side.

Further, the parafoveal region surrounding the fovea also plays a critical role in reading. The parafovea contains the rod cells that are sensitive to word shape and word length. Information received in the periphery of the eye helps guide the eye to its next fixation destination (Rayner & Sereno, 1994). The spaces surrounding words are important clues as to word boundaries and length, and this information is used by peripheral vision to plan the distance the eye should jump with each saccade. In essence, the rod cells are part of the eye's guidance system.

Because the area of cone cell concentration is small, the number of letters that can be in focus within a single eye fixation is limited. We contacted two leaders in the field of eye movement research to get information on how many letters can be perceived by the fovea at any given time. Keith Rayner (personal communication, May 10, 2009) stated,

The number of letters falling in the fovea depends on letter size and viewing angle. In general, 3-4 letters usually occupy 1 degree of visual angle. Because the fovea is about 2 degrees, it would be 6-8 letters in the fovea.

The second expert, George McConkie (personal communication, May 10, 2009) stated,

The foveal region is the area where...visibility of letters drops off pretty fast as they move outward from the center of vision. Thus, the problem in answering this question is setting a "clarity" criterion. I suppose that a criterion might be even the most similar letters such as *v* and *u* or *o* and *c* can be distinguished at this distance.... What Keith and I were after in our original studies was to determine the region within which letter distinctions make a difference. We found this to be about four letters to the left and eight to the right of the directly fixated letter. The greater distance on the right is probably (there is some supportive evidence) an attentional factor rather than retinal resolution differences to left and right.

Legge et al.'s (2007) research suggests that only six or seven letters surrounding the fixation point on the fovea can be identified with 80% accuracy and, as the eye moves farther away from the fixation point,

accuracy of identification decreases more. For example, within four letter spaces to the left of the fixation point, or eight letter spaces to the right of the fixation point, accuracy of identification drops to about 60%. In summary, the evidence from the experts is that the size of the perceptual span from which letters can be seen with accuracy and clarity falls in a range of six to eight letters. It also appears that the shape of the window is asymmetrical, with fewer letters in focus to the left of fixation and more letters in focus to the right of fixation. It is also commonly acknowledged that there is a rapid drop-off of acuity from the point of visual focus that makes word recognition difficult (Feinberg, 1949).

The experts make a good point when they say that there is no hard and fast rule about the number of letters that are in focus on the fovea. The number of letters in focus is a function of letter size and the distance at which they are being viewed. However, the experts agree that the number of letters in focus is not large. Thus, one bottleneck in the reading process is that only a small portion of the text on a line can be clearly identified. Other letters that fall to the right and left of the fovea experience a steady and rapid decline in clarity. One way to overcome this rapid loss of clarity is to shift focus through eye movements so that different parts of words that are not clear come into focus on the next eye movement. The problem of attempting to shift focus so that the desired part of the text is in focus on the fovea is somewhat analogous to the problem facing hunters who try to keep a moving target within the cross hairs of a riflescope. It is a difficult task, because eye movements make it easy to overshoot or undershoot and miss the target. Rod cells shift to the next eye fixation so that different parts of a text are in focus (Smith, 1971). While the cone cells aid in identifying letters and words, the rod cells help the brain plan the trajectory of how far to move the point of focus for each new eye fixation. In addition, when words are printed in lowercase letters, the words take on skyline and shape characteristics, and the rod cells are capable of picking up word shape information (Lee, Legge, & Ortiz, 2003).

The span of apprehension refers to the number of letters the eye can see in a single fixation. One might think that the span would be symmetrical around the fixation point, but this seems not to be the case. Instead, the span is asymmetrical with more letters recognizable to the right of fixation for those reading in English, whereas the span of apprehension is greater to the left of the fixation point for those reading in Hebrew or Arabic. These asymmetrical differences in the span of apprehension reflect how text is written and processed in each of the languages (e.g., English from left to right and Hebrew and Arabic from

right to left). The span, then, is attention driven or learned rather than the result of a fixed pattern (Hebb, 1930). Moreover, if the words to be recognized are low frequency and unfamiliar to the reader, the span of apprehension is smaller than if the words are familiar high-frequency words (Rayner, 1998).

Eye Physiology and Cognitive Psychology

The eye has additional properties that are relevant to the reading process. Sperling's (1960) research convinced skeptical psychologists that the eye has a memory, or what has come to be called iconic memory. This memory buffer can be viewed in terms of four characteristics: (1) speed of input, (2) speed of output, (3) capacity, and (4) longevity. Speed of input and output refers to how fast the images in the visual field can be placed on the retina and how fast the images can be retrieved when needed. The capacity of the eye to take in the visual scene is huge and includes everything in a circular visual field. As Sperling's research indicates, however, the longevity of the image that is placed on the retina is short. The typical image lasts only for one or two seconds at most, and then the image is gone. Although the longevity of the word's image on the fovea is short, the reader can overcome this problem by fixating the same word several times. When this is done, image two can function as an erasing image and the second image can obliterate image one (Gilbert, 1953). With each repeated fixation of the same word, additional decoding occurs and the process is repeated until the word is recognized.

Because the capacity to take in information from the visual field is so huge, it presents a processing problem. The human brain can process only a limited amount of information in a short amount of time. In a single eye fixation, there is more information than the brain can handle. Consequently, the brain filters the unwanted information and focuses on what it needs. For example, when viewing a scene outside, the view is circular. In reading, however, the reader does not want circular information from the visual field because he or she would then be getting information from lines of print that are located above and below the line on which he or she is focusing. If the reader were aware of the information on lines of print above and below the fixation point, it would lead to confusion. Thus, the eye needs to filter out information above and below the line that is being processed (Willows, 1974).

Selective Attention

The mechanism the brain uses to filter out the unwanted information is selective attention (McConkie & Rayner, 1976; Posner, 1980). Selective attention is an internal mechanism that filters out of the visual field that which is not important, and by doing so it also allows the reader to focus attention on the areas that are important (LaBerge & Samuels, 1974). To illustrate the nature of selective attention, researchers have created laboratory analogs to what has been described as the “cocktail party phenomenon”—where an individual engaging in a conversation continues to maintain eye contact with another person even when he or she has switched his or her internal cognitive attention to an adjunct conversation. Even though the person is essentially filtering out the immediate conversation to attend to the adjunct one, there is no observable change in body position or receptors such as eyes or ears.

Evidence that the eye has a memory buffer that enables selective attention came from Sperling (1960). The experiment involved flashing 12 numbers on a screen in three rows of four numbers, one each at the bottom, middle, and top of the screen. As the 12 numbers disappeared from the screen, they were followed by a tone (e.g., high, middle, low) that signaled at which level the numbers were to be read back. This task is difficult on several levels. First, the number of numbers—12—exceeds the capacity of short-term memory, so they cannot be memorized. Second, subjects needed to read back the numbers even though they were no longer on the screen. Third, they needed to provide the tones that came after the visual image had disappeared. To succeed at this task, subjects had to do an internal scan using selective attention and read the numbers off their visual memory buffer. Subjects were successful with this task before the visual image disappeared from memory.

Several characteristics of selective attention allow it to be a useful companion to eye movements and reading. One feature is filtering out unwanted information, such as text above and below the line the reader is on. Although it is true that only about 6-8 letters that fall across the fovea can be seen clearly, surrounding letters can be discerned. Without resorting to another eye movement, the reader can do an internal attentional scan to read the other letters of interest. If, however, the letters cannot be identified, another eye movement is required. For readers who are not yet at the point where the entire word has become the unit of word recognition, selective attention allows them to attend to parts of words as they decode them (LaBerge & Samuels, 1974). Selective attention also

allows readers to switch attention—whether on parts of a word or the entire word, on decoding and then on comprehension, or on the letters in focus and then on word shape in the periphery. Attention switching is fast, under cognitive control, and can be automatic or controlled (Shiffrin & Schneider, 1977).

Visual Unit of Word Recognition

If the processing problems created by the fact that only 6-8 letters strung across the fovea are seen with acuity were not enough, there is an additional problem that strikes hard on the beginning reader and less so on the skilled reader. This problem relates to the size of the visual unit used in word recognition. Cattell (1947), using a convenience sample of German graduate students, concludes that the unit of word recognition was the word. Gough (1971), on the other hand, concludes that the unit of word recognition was the letter, and each subsequent letter in a word added approximately 50 milliseconds (ms) to how long it took to recognize it. To determine which researcher was correct with regard to the size of the visual unit used in word recognition, the first author and colleagues (Samuels, LaBerge, & Bremer, 1978) devised an experiment in which words were shown on a computer screen. If the word on the screen was an animal word, the student pressed a handheld button, and the computer measured latency of response (i.e., reaction time) and accuracy. The animal words were all controlled for word frequency, and there were three-, four-, five-, and six-letter words.

The rationale for the study was simple. If Gough (1971) was correct and the unit of recognition was the letter, then the longer animal words should take more time to recognize, but if Cattell (1947) was correct, there should have been no difference in processing time related to word length because a chunk is a chunk and a word is a word. In our study, we had subjects from second, fourth, and sixth grades and college. Our results were fascinating, because Cattell and Gough were each correct, but for different age groups. The beginning readers were processing words letter by letter as Gough had predicted and longer words were taking them more time to process, while the sixth graders and the college students were using the whole word as the unit of recognition, supporting Cattell's contention that the unit of recognition was the whole word. For sixth graders and college students, there was no significant difference in processing time related to word length. In other words, they were processing words as entire units, and a chunk is a chunk. Fourth graders in this study showed

increases in the size of the unit of word recognition. Their unit of word recognition was larger than the single letter but not yet whole word (Samuels et al., 1978).

There have been several replications of this study, including a significant change in method, and the results are robust and hold up. In fact, the research by Taylor (1971) on the number of eye fixations required to read a 100-word text supports the research findings on the size of the visual unit used in word recognition. By 11th grade, only 96 eye fixations were required, implying that with each eye fixation the unit of recognition was the word. In first grade, however, 224 fixations were required, suggesting that the unit of recognition was smaller than a word. As we have already discussed under selective attention, there is an internal scan mechanism that is used for processing the letters that are on the fovea and letters that extend slightly beyond.

The size of the visual unit used in word recognition is an important factor in eye movement. Imagine how hard it must be for the beginning reader to place the target word on the fovea and then to process a word unit that is smaller than the entire word. To add to the difficulty, the processing must be fast enough that the word fragments put into short-term memory are not lost (Peterson & Peterson, 1959). The student must figure out the meanings of the words that are placed in short-term memory in less than 10 seconds or what was placed there will be lost and the process must be repeated. In addition, the wrong part of the text may fall on the fovea with the fixation pause. Considering some of the processing bottlenecks that have been identified, such as the fact that only about six letters are in focus in a single eye fixation, beginning readers may have difficulty with the accuracy of eye movements. Given these problems that must be overcome by beginning readers, it is not surprising that learning

Table 2.1. Development of Ocular-Motor Skills From 1st Through 12th Grades

Ocular-Motor Skills	1st Grade	12th Grade
Fixations (including regressions per 100 words)	224	96
Regression per 100 words	52	18
Average span of recognition in words	0.45	1.06
Average duration of fixation in seconds	0.33	0.25
Reading rate with comprehension	80	250

to read with skill takes time and practice. Table 2.1 shows the development of ocular-motor skills from first through 12th grades.

The Fixation Pause

Eye fixations in reading are critical because it is during a fixation that the eye takes in information from the printed page and begins to process it for meaning. The duration of the typical fixation pause is about 300 ms, which is about one third of a second (pauses can be as short as 100 ms or as long as 500 ms, which is 1/10 to 1/2 of a second). It is assumed that during longer fixations considerable cognitive processing is going on, such as attempting to grasp the meaning of a sentence or integrating information across several sentences. “While the word *fixation* implies that the eye is motionless, this is not the case. There is a slight eye tremor that serves to activate the neurons in the retina so they will continue firing (Gilbert, 1959). Taking into account the brief amount of time it takes to make a forward saccade in which the eye moves from one fixation pause to the next, in a single second the eye can make approximately three fixations. “When viewing a scene or a page of printed material, the typical person seems to be unaware that the information being processed by the brain has been coming in at a rate of three bursts a second and that each burst must be processed rapidly, because the visual image coming with each burst survives for less than a second and then it is lost. If, however, the processing is too slow and the visual image disappears from the retina, all is not lost. The reader can refixate the original image. The term *eye fixation pause* represents the time spent on a single fixation, whereas the term *gaze duration* suggests the total amount of time the reader spends on a word across several eye fixations.

Because of the rapid loss of the visual image from a fixated word or word part, what the reader must do is transform the visual image into its sound representation. For example, when the reader encounters the printed word *cat*, it is transformed into its phonological form /c-a-t/ and then placed in short-term memory.

The advantage gained by transferring visual into phonological information and placing the phonological information in short-term memory is that the shelf life of the acoustic information in short-term memory is about 10 seconds, which is considerably longer than the duration of visual information in iconic memory, which is less than 1 second (Peterson & Peterson, 1959). For the acoustic information that is in short-term memory, 10 seconds is usually sufficient time (in most

cases) for skilled readers to complete tasks such as decoding the text, integrating sentence meaning, and finally, moving the information that was temporarily stored in short-term memory into long-term-semantic memory.

Because eye physiology is such that the eye takes in different kinds of information from three areas—foveal, parafoveal, or peripheral—the total span of information is large. Beginning readers have a span of apprehension that is 12 letters to the right and skilled readers have a span of 15 spaces, although we cannot assume that words can be recognized that far out, but word length and shape information is obtained (Ikeda & Saida, 1978; Rayner, 1998). Foveal information enables one to identify words, whereas the parafoveal area provides information about shape and length (Rayner, Well, & Pollatsek, 1986). McConkie and Rayner (1976) have shown that as skill increases the span of recognition increases, but not beyond one or two words.

To the person reading a text or viewing a scene outdoors, the entire operation appears to be seamless. It is the seamless nature of the operation that led to the mistaken belief before Javal's (1879) time that the eye continuously took in information and simultaneously processed it as the eye swept smoothly across a page of print. In terms of transfer of training, it seems as if several of the eye-movement mechanisms used in viewing a scene outdoors are also used in reading a text. The saccade serves as the setup to get the right information in focus. However, information is not taken in while the eye is in motion during a saccade. It is only during the fixation pause that the brain assembles information for processing. The number of fixation pauses per second for viewing a scene outdoors is about the same as for reading a text (Taylor & Robinson, 1963) and comes to about three fixations per second. According to Feinberg (1949), the number of letters that fall on the fovea that can be seen clearly comes to about four or five—the same as the number of letters in a high-frequency word. Thus, if the reader is skilled and the unit of word recognition is the word, he or she should be able to process three words per second and be able to read at a rate of about 180 words a minute with comprehension, which is a little short of the figure that Germane and Germane (1922) report as the silent reading rate for good readers in the eighth grade.

An important question that eye movement researchers have addressed is whether the eyes fixate on each word in a text or skip certain words. It appears that the eye skips certain words, and the words that are skipped are determined in part by word length and skill level. Short words, high-frequency words, and words that can be predicted from context may

be skipped (Brysbaert & Vitu, 1998; Paulson & Goodman, 1999).

Gilbert (1940, 1959) notes that oral reading is slower than silent reading. This simple fact poses a problem in many classrooms where round robin reading is practiced. In round robin reading, one student reads orally from a text while the other students follow along reading silently from the same text. However, when a poor reader reads orally, with typical slow reading rate and lack of expression, it forces the better readers who are reading silently into twice as many eye fixations and regressions. Gilbert's concern was that this round robin reading practice was training poor ocular-motor habits in students. Gilbert cautions teachers that this common practice should be discontinued. The practice, however, was deeply entrenched in typical reading instruction.

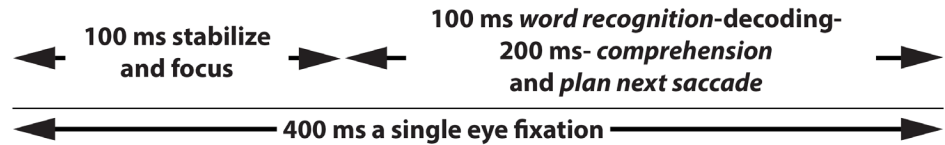
One reason for the entrenchment of round robin reading in practice is that it provided a means whereby teachers could get a sense of how numerous students were progressing in their reading from day to day. Informal reading inventories, given in a one-to-one setting, might provide more valid information, but they were costly in terms of teachers' time and, consequently, given infrequently (Pikulski & Shanahan, 1982). Deno (1986) provides a viable alternative for monitoring that was not as costly in terms of teachers' time and, as a result, freed up instructional time for other forms of oral reading, such as choral or echo reading or even scaffolded silent reading. Dena's solution was to have students read orally for one minute and to count the number of correct words read in that brief period of time. By keeping a running record on each student's reading rate over a period of time, teachers could determine if there was improvement in rate up to some asymptote. As good as Dena's method is, there is a problem. The problem is that comprehension is not measured, only rate. Despite warnings that meaning should not be sacrificed for the sake of reading rate, some teachers continue to encourage rate and students fail to put attention on meaning. Because of the problems associated with using only reading rate to measure progress, the time has come for researchers to develop a testing method that focuses attention on comprehension as well as reading speed.

As we noted, the typical eye fixation pause lasts for about 300 ms, or about one third of a second. Even a pause this short can be separated into components representing the different processing tasks that must be performed to read with understanding (Abrams & Zuber, 1972). The typical pause comes at the end of an eye movement when the eye has just completed a rapid movement from one spot on a text to the next spot,

somewhat like an automobile that comes to an abrupt stop at a stop sign. There is still residual motion that must be halted and stabilized, and in the case of the eye, it must be stabilized so that it can focus on the print.

Figure 2.2 is a simplified rendition of what takes place during an eye fixation of a skilled reader. Essentially, five tasks are performed with each fixation pause. The first task following a saccade is to stabilize the eye. Once the eye is stabilized sufficiently, the next task is to focus the visual images from the page on the fovea of the retina. With the visual image from the page focused on the retina, the third task is to engage in word recognition, or what many call the decoding process—converting the word into its sound representation. If the reader is highly skilled and automatic at word recognition, the task is done quickly and accurately, and it requires a minimal amount of cognitive resources and attention. Although the typical duration of an eye fixation is 0.33 seconds, in many instances it may take longer.

Figure 2.2. Activities That Occur During a 400 Ms Eye Fixation by a Fluent Reader



If a person is a skilled reader, the amount of time required for the word recognition process may be only 100 ms, leaving 200 ms for comprehension—the fourth task. In fact, the defining characteristic of fluency is the ability to decode and comprehend in the same eye fixation. For skilled fluent readers, the decoding task is done so quickly and requires so little of the cognitive resources that comprehension can take place at the same time (LaBerge & Samuels, 1974). The final task for the reader is to plan the next saccade (Abrams & Zuber, 1972). For fluent readers, the usual unit of word recognition is the word. A word is defined as a letter or series of letters surrounded by space. Space is a critical cue used by the rod cells in planning the trajectory for the next leap, which is probably the next word. For readers not at the automatic stage of word recognition, there are some important differences in what happens during an eye fixation. First, the word recognition process is usually slower, less accurate, and may use up all of the cognitive resources available at the moment. Thus, during that one eye fixation, the single major accomplishment for nonfluent readers is word recognition. To add to the complexity of word

recognition for the nonfluent reader, the unit of word recognition is smaller than the entire word, leaving the student in the position of having to piece together the letter clusters that in combination make up the word. Because only 6-8 letters that are on the fovea are in focus along with some other letters to the right that are not so distinct, the student may resort to selective attention to process the letter cluster. However, once the student has recognized the word, his or her next task is to switch attention to the comprehension process. This constant switching of attention from decoding to comprehension places a heavy load on short-term memory and makes learning to read so much harder for the less skilled reader than for the accomplished reader.

Figure 2.2 is an important visual because it strikes at the heart of the debate on what is reading fluency. This visual shows that, within an eye fixation, skilled readers can decode and comprehend what is in the text. Unfortunately, beginning readers cannot do both tasks simultaneously. They first decode the text, and then they try to comprehend what they have decoded. The products of the dual process are stored briefly in short-term memory, but the decoding and comprehension tasks must be completed within 10 seconds of that memory system or what was briefly stored is lost. If beginning readers lose what was stored, they repeat the process. The second attempt is faster because of the previous encounter with the text.

During an eye fixation, exactly what part of the word the eye is focused on is important if the reader wants to infer the word using only partial information. Different parts of a word vary in the amount of information they provide the reader. Broerse and Zwaan (1966) found that not all parts of a word are equally informative for purposes of word recognition. They found that it is the beginning part that carries the most information for purposes of word recognition. For example, if the reader has already identified the following context "Father was cutting the green....," and the letter string on the fovea for the next eye fixation contains the following "gr_ _ _," it is an easy task to infer that the next word is grass. Paulson and Goodman (1999) believe that under certain conditions the reader may skip words in a text and, if context is strong enough, use partial information to infer the word. However, Taylor (1971) is of the opinion that typically the reader does not try to infer a word from its parts. Despite this claim, Paulson and Goodman report there are times when words are inferred and recognized through their parts. In planning an eye movement, the preferred location for a fixation is halfway between the beginning and middle of a word (McConkie, Kerr, Reddix, & Zola, 1988) because, given the span of apprehension, and the typical length of common

words, the highly informative beginning of a word would be on the fovea.

An important question about the role of eye fixations in reading is how much information is taken in and processed with each fixation pause. The answer is that the eye provides the brain with information from three areas, the foveal, parafoveal, and the peripheral. Of the three, the foveal area is most important because it is here that the letters are in focus. The foveal area extends 2 degrees of visual angle for a maximum of 8 letter spaces asymmetrically distributed around the point of focus, with fewer letters in focus to the left of fixation and more to the right (McConkie & Rayner, 1976), but, as Feinberg (1949) has noted, beyond 4-5 letter spaces from the fixation point, there is a sharp drop-off in clarity. However, for skilled readers the amount of information that is available in a single eye fixation is usually sufficient to permit rapid identification of the word.

Although the parafoveal and peripheral areas do not provide sharp, detailed information, they provide important information in a number of ways. First, there is word length information (e.g., short words may be skipped). Word length information is provided by the spaces that skilled readers use in the decoding process. Second, there is word shape information. Words printed in lowercase have a characteristic shape or skyline that aids word recognition. In addition, the space surrounding words is used in planning the trajectory for the next saccade. To illustrate how difficult reading becomes when word shape and length information are eliminated, try reading the following sentence:

ONLYRECENTLYHAVEEYEMOVEMENTSAND EYEFIXATIONSBEENRE
COGNIZEDFORWHATTHEYREALLYARETHEYAREUSEDINTHEWORD
RECOGNITIONPROCESS

The division of printed words by spaces is a relatively recent invention that turns out to be a most useful cue to readers. Gaur (1992) has stated that the division of words and sentences developed only gradually, and these changes occurred between 600 and 800 AD. The majority of ancient scripts did not use space to divide words and sentences. The reason is that the scribes who wrote the texts were so well versed they did not need any aids as to word boundaries. It was not until about the year 1200 AD that monks preparing medieval manuscripts began to include spaces so that readers who were less skilled could determine where the word boundaries were, and it is this very word boundary information that is used today when the brain plans the next saccade. If the saccadic movement is incorrect and the eye overshoots the target, the flow of meaningful information can be interrupted and the reader may have to self-correct by means of regressive eye movements. Just as the duration

of eye fixations varies as a function of reader skill, the number of eye fixations reflects reader skill as well. From Taylor's (1971) research, we learned that to read a 100-word text, 1st graders needed 224 eye fixations and 12th graders needed only 94 fixations. There are yet other factors that should influence eye fixations. For example, how do the goals of the reader influence eye movements? At times, the goal may be to study a text carefully to pass an exam, while at other times the reader desires only a casual, surface-level overview of a text. Surely, we might wonder how these differences in goals for reading influence visual factors such as duration of eye fixations, span of apprehension, the distribution of attention over the text, the length of a saccade, and regressions.

Regressions and Rereading

To advance through an English text from beginning to end, the direction of eye fixation advances from left to right. There are eye fixations when reading English texts that move in the opposite direction, and go from right to left. Fixations following right-to-left eye movements, excluding return sweeps from one line of print to the next, may be considered regressions. Some scholars differentiate between regressions and rereading. Taylor (1971) believes that some regressions that serve no purpose may reflect poor habits formed during the learning-to-read stage, and these inefficient habits may persist for long periods of time. Other regressions, however, may be purposeful and indicate that the reader has encountered an unanticipated word and is going back to do a comprehension check.

Regression may occur for any number of reasons. For example, in the earliest stages of learning to read, the student must learn how to adjust the accuracy of each eye movement. Pointing to the words is a strategy that many beginning readers—and even more proficient readers when the task is challenging—use (Ehri & Sweet, 1991). This aid (i.e., fingerpointing) is evidence that some young children are aware that they need to train themselves where and how to look at print.

Taylor's (1971) research uncovered sizable differences in the number of regressive eye fixations made as a function of reading skill. In 1st grade, for example, Taylor found that for every 100 words read, the students made 52 regressions while the 12th graders made only 17 regressions. How does one account for this large difference in backward eye movements between unskilled and skilled readers? One reason identified by researchers for such regressions pertains to poor habits that are acquired in the early grades

that need to be overcome to some extent with increased skill in the later grades. A second reason acknowledges the need for comprehension checks, which may require regressions. A third reason to regress may occur when the eye misses its mark during a saccade and the reader tries to adjust with a regressive eye correction.

We suggest a fourth possibility. When beginning readers attempt to construct meaning from the text, they engage in a two-step process: They decode the words, and then they attempt to get their meanings. During this process, the decoded words or word parts are moved to short-term memory where they are held for 10 seconds before they are lost. Once lost, the readers must start the process again. Speed is of the essence in this process. We have all noticed during oral reading how beginning readers laboriously work their way through a sentence, stop, and then regress back to an earlier section of text and start over. What has happened is the students took too long and ran out of time, and what was temporarily stored in short-term memory was lost. Therefore, the students had to regress and start over. Twelfth graders make only 17 eye regressions per 100 words read but first graders make about 52. Not only do less skilled readers make more backward eye movements but also the duration of each eye fixation is longer, which accounts in part for the slower reading speeds of the less skilled readers. Text difficulty also influences eye movements, with increases in text difficulty usually accompanied by increases in the duration of the fixation pause. Low-frequency, unfamiliar words in text are fixated longer, the distance the eye moves with each saccade decreases, and more regressions occur as more comprehension checks are needed.

Forward Saccades

When reading English, forward saccades are characterized by left-to-right eye movements. During an eye movement, vision is suppressed because the movement is so fast that the brain cannot process the information. The amount of time required to move the eye from fixation to fixation requires only 1/20th of a second. The distance the eye moves in each forward saccade ranges between 1 and 20 letter spaces, with the average being 4-5 letter spaces—the length of a shorter word. It would appear, then, that for skilled readers, for whom the unit of word recognition is the word, the eye jumps from word to word. For skilled readers, what controls the distance the eye jumps with each saccade are the rod cells, which are sensitive to the spaces that mark word boundaries. Ideally, the saccade would place the image of the word so that the letters

are spread across the fovea of the eye where letters are in focus. As we move away from the focal point, clarity of the letters decreases, and in fewer than 10 letter spaces out from the point of fixation, visual acuity has dropped by 45% and ease of word recognition becomes more of a problem (Feinberg, 1949; Legge et al., 2007). Consequently, as Rayner (1983) states, the planning of how far to move the eye with a forward saccade is critical.

It is by means of the forward eye movements that the reader is able to advance through a text from its beginning to its end. As important as the forward eye movements are, they exact a heavy price. The price is that they slow down reading speed and impair comprehension. It has been shown, for example, that when readers look at a point on a computer screen and all the words from a text are presented one at a time to that point, very high rates of reading accompanied by modest comprehension can be obtained—somewhere between 700 to 1,000 words per minute. This procedure, however, that requires no eye movements embodies a serious problem. It prevents the reader from making regressions that are essential for comprehension checks (Rayner & Pollatsek, 1989).

In summary, readers are able to overcome the limitations presented by the fact that in any given instant the eye can only see with clarity about one short word or eight letter spaces—through several kinds of eye movements: forward saccades, regressions, and fixations. In the next section, we examine problems that readers can experience with eye movements.²

What Educators Should Know About Eye Problems in Reading

Although there is some disagreement about the extent to which abnormalities of the eye itself and eye movement deficits lead to reading problems, we take the position that eye abnormalities and ocular-motor deficiencies can contribute to reading problems for inexperienced as well as experienced readers. Certainly, lack of visual acuity for distance viewing can be picked up through Snellen eye charts and corrected through properly fitted glasses, but the charts are not useful for detecting problems in the close-up viewing that is required in reading. Tracking can be a common and persistent source of difficulty, where readers have trouble maintaining the focus of the eye on a line of print. Some readers who have a tracking problem may skip entire lines. Even skilled readers

² Reference to the next section is referring to materials in *Revisiting Silent Reading: New Directions for Teachers and Researchers*.

may have this problem, especially when the lines of text are long. In fact, the tendency to lose one's place when reading long stretches of text across the page led many newspapers across the United States in the 1950s to adopt the practice of using narrow columns of text as a way to reduce eye-tracking problems in reading (Tinker, 1958).

Ideally, each eye should coordinate with the other, and both eyes should work as a team. When both eyes are working properly, we have binocular coordination. When there is a lack of binocular coordination, the effort it takes to read can become prohibitive. A somewhat related eye problem is convergence insufficiency. When reading, it is necessary for the eyes to turn inwards toward each other as well as to focus on the letters of the words that are being read. Convergence insufficiency is the condition whereby the ability of the eyes to converge and focus properly is compromised. If this occurs, students may experience blurred or double vision, headaches when trying to read, and burning and tearing of the eyes.

There can be other issues that stem from the fact that eye movements such as forward saccades and regressions are controlled by the muscles of the eye and, as such, are similar to motor activities found in skiing, golf, and tennis, all of which respond positively to training and practice. Each of these sports has a learning curve, starting with nonaccuracy and advancing with extended practice to mastery and beyond that to automaticity. When a skill reaches the automatic stage, it can be performed accurately and without conscious thought about its execution. Eye movements such as the forward saccades and regressions also require learning and practice to perform them accurately and automatically. These ocular-motor eye movements are difficult for the reader because he or she must estimate how far to move the eye from one fixation point to the next. Gauging how far the eye is to jump is not an automatic activity and is a complex skill that must be learned over an extended period of practice. The sports examples mentioned earlier may be easier to learn than gauging how far to move the eye, because skiing, tennis, and golf are open to observation by a coach who can see what is being done wrong and correct the athlete's mistakes. Under ordinary classroom conditions, the eye movements are invisible to the outsider. This "black box" phenomenon makes coaching the reader virtually impossible. The black box phenomenon is a term psychologists use to describe the workings of the mind that are hidden to an observer.

The sophisticated equipment that researchers such as Rayner and McConkie—the experts we contacted for this chapter—have used to study eye movements is unlikely to be affordable or practical in most

classroom settings. However, digital technology is advancing rapidly, as video recording on cell phones and flip camcorders become cheaper and more accessible. The capacity of the technology will need to be matched with ways of analyzing eye movements. There is equipment that is more reasonable in cost and provides data almost as accurate as that of the equipment typically used in research settings (Spichtig, Vorstius, Greene, & Radach, 2009). It is unlikely that equipment such as this will be used routinely in classrooms but, for comprehensive diagnoses of struggling readers in particular, such information may well add insight that is not available from typical assessments. Diagnosis needs to be coupled with plans for addressing the inefficient patterns. There, too, digital programs hold promise in developing more productive eye movements that underlie efficient reading (Marrs & Patrick, 2002).

Conclusions About the Essential Role of Eye Movements in Reading

At one time it was thought that the eye of the reader swept along the lines of text and continuously processed texts for meaning. However, Javal's (1879) research more than a century ago showed that in reading the process was not characterized by a continuous uptake of information. Rather, it appeared as if the eye jumped from one point on a line of text to another, and the critical components of reading such as word recognition and comprehension occurred only during the brief fixation pauses. These fixation pauses occurred at a rate of about three fixations per second. Because the physiology of the eye is such that only about 6-8 letters at a time can be seen with clarity, eye movements are the means by which different parts of the text can be sequentially processed.

A critical aspect of eye movements is the guidance system used to gauge how far to move the eye from one focal point to the next. For skilled readers for whom the unit of word recognition is the word, the spaces surrounding words are used as cues to guide the jumps as the reader advances from one word to the next. For beginning and less skilled readers, however, the unit of word recognition is some unit that is smaller than the entire word. What guides the beginning reader is not clear, because the bulk of the eye movement research that has been done has used skilled adult readers. With the many advances that have come with eye-tracking equipment, the time has come to learn more about the ocular-motor processes of less skilled readers. One possibility as to how less skilled readers process a text is that in addition to the eye fixation pause

when letters from a word are displayed in focus across the fovea, they rely partly on an internal shift of attention, or an internal scan, to process parts of words that are on their retina.

One thing seems certain when considering all involved in the learning-to read process effective and efficient eye movements are critical. Although many students master the complexities of eye movements on their own, there are many others who require additional help. Not only do less skilled readers need help with eye movements, but also even skilled readers working on advanced academic degrees can show marked improvement in their ocular-motor efficiency and reading achievement after receiving additional training and practice with eye movements.

Reference

- Abrams, S.G., & Zuber, B.L. (1972). Some temporal characteristics of information processing during reading. *Reading Research Quarterly*, 8(1), 40-51.
- Broerse, A.C., & Zwaan, E.J. (1966). Informative value of initial letters in the identification of words. *Journal of Verbal Learning and Verbal Behavior*, 5(5), 441-446. doi:10.1016/S0022-5371(66)80058-0
- Brysbaert, M., & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 125-148). Oxford, England: Elsevier.
- Cassidy, J., & Cassidy, D. (2009, February/March). What's hot for 2009. *Reading Today*, 26(4), 1, 8-9.
- Cattell, J. (1947). *Man of science: Psychological research*. Lancaster, PA: Science.
- Deno, S.L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review*, 15(3), 358-374.
- Dodge, R. (1900). Visual perceptions during eye movement. *Psychological Review*, 7(5), 454-465. doi:10.1037/h0067215
- Ehri, L.C., & Sweet, J. (1991). Fingerprint-reading of memorized text: What enables beginners to process the print? *Reading Research Quarterly*, 26(4), 442-462. doi:10.2307/747897
- Feinberg, R. (1949). A study of some aspects of peripheral visual acuity. *American Journal of Optometry and American Academy of Optometrists*, 26(2), 49-56.
- Gaur, A. (1992). *A history of writing*. New York: Cross River.
- Gelzer, A., & Santore, N.J. (1968). A comparison of various reading improvement approaches. *The Journal of Educational Research*, 61(6), 267-272.
- Germane, C.E., & Germane, E.G. (1922). *Silent reading: A handbook for teachers*. Chicago: Row, Peterson & Co.
- Gilbert, L.C. (1940). Effect on silent reading of attempting to follow oral reading. *The Elementary School Journal*, 40(8), 614-621. doi:10.1086/457813
- Gilbert, L.C. (1953). Speed of processing visual stimuli and its relation to reading. *Journal of Educational Psychology*, 50(1), 8-14. doi:10.1037/0945592
- Gilbert, L.C. (1959). Saccadic movements as a factor in visual perception in reading. *Journal of Educational Psychology*, 50(1), 15-19. doi:10.1037/h0040752
- Gough, P.B. (1971). One second of reading. In J.F. Kavanagh & I.G. Mattingly (Eds.), *Language by ear and by eye: The relationship between speech and reading* (pp. 331-358). Cambridge, MA: MIT Press.
- Hebb, D.O. (1930). *Organization of behavior: A neuropsychological theory*. New York: John Wiley & Sons.
- Hochberg, J. (1970) Components of literacy: Speculations and exploratory research. In H. Levin & J.P. Williams (Eds.), *Basic studies on reading* (pp. 74-89). New York: Basic.
- Huey, E. (1968). *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press. (Original work published 1908)
- Ikeda, M., & Saida, S. (1978). Span of recognition in reading. *Vision Research*, 18(1), 83-88. doi:10.1016/0042-6989(78)90080-9
- Javal, L.E. (1879). Essai sur la physiologie de la lecture [Essay on the physiology of reading]. *Annales d'Oculistique*, 82, 242-253.
- Just, M.A., & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329-354. doi:10.1037/0033-295X.87.4.329
- Kaakinen, J.K., & Hyönä, J. (2008).

Perspective driven text comprehension. *Applied Cognitive Psychology*, 22(3), 319-334. doi:10.1002/acp.1412

LaBerge, D., & Samuels, S.J. (1974). Towards a theory of automatic information processing in reading. *Cognitive Psychology*, 6(2), 293-323. doi:10.1016/0010-0285(74)90015-2

Latour, P.L. (1962). Visual threshold during eye movements. *Vision Research*, 2(7-8), 261-262. doi:10.1016/0042-6989(62)90031-7

Lee, H.-W., Legge, G.E., & Ortiz, A. (2003). Is word recognition different in central and peripheral vision? *Vision Research*, 43(26), 2837-2846. doi:10.1016/S0042-6989(03)00479-6

Legge, G., Sing-Hang, C., Yu, D., Chung, S., Lee, H., & Owens, D. (2007). The case for the visual span as a sensory bottleneck. *Journal of Vision*, 7(2), 1-15.

Marrs, H., & Patrick, C. (2002). A return to eye movement training? An evaluation of the Reading Plus program. *Reading Psychology*, 23(4), 297-322. doi:10.1080/713775286

McConkie, G.W., Kerr, P.W., Reddix, M.D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations on words. *Vision Research*, 28(10), 1107-1118. doi:10.1016/0042-6989(88)90137-X

McConkie, G.W., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the Psychometric Society*, 8, 365-368.

Paulson, E., & Goodman, K. (1999). Eye movements and miscue analysis: What do the eyes do when a reader makes a miscue? *Southern Arizona Review*, 1, 55-62.

Peterson, L.R., & Peterson, M.J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 58(3), 193-198. doi:10.1037/h0049234

Pikulski, J.J., & Shanahan, T. (1982). Informal reading inventories: A critical analysis. In J.J. Pikulski & T. Shanahan (Eds.), *Approaches to the informal*

evaluation of reading (pp. 94-116). Newark, DE: International Reading Association.

Posner, M.I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, 32(1), 3-25. doi:10.1080/00335558008248231

Rayner, K. (1983). *Eye movements in reading: Perceptual and language processes*. New York: Academic.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372-422. doi:10.1037/0033-2909.124.3.372

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Hillsdale, NJ: Erlbaum.

Rayner, K., & Sereno, S.C. (1994). Eye movements in reading: Psycholinguistic studies. In M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 57-82). San Diego, CA: Academic.

Rayner, K., Well, A.D., & Pollatsek, A. (1986). Asymmetry of the effective visual field in reading. *Perception & Psychophysics*, 27(6), 537-544.

Ruddell, R.B., & Unrau, N.J. (Eds.). (2004). *Theoretical models and processes of reading* (5th ed.). Newark, DE: International Reading Association.

Samuels, S.J., LaBerge, D., & Bremer, C.D. (1978). Units of word recognition: Evidence for developmental changes. *Journal of Verbal Learning and Verbal Behavior*, 17(6), 715-720. doi:10.1016/S0022-5371(78)90433-4

Saxe, J.G. (1873). *The poems of John Godfrey Saxe* (Complete ed.). Boston: James R. Osgood.

Shiffrin, R.M., & Schneider, W. (1977). Controlled and automatic human information processing. *Psychological Review*, 84(2), 127-190. doi:10.1037/0033-295X.84.2.127

Sipel, B., & van den Broek, P. (2009). *Where readers look when reading recently learned and unknown words*. Unpublished

manuscript. Holland: Leiden University.

Smith, F. (1971). *Understanding reading: A psycholinguistic analysis of reading and learning to read*. New York: Holt, Rinehart and Winston.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 74(11), 1-30.

Spichtig, A., Vorstius, C., Greene, A., & Radach, R. (2009, August). *Visagraph III and Eyelink2K: A comparison study*. Poster session presented at the 15th European Conference on Eye Movements, Southampton, UK.

Taylor, S.E. (1971). *The dynamic activity of*

reading: A model of the process. New York: Educational Developmental Laboratories.

Taylor, S.E., & Robinson, H. (1963, February). *The relationship of the oculo-motor efficiency of the beginning reader to his success in learning to read*. Paper presented at the meeting of the American Educational Research Association, Chicago, IL.

Tinker, M.A. (1958). Recent studies of eye movements in reading. *Psychological Bulletin*, 55(4), 215-231. doi:10.1037/h0041228

Willows, D.M. (1974). Reading between the lines: Selective attention in good and poor readers. *Child Development*, 45(2), 408-415. doi:10.2307/1127962

CHAPTER 3

Are Students Really Reading in Independent Reading Contexts? An Examination of Comprehension-Based Silent Reading Rate¹

Elfrieda H. Hiebert

TextProject & University of California, Santa Cruz

Kathleen M. Wilson & Guy Trainin

University of Nebraska, Lincoln

After a recent presentation by one of the authors (Kathleen), a teacher asked, “My students act like they are reading when reading silently, but how do I know if they are really reading?” This teacher’s question reflects a concern of many teachers. Recently, however, teachers have not been the only ones asking questions about the efficacy of silent reading. As a result of the conclusions of the National Reading Panel (NRP; National Institute of Child Health and Human Development, 2000) that sustained silent reading has not proven particularly effective in increasing fluency and comprehension, policymakers and administrators have raised questions about the effectiveness of silent reading during instructional time. The NRP’s conclusions regarding the efficacy of oral, guided repeated reading have meant an emphasis on oral reading experiences in the primary grades as evident in classroom observations (Brenner, Hiebert, & Tompkins, 2009) and in textbook programs (Brenner & Hiebert, 2010). At the same time, the Panel’s conclusions regarding the lack of substantive empirical literature that confirms the efficacy of independent, silent reading experiences on comprehension have meant, at least in the primary grades, a deemphasis on silent reading (Brenner et al., 2009).

Ultimately, however, most of the reading that adults, adolescents, and even middle- and upper elementary-grade students do is silent.

¹ This chapter was previously published in *Revisiting Silent Reading: New Directions for Teachers and Researchers*. The definitive publisher-authenticated version published in 2010 and in 2014 is available online at: <http://www.reading.org/general/Publications/Books.aspx> & <http://textproject.org/library/books/revisiting-silent-reading/>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Unarguably, the ability to read extended texts on one's own (i.e., silently) with comprehension is the foundation of proficient reading. The products and processes of comprehension are frequently the focus of researchers and educators. However, one dimension that is infrequently addressed is the rates at which students are reading with meaning. The topic of rate of silent reading has often been equated with speed reading. We are not suggesting a return to the speed-reading craze of the 1960s, nor are we advocating the obsession with speed that has become the interpretation of oral reading fluency during the last decade.

There can be little doubt that demands for efficient and effective silent reading have increased as the amount of information available to citizens of the digital-global age increases. The form of reading in which we are interested has comprehension at its center. Within a focus on comprehension, we believe that there is room for attention to the rates at which students are reading, particularly whether students are reading at appropriate rates. The digital revolution has meant that there are potential ways to address these reading rates and for determining whether they are appropriate for the tasks confronting students. We have termed the construct in which we are interested as comprehension-based silent reading rate (CBSRR).

Teachers in our graduate courses and workshops have asked numerous questions about CBSRR, such as the one that introduces our chapter. We delved into the research literature to answer these questions as well as our own questions. Our search for answers, however, produced few definitive responses. With only a few exceptions (e.g., Carver, 1990, 1992), researchers have not addressed CBSRR over the past decades. While the lack of a robust research surprised us, it also served as an impetus. We initiated a study that considered several persistent questions about CBSRR. We could not address all of the critical questions in a single study, so we raise some of our many remaining questions at the end of the chapter. We were able, however, to provide preliminary answers to some critical questions about CBSRR in the study we describe here.

This chapter provides a summary of responses to the three foci of our study:

(1) How do students of different quartiles vary in their CBSRR? (2) How well do students sustain their CBSRR across an extended text? (3) How consistent is the CBSRR of students in a digital context relative to a paper-and-pencil context? Before describing the design and findings of this study, we provide an overview of what is and is not known about CBSRR and our three foci.

A Review of CBSRR

The term *comprehension-based* is central to our definition of CBSRR. The digital age has made an abundance of information available to human beings, unlike any volume experienced by previous generations. While offering unique opportunities for learning and communication, this surfeit of information places demands on readers for higher level comprehension processes more than those demands of previous eras. Full participation in the digital-global marketplace and community demands deep and broad background knowledge and comprehension skills that are finely honed to evaluate and integrate information. A fast reading rate without higher order comprehension skills falls far short of the literacy standards needed for full participation.

The term *silent reading rate* is also a critical consideration in developing readers who can participate fully in the tasks of the digital-global age. Readers who stop and tediously sound out numerous words in texts are unlikely to have the cognitive resources to employ higher level comprehension processes. They are also individuals who will likely not have the stamina to read and integrate information from several sources or read extended texts.

Literacy researchers have shown an interest in two of the words within these terms—*comprehension* and *rate*. There has been substantial research on comprehension and comprehension processes (e.g., Duke & Pearson, 2002) and considerable work on rate. Almost all of this work, however, has been done on oral reading rate (e.g., Fuchs, Fuchs, Hosp, & Jenkins, 2001; Kame'enui & Simmons, 2001). Rarely, however, have the two constructs been examined in the same study. In particular, attention on the rates at which students are reading with meaningful comprehension has been scant.

When the topic of silent reading rates is raised among literacy researchers, the general response is one of skepticism (e.g., Brozo & Johns, 1986) or disinterest (see, e.g., Cassidy & Cassidy, 2009). In our case, especially for the two of us who have been teachers or teacher educators in U.S. contexts since the early 1970s, we know that this describes our perspective. As teachers and graduate students, we watched with skepticism the claims of and the techniques on speed reading (e.g., Frank, 1992). Continued spurious claims of speed-reading programs, such as that of reading 25,000 words a minute, have only reinforced a sense of skepticism for a new generation of researchers. As a result, the study of

rate, with respect to silent reading at least, has not been a popular topic for research.

Although there are several sets of oral reading norms (e.g., AIMSweb, 2008; Good & Kaminski, 1996; Hasbrouck & Tindal, 2006), there is a single set of silent reading norms that are based on data gathered in the late 1950s and reported in 1960 (Taylor, Frankenpohl, & Pettee, 1960). These silent reading norms are presented in Table 3.1. This set, although based on a large sample, is for the 50th percentile. How the 25th or 75th percentile groups do in comparison is uncertain. Such generic norms stand in contrast to the oral reading norms like those of Hasbrouck and Tindal (2006) that are also included in Table 3.1. As is the case with the various oral reading norms that have proliferated over the past 20 years in the wake of the advent of curriculum-based measurement (CBM; Deno, 1985), these oral reading norms are not based on assessments that include comprehension. Although dated and not as detailed as the Hasbrouck and Tindal (2006) oral reading norms, the silent reading norms (Taylor et al., 1960) are based on comprehension. This distinction is an important one, and it served as a primary incentive for our interest in CBSRR rather than simply on silent reading rate.

Table 3.1. Silent Reading and Oral Reading Rates

	Percentile	Grade												
		1	2	3	4	5	6	7	8	9	10	11	12	College
Silent reading rates (Taylor et al. (1960))	50th	80	115	138	158	173	185	195	204	214	224	237	250	280
	25th	23	65	87	92	100	122	123	124	NA				
Oral reading rates (Hasbrouck & Tindal, 2006)	50th	54	94	114	118	128	150	150	151	NA				
	75th	82	117	137	153	168	177	177	177	NA				

How Do Students of Different Quartiles Vary in Their CBSRR?

Although the Taylor et al. (1960) comprehension-based silent reading norms do not give an indication of the variation across a cohort of students, all available evidence leads to the expectation that differences across students within a cohort would be great. On the National Assessment of Educational Progress (NAEP; Lee, Grigg, & Donahue, 2007), the differences within a cohort of students in their comprehension

performances on a silent reading test are substantial.

There is evidence that rate figures into these performance differences on the NAEP silent reading assessments, insofar as the evidence comes from oral reading assessments. In a special study associated with the NAEP, researchers had a representative sample of students read orally the texts on which their silent reading comprehension had been assessed (Pinnell et al., 1995). Oral reading rate correlated moderately well with comprehension. Differences in students' word recognition accuracy were not statistically significant. Differences in students' oral reading rates were substantially different, with students who comprehended less well having much slower oral reading rates than students whose comprehension was higher. Similar patterns were found in a recent replication of the Pinnell et al. study (Daane, Campbell, Grigg, Goodman, & Oranje, 2005).

Table 3.1 includes the rate of growth that occurs in words per minute (wpm) in oral reading for students at three percentile levels across first through eighth grades, according to the Hasbrouck and Tindal (2006) norms. What is remarkable is the degree of consistency across the percentile groups once students move beyond first grade. They start at different points in first grade, but their growth occurs at the same pace after this point. Once students get to the middle grades, they level off. This rate of oral reading—150 wpm—is the same as the typical speech production rate of adults in the United States (Schmidt & Flege, 1995). The students in the 75th percentile have attained a level slightly higher than this rate, but the 50th percentile is on target in terms of speech production speed. The 25th percentile, at least through eighth grade, performs approximately 25 words slower than the average speech production rate.

In considering the potential patterns of CBSRR for readers at different levels, it is critical to recognize the differences between oral and silent reading. Oral reading is a performance-based situation. If a word is unknown, students cannot gloss over it in the manner that is possible in silent reading. Further, oral reading speed is governed by the speed with which individuals talk. Humans can speak faster than 150 wpm, and students can likewise read faster orally, especially if there is no concern with prosody or comprehension. These higher than expected rates may be the case as a result of the assessment expectations and practices of the past decade. Typically, as the norms in Table 3.1 indicate, proficient oral reading keeps pace with the rate at which human beings speak.

The oral production factor and the need to produce each word when reading orally, especially to a teacher or evaluator, leads to the suggestion that there may be more similarities among individuals in oral

reading than in silent reading. Silent reading contexts, however, also have constraints. There are limits to what the brain can do (Cunningham, Stanovich, & Wilson, 1990) and what the eye can do (see Chapter 2², this volume). Claims that someone can take a mental photograph of a page of text at 25,000 words a minute do not require extensive investigation to be deemed as spurious (McNamara, 2000).

What is clear from the data in Table 3.1 is that, not long into the reading acquisition process, silent reading rates surpass oral reading rates. The comparison of students at the 50th percentile in oral and silent reading attest to this conclusion, even at first grade. By fourth grade, silent reading for 50th percentile students is approximately one third faster than it is for oral reading. Further, once oral reading rates stabilize (reflecting the oral production factor) at the end of elementary/ middle school, silent reading rates continue to increase. By the time they are in college, readers at the 50th percentile read silently at almost twice the rate that they read orally.

With a greater range in reading rates, as is the case with silent reading, there may be greater variability among students of different proficiency levels. One factor that has sometimes created problems in the measurement of silent reading is the tendency for struggling readers to inflate their self-reports of reading rates (Fuchs et al., 2001). By making comprehension performances the ultimate criterion for determining appropriate rates, we are eliminating the potential of “fake” reading (Griffith & Rasinski, 2004).

How Well Do Students Sustain Their CBSRR Across an Extended Text?

We are especially interested in a construct called “reading stamina”—the ability to sustain attention and proficiency across a text. Even though educators refer to stamina as a critical aspect of reading (e.g., Johnson, Freedman, & Thomas, 2008; Qualifications and Curriculum Authority, 2005), it is rarely addressed directly in research. For example, in reviewing the three volumes of the *Handbook of Reading Research*, we found no references to or descriptions of stamina. Despite this lack of attention, a strong case can be made for hypothesizing that stamina could be an issue in both oral and silent reading. Students, particularly those in the bottom quartile, may quickly become fatigued when asked to read longer texts. Conversely, it could be argued that once students become familiar with the content and the vocabulary of an extended text,

2 References to chapters, this volume are referring to materials in *Revisiting Silent Reading: New Directions for Teachers and Researchers*.

their reading rates would increase. Texts are frequently written so that the principal ideas-and the vocabulary that represents those ideas-are presented early in a text. Once students have been introduced to a text's vocabulary and principal ideas, their reading rates might increase as they move through the remainder of the text.

Another perspective is that stamina would be challenged most directly in silent reading. Silent reading involves managing one's strategies and comprehension. A strategy that illustrates such comprehension management is clarifying confusing parts of text, one of a handful of strategies that has been found to distinguish proficient and challenged readers (Brown & Smiley, 1978). Thus, slow silent reading may be an indication of comprehension monitoring. Evidence for this hypothesis is limited. There is a need to find out more about silent reading rates, especially those of students in different proficiency groups. Rather than glossing over silent reading, interventions may need to focus directly on the nature of dysfluent silent reading patterns of low-performing students.

Stamina may be a particularly critical construct to consider in relation to the "iGeneration" (Rosen, 2010). For these students, whose lives have involved a barrage of information presented in several modalities simultaneously, attending to the fine print in rather solitary situations may be challenging. These students may have high levels of word recognition and may be facile with a variety of background knowledge. What may be challenging for them is sustained involvement with a text. The average length of a text on the fourth-grade NAEP is 800 words (Lee et al., 2007), while the average length of texts in the fourth-grade anthology of a widely used core reading program is approximately 2,000 words (Afflerbach et al., 2007).

A particular shortcoming of assessments that have typified the CBM movement, whether the mode is oral or silent reading, is the brevity of assessments—one minute or two minutes at most. The oral reading norms summarized in Table 3.1 reflect the shorter tasks. The silent reading norms, by contrast, reflect substantially longer tasks.

How Consistent Is the CBSRR of Students in a Digital Context Relative to a Paper-and-Pencil Context?

Teachers' interest in answers to this question derive from the recognition that reading in digital contexts is central to success in the digital-global age. Reading in digital contexts involves a myriad of issues that are not present in paper-and-pencil contexts (see Chapter

13, this volume). Even elementary students need to make numerous choices as they negotiate online reading tasks. In the face of a paucity of information on students' comprehension and rate of reading, our interest was straightforward: We wanted to know if students were able to read with similar levels of comprehension and at similar rates when they were reading texts presented digitally and in conventional contexts with printed texts.

Students' ability to transfer their reading skills to a new and critical context was one reason for including this component in our study. As researchers, we had a second reason. If teachers are going to support students' stamina and capture whether students are improving in their CBSRR, they need ways to gather information on students' CBSRR regularly and with *authentic* data. At the present, the typical form of assessment that is used for capturing CBSRR is the maze technique (Deno, 2003). The maze technique emanates from the CBM perspective that also spawned the widely used one-minute oral reading assessments (e.g., Good & Kaminski, 1996). A maze assessment for the primary grades consists of a passage slightly longer than what is anticipated would be read by the fastest grade-level readers (e.g., 300 words for second grade). Every seventh word (although the number can be varied) is replaced with a blank, and three or four words are listed underneath. The choices include the correct word as well as words that vary in their semantic, syntactic, or graphophonemic similarity to the target word. Students mark their choices. Their CBSRR is based on the number of words represented by their correct choices. As with oral reading fluency assessment, the typical length of time is one minute.

Studies have been conducted on the reliability of the maze relative to other assessments and have shown that the maze is positively related to performances on standardized tests (Shin, Deno, & Espin, 2000). Questions of validity have persisted around the maze, such as the effects of needing to stop and mark choices (Guthrie, Siefert, Burnham, & Caplan, 1974; Parker, Hasbrouck, & Tindal, 1992). Maze developers have identified particular rules for guessing, but the technique's success depends on carefully crafted alternatives for the target words.

The crafting of questions is a challenge for any assessment, but we are interested in the use of comprehension texts and questions that are typical of those used in classroom experiences, including typical tests. The tests that currently form such a central part of the classroom lives of students and teachers often contain highly crafted questions. Unfortunately, information from such tests is reported as summary scores,

usually in the form of norms. If data on CBSRR are to be brought to bear on instruction, teachers and students require information about specific texts and questions. They also require this information quickly to make informed instructional decisions—in hours rather than in the weeks or even months it can take to get back test results.

Because recent advances in digital environments have been notable (PytlikZillig, Bodvarsson, & Bruning, 2005), we believe that new technologies offer a viable approach to the problem of assessing CBSRR. In particular, the interactivity of the computer “page” could permit educators to measure students’ CBSRR reliably, frequently, and with authentic texts and tasks. A question that remained unanswered was whether students would perform with similar rates and comprehension when reading text on a computer screen and in the more typical school contexts of a printed text.

Designing and Implementing a Project to Answer Questions About CBSRR

In the study that we designed to address our questions about CBSRR, we had students representing a range of reading proficiencies read silently sections of an extended text in two different reading contexts. Our interest lay in similarities or differences in the performances of students of different quartile groups, at different points in reading an extended text, and between two contexts (digital and paper and pencil).

Method

Eighty-three students from five fourth-grade classrooms in a Midwestern, urban school district participated in the study. The participants were 65% Caucasian, 13% African American, 12% Asian American, and 9% Hispanic. More than 60% of the students in the schools receive free-or reduced-cost lunch. Participants included 15% English Learners and 13% special education students (i.e., those with speech-language disorders or specific learning disabilities).

We wrote two comparable sets of informational texts, each containing 1,000 words. Each set consisted of five texts connected by a common theme. The content of both themes came from a similar domain—communication. The underlying theme of one set of texts had to do with the role of posters in the past and present (e.g., posters as a source of information and announcements before the printing press). The theme of the second set was on nonverbal language (e.g., military hand signals,

Braille).

Texts were created over numerous iterations to ensure that the two sets were as comparable as possible on several measures. The first was sentence length. As the readability levels for the Flesch-Kincaid and Fry indicate in Table 3.2, texts were comparable on that dimension. A second consideration in the creation of the texts was the comparability of vocabulary. Data on the distribution of words in word zones established by frequency of appearance in written English (Hiebert, 2005) indicate that the distribution of words that were highly frequent (i.e., Word Zones 0-2), moderately frequent (Word Zones 3-4), and rare (Word Zones 5-6) was comparable across the two sets of texts.

The readability levels on both the Flesch-Kincaid and Fry suggest that the texts were approximately 1.5-2.5 grade levels above the mid-fourth-grade (the grade-level placement of students in the study). This difficulty level, however, is an artifact of a feature of readability formulas that has long been recognized as inflating the difficulty of informational texts (Cohen & Steinberg, 1983). This feature is that each appearance of a word counts in the establishment of readability with formulas such as the Flesch-Kincaid or Fry. In informational texts, rare (and often multisyllabic words) are repeated frequently when they are central to the content. Thus, informational texts typically are assigned high readability levels.

Table 3.2. Features of Texts Used in Study

Feature	Text A (Posters)	Text B (Nonverbal Language)
Number of words	1,000	1,000
Flesch-Kincaid readability	6.1	5.9
Fry readability	7	7
Unique words:		
Word Zones 0-2	85%	83%
Word Zones 3-4	13%	16%
Word Zones 5-6	1.5%	1%
Type-token ratio	0.28	0.28

The texts in this study had been written to be representative of informational texts and to comply with components of the TExT model (Hiebert, 2002) in which cognitive load (i.e., the ratio of unique words to total words or type-token ratio) and the percentage of rare words (i.e., Word Zones 5-6) are seen to influence text difficulty. The texts, as can be seen in Table 3.2, had type-token ratios of 0.28. A typical assessment text, such as those on the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 1996) has a type-token ratio of 0.50 or higher

(Hiebert, Stewart, & Uzicanin, 2010). Further, the percentages of rare words were low (1-1.5%) and the percentages of words in the 1,000 most frequent words (i.e., Word Zones 0-2) of 83-85% were high, leading to the expectation that most fourth graders should be able to read the majority of words.

To accompany the two text sets, we created two short sample passages of 200 words, each on familiar informational subjects: U.S. parks and dinosaurs. Each sample passage had two multiple-choice comprehension questions. As with the main text sets, the vocabulary in the sample passages was controlled. The purpose of the sample passages was to familiarize the participants with the assessment's format.

Each passage within a theme was immediately followed by four comprehension questions specific to the passage that students needed to answer before continuing to the next passage. Each set of passages, therefore, included 20 questions. Each set of questions for a passage included two literal questions, one inferential, and one interpretive.

We conducted a pilot study to ensure the validity and reliability of the comprehension questions and to ensure that the special Internet-based application that had been created for the computer condition of the study was student friendly. The pilot study sample consisted of two fourth-grade classes with demographics similar to those in the main study. One class of students ($n = 19$) was administered the full texts with comprehension questions in the computer context. A second class ($n = 21$) responded to the questions about the texts without exposure to the texts. The data from the pilot study was used to refine both the computer program and the comprehension questions. For example, questions that students in the latter group could answer with high levels of success were eliminated from the final test set.

Students were assessed in spring of fourth grade. Computer administration was conducted in the school's computer lab with two observers who read directions, assisted with technical problems, and redirected students. The individualized paper-and-pencil administration followed the same format and organization but added a third observer who aided in recording students' start and stop times for text sections.

Texts were counterbalanced for order of administration (i.e., computer vs. paper-and-pencil) and topic (i.e., nonverbal language vs. posters). Comprehension scores were corrected for guessing. Reliability of the 20-item comprehension items for each set of passages was established using coefficient α . The reliability for both scales was 0.74, an acceptable

range for research measures.

Results

Outlier analysis showed that there was a group of students with extremely high reading rates and very low comprehension performances. The performances of the outlier students can be seen in Figure 3.1. The observers who had been present during the task administration to ensure students' ease with the computer interface confirmed that particular students appeared to move rapidly through the task. As a result of this analysis, the data used in the subsequent analyses was limited to 65 students.

Descriptive statistics that appear in Table 3.3 indicate that silent reading rates were precisely the same on the two different sets of passages. This silent reading rate of approximately 154 wpm is similar to the average of 158 wpm reported by Taylor et al. (1960) for fourth graders almost 50 years ago. Comprehension performances were slightly lower on the posters text than that on nonverbal language.

A repeated-measures ANOVA was used to compare performances in the paper-and-pencil and computer administrations. For reading comprehension, there were no significant differences: $F(1, 77) = 1.19, p = 0.28$ $MSE = 6.32$. For silent reading rate, there was a significant effect for mode of presentation $F(1, 61) = 5.43, p = 0.02$ $MSE = 873$. This difference was not massive, but the context in which the slightly faster rate occurred is of interest—the computer context as is evident in Figure 3.1. Further, the lack of significant differences in comprehension indicates that this somewhat higher rate did not compromise comprehension.

The next set of analyses considered differences across quartile groups. Quartile groups were established on the basis of comprehension scores. Repeated-measures ANOVA revealed that rates for different comprehension quartiles were significantly different overall $F(3, 72) = 2.7, p = 0.05$ $MSE = 210035$.

The interpretation of rates by different groups is difficult because of different patterns of performance by the quartile groups on different parts of the texts. These patterns are provided for the first text (Posters) in Figure 3.2. For the first section of the assessment, the highest quartile performed approximately 30 wpm faster than the other three quartiles. The rates of Quartiles 1 and 2 were slightly lower than those of Quartile 3 but not substantially so on the first section of the text.

Figure 3.1: Average reading rate by group and context

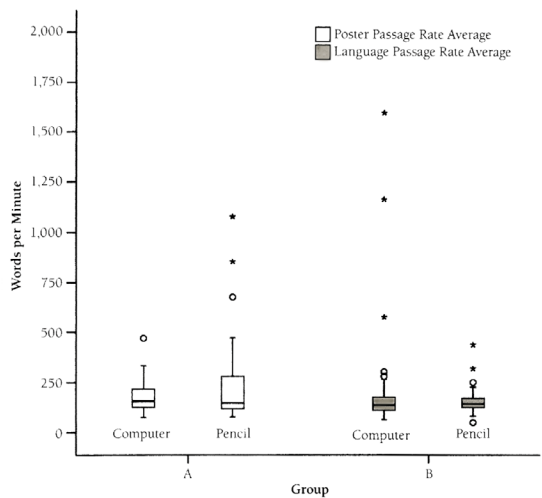
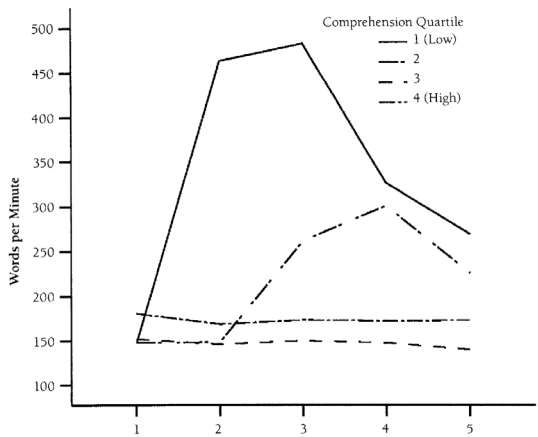


Table 3.3. Descriptive Statistics for Comprehension and Silent Reading Rate for Texts

	Mean	SD
Corrected comprehension score Text A (posters)	6.3	4.1
Corrected comprehension score Text B (nonverbal language)	7.9	3.7
Silent reading rate Text A	153.5	63
Silent reading rate Text B	153.5	60

Figure 3.2: Silent reading rate for text A (posters) by section



A repeated-measures ANOVA verified the pattern that can be seen in Figure 3.2 of performances of different quartile groups across sections of the text. Although students in the two lower quartiles started out at a reasonable rate, their rates changed dramatically over the sections of the assessment (but not with in creases in comprehension). The effect was nonlinear. The lowest quartile readers increased their speed after one passage (but without commensurate gains in comprehension). The second-lowest quartile increased their speed after two sections (again, without commensurate gains in comprehension). The students in the top two quartiles had stable rates that changed very little across sections of the text. Further, their comprehension remained stable.

Some Conclusions About Silent Reading

Silent reading has been an area in which educational practices have swung from one extreme to another (see Chapter 1, this volume). At particular times, all reading—even for first graders—was mandated or advocated to be silent. The opposite swing of the pendulum has been evident in the past decade, when oral reading has been emphasized as the primary mode. When one solution is found wanting, it is replaced by another solution. In a domain as complex as reading, single solutions will always be found wanting. A single study on CBSRR cannot produce all of the answers to a very complex set of issues. We can, however, give some tentative answers to a critical set of questions. These answers are offered in the spirit of continuing investigation, both by researchers and teachers, of what works best with particular kinds of texts and at particular points in development.

We begin by answering the question that we raised in the title of this chapter—are students really reading in independent reading contexts? The answer: Yes, most students are. Many students read at fairly consistent rates across different sections of a text. They comprehend at a fairly consistent level as well. Their rate is somewhat faster when they are reading digital text rather than a paper text, but with similar levels of comprehension.

This pattern—where most students are reading consistently in different silent reading contexts—is an important one to consider when thinking about the design of instruction. We are in the midst of the greatest knowledge revolution in human history. In a world where knowledge is the critical commodity, reading is a primary means whereby knowledge is acquired. We are not suggesting by any stretch of the

imagination that all reading should be silent reading (see the Conclusion for an expansion on the functions of oral and silent reading). Oral reading serves several essential roles, particularly at critical periods in students' reading acquisition. By the same token, to limit silent reading opportunities of all students because a portion of a cohort struggles with the task does a great disservice to all students. For struggling readers, such prohibitions mean that there is no opportunity to develop capacity in silent reading. For proficient readers, opportunities to learn are constrained when silent reading is limited.

Consider the greater amount of new vocabulary that students can acquire through silent rather than oral reading. If fourth graders read orally for 30 minutes daily at a speed of 118 wpm, they will read approximately 3,540 words daily or 637,200 words over a school year of 180 days. If they spend the same length of time reading silently, they will read 4,590 words daily or 826,200 words over the school year—approximately 189,000 more words. Based on existing research, it is estimated that 2-5% of these words will be unknown to students (Stahl, 1999) and, of these unknown words, students can be expected to remember approximately 5-10% from a single reading (Nagy, Anderson, & Herman, 1987). Using estimates of 3.75% unknown words and 7.5% remembered words, students will learn approximately 532 additional words in silent reading contexts. In that it is estimated that fourth graders acquire approximately 2,000 new words a year (Graves, 2006), this amount is significant. Further, because a primary way in which oral reading occurs is through round robin reading (Brenner et al., 2009), it is not at all clear that students will be attending to the texts to the same degree during oral reading as in silent reading.

But not all students' performances are consistent and reliable in silent reading contexts. Approximately 20% of the students did not stay "on the page." Another group of students read the first one or two texts conscientiously but changed their strategy at that point, moving quickly to answer the comprehension questions without careful reading of the text. Considerable attention is required on the kind of experiences that underlie consistency in silent reading, particularly the stamina that is required to sustain interest and monitor comprehension through extended texts. We hypothesize that stamina is part of the cycle of poor reading that Stanovich (1986) describes. As poor readers read less, their skills become increasingly inadequate for new developmental tasks such as reading chapter-long texts. Even if the texts are not overly difficult (which was the case with the texts in the present study), poor readers approach reading tasks with low levels of motivation and interest. As Swan, Coddington, and Guthrie (see

Chapter 6, this volume) describe, these students have poor identities of themselves as readers and low levels of intrinsic motivation.

Effective silent reading habits are not automatic outcomes of proficient word recognition and oral reading fluency. There are aspects of silent reading that make it unique from oral reading: vocalization, the need for self-monitoring, stamina, and interest. Numerous chapters in this volume highlight the components of instruction that support these components of effective silent reading. We will not review all of these components, but we do underscore one point: Just as the development of poor reading habits occurs over an extended period of time, so too development of good reading habits likely reflects many experiences over an extended period of time.

For the students who engage in what Griffith and Rasinski (2004) have described as “fake reading” behaviors, efforts to develop proficiencies such as self-monitoring, stamina, and interest are interwoven with the need to develop students’ identities as readers and their intrinsic motivation. Most students have acquired fundamental word recognition by the end of second grade (Hiebert et al., 2010) and definitely by the middle of fourth grade (Pinnell et al., 1995). For a significant portion of these students (approximately a third of a grade cohort), this recognition is tedious and time consuming. They have not developed perseverance or stamina for the task. They need considerable support if they are to sustain attention to the texts and tasks of daily classroom life.

There are likely limits to what teachers can do—especially in classrooms where large groups of students have such behaviors. Hiebert, Menon, Martin, and Bach (2009), in considering the research on silent reading, suggest that digital contexts may be one means whereby support can be provided for struggling readers. In a computer context, the text can be fine tuned. The length of time can be monitored. Content can be chunked and periodic check-ins can be made. The architecture can be designed so that the length of time, the accessibility of text, and the tasks can be carefully adjusted to students’ growing capacity as readers. Not much data have been gathered on current efforts, especially for struggling readers, but there is suggestive evidence that digital technology may provide the scaffolding that supports struggling readers in becoming stronger readers (Moran, Ferdig, Pearson, Wardrop, & Blomeyer, 2008).

At least in terms of our interest in providing classroom teachers with authentic and reliable assessments, the findings of this study leave us optimistic that digital contexts can serve as a means for providing teachers and students with consistent and usable information. Students responded

well to the digital context with over-all reading rates higher in that context than in the paper-and-pencil context. What we found to be particularly encouraging about this result is that students' faster rates did not compromise comprehension. This finding of students' somewhat superior performances in the digital context also bodes well for their flexibility as readers and their adaptation to a context that will be a critical one in their futures.

The study that we report in this chapter offers a window on variations of silent reading rate and comprehension of fourth-grade students when they are asked to read informational text. There are numerous questions that remain: How does this relationship change when similar assessments are administered to students in other elementary grades? Will the rates level off, as has been observed with oral reading fluency as the grades increase? Will reading rates change when comparing matched narrative and informational texts? When is it possible to gather reliable data based on students' developmental reading patterns? How should meaningful benchmark reading rates across the grades be created that are related to comprehension performance? Are students reading at appropriate rates? Are there optimal silent reading rates? Does oral reading practice improve CBSRR? Although this list of unanswered questions is sizable, it is not exhaustive. It illuminates the need for much more work in the area of silent reading assessment. Educators at all levels would benefit from a more nuanced understanding of the factors that affect students' learning when reading silently. Greater understanding of this little-studied reading mode will help to inform the instructional choices teachers make as students progress across the grades.

References

- Afflerbach, P., Blachowicz, C.L.Z., Boyd, C.D., Cheyney, W., Juel, C., Kame'enui, E.J., et al. (2007). *Scott Foresman Reading Street: Grade 4* (4th ed.) [Student edition]. Glenview, IL: Scott Foresman.
- AIMSweb Assessment and Data Management for RTL (2008). Establishing curriculum-based measurement oral reading fluency performance standards to predict success on local and state tests of reading achievement. Retrieved June 22, 2010, from www.aimsweb.com/uploads/news/idl9/orf_benchmarks.pdf/
- Brenner, D., & Hiebert, E.H. (2010). If I follow the teachers' editions, isn't that enough? Analyzing reading volume in six core reading programs. *The Elementary School Journal*, 110(3), 347-363. doi:10.1086/648982
- Brenner, D., Hiebert, E.H., & Tompkins, R. (2009). How much and what are third graders reading? Reading in core programs. In E.H. Hiebert (Ed.), *Reading more, reading better* (pp. 118-140). New York: Guilford.
- Brown, A.L., & Smiley, S.S. (1978). The development of strategies for studying texts. *Child Development*, 49(4), 1076-1088.
- Brozo, W.G., & Johns, J.L. (1986). A content and critical analysis of 40 speed reading books. *Journal of Reading*, 30(3), 242-247.
- Carver, R.P. (1990). *Reading rate: A review of research and theory*. San Diego: Academic.
- Carver, R.P. (1992). Reading rate: Theory, research, and practical implications. *Journal of Reading*, 36(2), 84-95.
- Cassidy, J., & Cassidy, D. (2009). What's hot for 2009: National Reading Panel influence wanes in 13th annual survey. *Reading Today*, 26(4), 1, 8-9.
- Cohen, S.A., & Steinberg, J.E. (1983). Effects of three types of vocabulary on readability of intermediate grade science textbooks: An application of Finn's transfer feature theory. *Reading Research Quarterly*, 19(1), 86-101. doi:10.2307/747339
- Cunningham, A.E., Stanovich, K.E., & Wilson, M.R. (1990). Cognitive variation in adult college students differing in reading ability. In T.H. Carr & B.A. Levy (Eds.), *Reading and its development: Component skills approaches* (pp. 129-159). San Diego, CA: Academic.
- Daane, M.C., Campbell, J.R., Grigg, W.S., Goodman, M.J., & Oranje, A. (2005). *Fourth grade students reading aloud: NAEP 2002 special study of oral reading* (NCES 2006-469). Washington, DC: Institute of Education Sciences, U.S. Department of Education.
- Deno, S.L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52(3), 219-232.
- Deno, S.L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education*, 37(3), 184-192. doi:10.1177/00224669030370030801
- Duke, N.K., & Pearson, P.D. (2002). Effective practices for developing reading comprehension. In A.E. Farstrup & S.J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 205-242). Newark, DE: International Reading Association.
- Frank, S.D. (1992). *Remember everything you read: The Evelyn Wood 7-day speed reading and learning program*. New York: Avon.
- Fuchs, L.S., Fuchs, D., Hosp, M.K., & Jenkins, J.R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, 5(3), 239-256. doi:10.1207/S1532799XSSR0503_3
- Good, R.H., & Kaminski, R.A. (1996). *DIBELS: Dynamic Indicators of Basic Literacy Skills*. Longmont, CO: Sopris West.
- Graves, M.F. (2006). *The vocabulary*

book: *Learning and instruction*. New York: Teachers College Press; Newark, DE: International Reading Association; Urbana, IL: National Council of Teachers of English.

Griffith, L.W., & Rasinski, T.V. (2004). A focus on fluency: How one teacher incorporated fluency with her reading curriculum. *The Reading Teacher*, 58(2), 126-137. doi:10.1598/ RT.58.2.1

Guthrie, J.T., Siefert, M., Burnham, N.A., & Caplan, R.I. (1974). The maze technique to assess, monitor reading comprehension. *The Reading Teacher*, 28(2), 161-168.

Hasbrouck, J., & Tindal, G.A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher*, 59(7), 636-644. doi:10.1598/RT.59.7.3

Hiebert, E.H. (2002). Standards, assessments, and text difficulty. In A.E. Farstrup & S.J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 337-369). Newark, DE: International Reading Association.

Hiebert, E.H. (2005). In pursuit of an effective, efficient vocabulary curriculum for the elementary grades. In E.H. Hiebert & M.L. Kamil (Eds.), *The teaching and learning of vocabulary: Bringing scientific research to practice* (pp. 243- 263). Mahwah, NJ: Erlbaum.

Hiebert, E.H., Menon, S., Martin, L.A., & Bach, K.E. (2009). *Online scaffolds that support adolescents' comprehension*. Seattle, WA: Apex Learning.

Hiebert, E.H., Stewart, J., & Uzicanin, M. (2010, July). *A comparison of word features affecting word recognition of at-risk beginning readers and their peers*. Paper presented at the annual conference of the Society for the Scientific Study of Reading, Berlin, Germany.

Johnson, H., Freedman, L., & Thomas, K.F. (2008). *Building reading confidence in adolescents: Key elements that enhance proficiency*. Thousand Oaks, CA: Corwin.

Kame'enui, E.J., & Simmons, D.C. (2001). Introduction to this special issue: The

DNA of reading fluency. *Scientific Studies of Reading*, 5(3), 203-210.

Lee, J., Grigg, W.S., & Donahue, P.L. (2007). *The nation's report card: Reading 2007* (NCES 2007- 496). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

McNamara, D.S. (2000). *Preliminary analysis of photoreading* (ODURF File No. 193021). Moffett Field, CA: NASA Ames Research Center. Retrieved April 27, 2010, from ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20000011599_2000009345.pdf

Moran, J., Ferdig, R.E., Pearson, P.D., Wardrop, J., & Blomeyer, R.L., Jr. (2008). Technology and reading performance in the middle-school grades: A meta-analysis with recommendations for policy and practice. *Journal of Literacy Research*, 40(1), 6-58. doi:10.1080/10862960802070483

Nagy, W.E., Anderson, R.C., & Herman, P.A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, 24(2), 237-270.

National Institute of Child Health and Human Development. (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. 00-4769). Washington, DC: U.S. Government Printing Office.

Parker, R., Hasbrouck, J.E., & Tindal, G. (1992). The maze as a classroom-based reading measure: Construction methods, reliability, and validity. *The Journal of Special Education*, 26(2), 195-218. doi:10.1177/002246699202600205

Pinnell, G.S., Pikulski, J.J., Wixson, K.K., Campbell, J.R., Gough, P.P., & Beatty, A.S. (1995). *Listening to children read aloud: Data from NAEP's Integrated Reading Performance Record CTRPR at grade 4*. Washington, DC: National Center for Education Statistics, U.S. Department of Education.

PytlikZillig, L.M., Bodvarsson, M., & Bruning, R.H. (2005). *Technology-based education: Bringing researchers and practitioners together*. Greenwich, CT: Information Age.

Qualifications and Curriculum Authority. (2005). *English: 2004/05 annual report on curriculum and assessment* (QCA/05/2167). Coventry, England: Author. Retrieved April 27, 2010, from <http://www.ttrb.ac.uk/attachments/31f8ef70-0516-4a54-878e-c4f268ef85c2.pdf>

Rosen, L.D. (2010). *Rewired: Understanding the iGeneration and the way they learn*. New York: Palgrave Macmillan.

Schmidt, A.M., & Flege, J.E. (1995). Effects of speaking rate changes on native and nonnative speech production. *Phonetica*, 52(1), 41-54. doi:10.1159/000262028

Shin, J., Deno, S.L., & Espin, C. (2000). Technical adequacy of the maze task for curriculum-based measurement of reading growth. *The Journal of Special Education*, 34(3), 164-172. doi:10.1177/002246690003400305

Stahl, S.A. (1999). *Vocabulary development*. Cambridge, MA: Brookline.

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360-407. doi:10.1598/RRQ.21.4.1

Taylor, S.E., Frankenpohl, H., & Pettee, J.L. (1960). *Grade level norms for the components of the fundamental reading skills* (EDL Research and Information Bulletin No. 3). Huntington, NY: Educational Development Laboratories.

II. INSTRUCTION AND PRACTICES THAT FOSTER STAMINA AND SILENT READING

CHAPTER 4

Stretching Elementary Students in Complex Text: Why? How? When?

Heidi Anne E. Mesmer
Virginia Tech

Stretching students in text? What does that mean? Put them on a rack? A third-grade teacher mischievously made the comment at a recent professional development workshop. I had to bite my tongue because, in truth, I find the phrase a little odd myself. I know that I certainly never used the term “stretch text” when I thought about challenging students with reading materials before the Common Core State Standards for the English Language Arts in History/Social Studies, Science, and Technical Subjects (CCSS) were established (National Governors Association Center for Best Practices (NGA Center) & Council of Chief State School Officers (CCSSO), 2010a). Instead, like many other teachers, I might have spoken of an instructional-level text but never a stretch text.

So where did this term come from? What does scholarship say about how to stretch—or challenge—students in text? This chapter focuses on these very questions. The chapter will begin with a discussion of the meaning of complex text, both how the CCSS define it as well as how it is defined from other perspectives. The second section discusses what is meant by stretch text in elementary school and how the introduction of the stretch notion will influence reader–text matching paradigms. The brief third section presents a series of rationales, both good and bad, used to bolster arguments to stretch students in text. Finally, the last section is an extended discussion of the factors that may contribute to or inhibit students being stretched in text. Each section pays attention to the gaps in the literature and the type of information needed for students to reach the high aspirations that the CCSS introduce.

What Is Complex Text?

The CCSS (NGA Center & CCSSO, 2010a) are unprecedented in their focus on text. No other standards document in recent history has addressed text with greater attention, specificity, or energy than has the CCSS. In some powerful ways, the CCSS thrust text into the spotlight and

challenge teachers, publishers, and researchers to think more carefully about students' reading materials. According to the CCSS, text complexity is "The inherent difficulty of reading and comprehending a text combined with consideration of reader and task variables" (NGA Center & CCSSO, 2010b, p. 43). Thus, the CCSS use the term "complexity" interchangeably with "difficulty" (a point with which I differ later in this section).

In some respects, understanding the CCSS definition of complex text comes into focus better by reviewing the three-part assessment of text complexity articulated in Appendix A (NGA Center & CCSSO, 2010b). This model illustrates the elements of text and of the reader–text match that the CCSS conceptualize as making a text difficult. The tripartite model addresses qualitative tools, reader and task variables, and quantitative tools that capture the complexity of a text to an individual student.

Through qualitative means, a discernible and experienced human reader applies professional judgment to evaluate a text in order to estimate its complexity for target readers. According to the CCSS, the text features best evaluated using human judgment include:

- Levels of meaning in literary texts and levels of purpose in informational texts
- Text structures (e.g., simple, well-marked structures vs. implicit and layered structures)
- Language conventionality and clarity (e.g., literal, clear language vs. figurative, academic, or domain-specific vocabulary)
- Knowledge demands (e.g., level of knowledge assumed by the text)

The qualitative leg of the CCSS tripod, while theoretically interesting, has not been reliably established by research.

The second leg of text complexity relates to reader and task factors, elements generally not inherent to the text itself. (From my perspective, these are part of the reader–text match but not really an assessment of text complexity.) Appearing to draw from the reader–text–task model found in the RAND report, the CCSS remind the field that reader variables, such as motivation, background, knowledge, and experiences, will all render a text more or less difficult to a group of readers (RAND Reading Study Group, 2002). Additionally, the CCSS address task variables, including purpose, assignment requirements, and teacher levels of expectation, reminding the reader that the analysis of text complexity as it relates to reader and task is best done by teachers.

The third leg of the text complexity assessment is the one with the most validation and reliability and the longest history, as it uses the quantitative systems of readability formulas (Harrison, 1980; Mesmer,

2008). These traditional formulas (e.g., Dale-Chall, Flesch-Kincaid) and their second-generation digitally calculated cousins (e.g., Lexiles, ATOS, degrees of reading power) are theoretically the same. Both estimate difficulty by using a word factor, usually an estimate of word frequency, and a syntactic factor, usually the length of sentences. Labels such as grades, Lexiles, or degrees of reading power are generated for texts and used to estimate difficulty. The CCSS also identify the Coh-Metrix tool, which measures text cohesion through a myriad of text features (e.g., anaphora, cross-sentence referents).

The quantitative guidelines for requisite text complexity across six grade-level bands are specified precisely in Appendix A of the CCSS (NGA Center & CCSSO, 2010b, p. 8). The bands’ range has been extended in a supplement to Appendix A (Nelson, Perfetti, Liben, & Liben, 2012). The staircase of text complexity, as illustrated in Table 4.1, moves from beginning reading to the college and career-readiness level (NGA Center & CCSSO, 2010b, p. 8). At each step or grade-level band, a precise range of text difficulty is prescribed in various readability formulas. Although the CCSS special study by Nelson et al. (2012) evaluated six readability systems to express text difficulty in grade levels, the Lexile Framework has figured prominently in text complexity determination. The framework is based on word frequency and sentence length, and it uses Lexiles (L) rather than grade levels as a unit of text difficulty. One Lexile is “1/1000th of the difference between the mean difficulty of mid-first grade material and the mean of difficulty of college and workplace passages” (Stenner, Burdick, Sanford, & Burdick, 2007, p. 6).

Table 4.1: The Common Core Staircase of Text Complexity¹

Common Core Band	ATOS	Degrees of Reading Power*	Flesch-Kincaid	The Lexile Framework*	Reading Maturity	SourceRater
2nd—3rd	2.75—5.14	42—54	1.98—5.34	420—820	3.53—6.13	0.05—2.48
4th—5th	4.97—7.03	52—60	4.51—7.73	740—1010	5.42—7.92	0.84—5.75
6th—8th	7.00—9.98	57—67	6.51—10.34	925—1185	7.04—9.57	4.11—10.66
9th—10th	9.67—12.01	62—72	8.32—12.12	1050—1335	8.41—10.81	9.02—13.93
11th—CCR	11.20—14.10	67—74	10.34—14.2	1185—1385	9.57—12.00	12.30—14.50

¹ From Coleman, D., & Pimentel, S. (2012). Supplemental Information for Appendix A of the Common Core State Standards for English Language Arts and Literacy: New Research on Text Complexity (p. 4). Washington, DC: National Governors Association & Council of Chief State School Officers. Retrieved from http://www.corestandards.org/assets/E0813_Appendix_A_New_Research_on_Text_Complexity.pdf

In the CCSS, kindergarten through first grade levels are not assigned a text difficulty range, but a default level is set by the entering value for the band for second to third grade. First-grade children must reach the minimal level at the bottom of that default entry by the end of first grade (420L level). Note that the levels of text complexity expected at various grades are somewhat accelerated. While schools would typically expect students at the end of the third-grade year to read at a fourth-grade level, the CCSS staircase sets that level at about fifth or sixth grade.

Lest anyone think the staircase is merely suggestive, the language in English Language Arts Standard 10 indicates otherwise. The phrasing within Standard 10 for any grade level indicates that the text ranges are not loose guidelines but concrete expectations. For instance, the language in the grade 3 informational text standard reads, “Comprehend information texts...at the high end of the 2-3 band independently and proficiently” (NGA Center & CCSSO, 2010a, p.12, emphasis mine). Thus, although the CCSS offer three ways to assess text complexity, the quantitative tools are the most specific and the most translatable into classrooms.

Treating the terms “text complexity” and “text difficulty” as interchangeable, however, as done in the CCSS, confuses causes with effects. Mesmer, Cunningham, and Hiebert (2012) distinguished between text complexity and text difficulty. Text complexity is simply the naturally occurring textual elements in a passage or book that can be analyzed, manipulated, or otherwise studied, and, as such, is an independent variable. On the other hand, text difficulty is not one-dimensional but a numeric expression of a relationship between text and readers, and it is not a feature intrinsic to the text. As Mesmer et al. stated, “The difficulty of a text or text feature always implies a dependent or criterion variable: the actual or predicted performance of multiple readers on a task based on that text or feature” (p. 236).

Text difficulty estimates, such as those created by readability formulas, connect the complexity of a text (e.g., word frequency and sentence length) to reader performance (i.e., readers’ comprehension of a text) or predicted performance (e.g., a formula’s estimate of difficulty or a teacher’s estimates of difficulty). Therefore, the estimate of text difficulty is only as good as the relationship upon which the estimate is based, and the complexity of a text is simply what is there. Therefore, if we are to “stretch” students in text, we must depend on the very best estimates of text difficulty that exist, and we must better understand the impact of various text complexity features on readers.

What Is Stretch Text?

The theme of challenging text is replete throughout the CCSS, but the term “stretch” appears only in Appendix A. In describing how students need opportunities to both stretch their reading abilities and engage in easy fluency reading, the CCSS writers stated, “Students deeply interested in a given topic, for example, may engage with texts on that subject across a range of complexity. Particular tasks may also require students to read harder texts than they would normally be required to” (NGA Center & CCCSO, 2010b, p. 9). Scholars such as Roskos and Neuman (2013) and Shanahan (2011) have endorsed the use of challenging texts that stretch students’ reading abilities.

As the text staircase indicates, the Common Core will infuse more challenging texts into U.S. classrooms. In fact, Common Core writers already have produced a document to guide educational publishers in the creation of materials (Coleman & Pimentel, 2012). This document clearly specifies that the complexity of texts should be aligned with the staircase in Appendix A (NGA Center & CCCSO, 2010b). Thus, the Common Core text parameters have clear and dramatic policy implications. As a country, we have witnessed the abrupt translation of policies like this by educational publishers in the past, such as the rapid swings from authentic or whole language texts to decodable texts during the final decade of the 20th century (Hiebert & Martin, in 2015). Rapid—and dramatic—changes in policy require a solid grounding in evidence. In the section that follows, I consider the evidence that underlies the perspective on challenging texts in the CCSS.

The emphasis on giving students challenging texts introduces a paradigm shift in reader–text matching that contrasts with decades-old emphases on the avoidance of frustration. The older paradigms focused on finding the “just right” text as measured at a specific point in time. The guidelines that most elementary teachers use for reader–text matching are the word accuracy and comprehension levels established by Betts (1946) in finding independent, instructional, and frustrational texts. These include

- independent: texts that students read without teacher support (word accuracy: 99–100%; comprehension: 90–100%);
- instructional: texts that students read with teacher support (word accuracy: 95–98%; comprehension: 75–80%); and
- frustrational: texts are inaccessible to students with or without support (word accuracy: $\leq$ 94%; comprehension: $\leq$ 74%).

Within these guidelines, if teachers were to call any texts “stretch

texts,” they would likely identify the instructional level text as such.

Inadvertently, the Betts’s labels and reader–text matching standards may have shaped the views of many text researchers and teachers; however, while these boundaries for text difficulty have become the essential guidance through the present day, many questions have been asked across the years about their empirical basis (Clay, 1985; Ehri, Dreyer, Flugman, & Gross, 2007; Ekwall, Solis, & Solis, 1973; Halladay, 2012; Morgan, Wilcox, & Eldredge, 2000; Pikulski & Shanahan, 1982; Stahl & Heubach, 2005). The intense concern about avoiding frustration may not have been balanced with the equally important message to encourage challenge and avoid stagnation. It is indeed possible to build capacity for readers to handle more difficult passages. Although the text complexity staircase introduces many valid concerns, the theme of the Standard 10, to embrace challenge, is a message long overdue.

Unfortunately, just as the reader–text standards of the previous decades lacked empirical basis, so also does the stretch paradigm. We simply do not have an empirically based paradigm for how to challenge students in texts. We do not know exactly how far students can be pushed before they break, reaching the point where reading becomes incomprehensible and cognitively, psychologically, or emotionally exasperating. We do not know which text features can be ramped up and which must only be gently accelerated. We do not know at which points students can be stretched developmentally and within which contexts. Of course, this all begs the question why the text complexity standard and surrounding verbiage were introduced to begin with. What exactly has happened to cause standards writers to be concerned about the levels of texts at which students are reading?

Why Stretch Text for Elementary Students?

A careful reading of the Common Core materials and analysis of the themes and messages coming from the NGA Center suggests that text difficulty or complexity was given focus in the standards overall and specifically in Standard 10 for three reasons. First, and perhaps most obviously, the United States has not been comparing well in literacy performance internationally in the last 10 years (Martin, Mullis, & Kennedy, 2007).

Second is what I call “the text complexity gap between high school and college.” The 2006 ACT report *Reading Between the Lines* indicated that the success of students in a college-level social science course (i.e.,

grade of B or better) was predicted by the difficulty of texts to which they were exposed in high school. In other words, the complexity of materials in high school was watered down and limited students' abilities to achieve in college. Before the introduction of the Common Core's text staircase, the difficulty of materials required at the end of high school (1215L) was much lower than the difficulty of materials required for college and career (1355L). In Appendix A (NGA Center & CCSSO, 2010b), the trend of dumbing down high school texts was further reinforced with a citation of two studies (Chall, 1977; Hayes, Wolfer, & Wolfe, 1996). Both of these studies confirmed the easing of text difficulty across secondary schools over several decades. But, as demonstrated by Gamson, Liu, and Eckert (2013) and Hiebert and Mesmer (2013), the elementary school texts have not increased in difficulty. Through grade three, texts appear to have gotten more difficult during the 50-year period from 1960 to 2010 identified by CCSS writers, and sixth-grade texts in the most recent decades appear to be as difficult, if not slightly more so, than those of earlier eras. Thus, there is some evidence to support the need for increases in text difficulty, but the evidence exists at the secondary level and not at the elementary levels, where text complexity has increased.

A third reason why CCSS writers believed that accelerated text levels are necessary stems from claims that typical texts are too easy for many students. Evidence underlying this view comes from Williamson (2006), who followed a cohort of more than 60,000 third graders, beginning in 1999, through their eighth-grade year. Using the North Carolina end-of-grade test in reading (measured in Lexiles), the study tracked the progress of students in Lexile levels and contrasted this progress with the levels of the typical textbooks. The results showed that the achievement of students was close to, and perhaps limited by, the difficulty of texts. The findings logically followed that, if text levels were increased, student levels of achievement might also increase. The findings do not indicate, however, the degree to which the student sample in North Carolina matched national samples. During this period, North Carolina was showing high levels of achievement on the National Assessment of Educational Progress (NAEP), and it is possible that this sample was unlike national populations.

In sum, the reasons provided for the CCSS move toward increasing the text levels for students and stretching students in challenging text is based on international literary performance progress comparisons, trends in secondary schools, and a study conducted with a sample of students in North Carolina. The empirical foundation for the CCSS text staircase,

especially at the elementary levels, may not be strong but the guidelines and recommendations regarding challenging texts at all grade levels promises to have important consequences for teachers and students.

How and When to Stretch Students

How, then, do we go about focusing on challenge in reading in elementary classrooms? When do we put the stretch concept into action? I begin with reviewing the basics of reader–text matching because if there is no system in place for this, then there is no basis upon which one might stretch a student. Stretching or challenging students in text must be based upon some starting point, and this is established through reader–text matching. The remaining factors relating to stretching students in complex text include text levels, text length, genre, and cohesion. In addition, I address the importance of research that gives focus to the program of text that students read across time longitudinally.

The Stretch Baseline: Reader–Text Matching

The first steps toward stretching students in text must include the basics of reader–text matching. When teachers challenge students in text, they should not arbitrarily ask the whole class to read a single designated stretch text. The reader–text matching process begins with knowing the students' reading levels and then having some estimate of the levels of the texts. For example, if the reading level is obtained in grade levels (e.g., 2.1, 2.2, 2.3) through the Standardized Test for the Assessment of Reading (STAR), then the estimate of text difficulty should be given in the same units.

I use the word estimate very purposefully. Until the student is actually reading the text, then both the reading level and the text difficulty are estimates. Text difficulty measures, in particular, only provide a basic approximation of text difficulty. Actual text difficulty emerges only when an actual reader is reading that text. Although knowing a student's reading level and a text's estimated difficulty level provides a good place to start, a baseline on which to stretch students has yet to be conclusively established.

Stretching Students Through Text Levels

Beyond the Betts (1946) criteria for word accuracy and comprehension, little is known about exactly how much above a student's instructional level a text may reach before it becomes frustrational text

(i.e., we do not know what the tipping point is). Several researchers have experimented with the degree to which word accuracy levels can dip below the accepted 90%. With a great deal of rereading and teacher support, younger students can get as much as 85% word accuracy and still show reasonable comprehension (Ehri et al., 2007; Stahl & Heubach, 2005).

Only a few studies have suggested the degree to which text levels can surpass students' instructional reading levels. Morgan et al. (2000) used a partner reading intervention in which low-performing second graders were randomly assigned to read at three different text levels: on-grade level (i.e., their instructional level), two grades above instructional level, and four grades above instructional level. Students read with partners for 15 minutes per day for 95 days, and results showed the most improvement for students reading two grades above instructional level. The reading levels of students before and after the intervention were not provided. Below-level students starting out at a pre-primer level are likely quite different from below-level students starting out at a first-grade reading level.

To more carefully consider the effects of differing text levels on student proficiency, Mesmer and Hiebert (2011) analyzed a large data set of 9,535 records of third-graders' text comprehension, giving attention to cases in which students were reading texts that were above or below their targeted instructional reading level. Essentially, the difficulty of each text read by each student was labeled relative to the student's instructional level (called "target" level). The target level was the point at which readers could comprehend 70% of the material. Thus, it was possible to identify cases in which the texts read were specific amounts above or below the student's target level. Because the Lexile Framework was used, we divided the records into the following categories:

- Easy texts (101L to 250L below level)
- On-level texts (100L below to 50L above level)
- Stretch texts (51L to 100L above level)
- Difficult texts (101L to 250L above level)

The stretch text levels were one standard deviation above the targeted on-level designation. We also separated students into two proficiency groups: those whose targeted levels were on or above grade level, and those whose targeted level was below the grade-level range. Below-level readers were defined as those reading below 450L because the CCSS define the range of text difficulty for the second-to-third-grade band as 450L to 790L (NGA Center & CCSSO, 2010a). Students reading at or above 450L were

designated as on-level readers.

Means for comprehension are provided in Table 4.2 by reader level (below vs. on or above level). There were main effects for the text difficulty and reader levels ($F(3, 9,531) = 207.34, p < .001$, $F(2, 9,532) = 10.55, p < .001$, respectively). The reader by text difficulty interaction also was significant ($F(3, 9,532) = 15.03, p < .001$). Pairwise comparisons were significant at the .001 level for all text and reader combinations except for the difficult texts. On average, all students achieved a 61% reading comprehension in stretch texts that averaged 76L above their target levels. Below-level readers comprehended at a lower level than did on-level readers at all text levels except in the difficult texts, where all readers comprehended at about 53%. Across reader levels, performance declined as text difficulty increased, with comprehension levels dipping below 70% in the stretch texts.

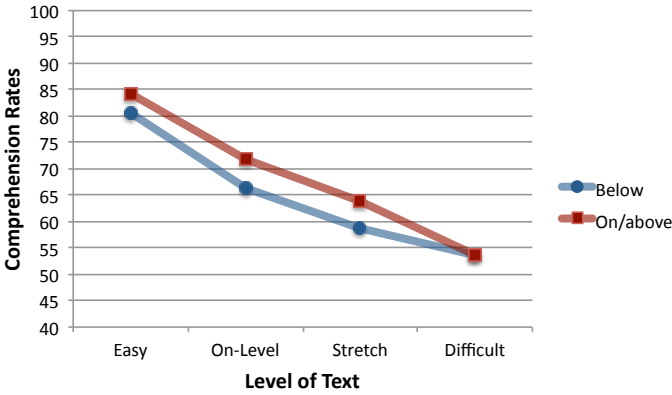
Table 4.2: Comprehension of Below-Level vs. On- or Above-Level Third-Grade Students on Texts of Different Levels

Text Levels (Lexile range relative to reader proficiency)	Reader Level	Comprehension: X (SD)
Easy (101L to 250L below)	Below	80.65 (14.83)
	On/above	84.23 (15.75)
On-Level (100L below to 50L above)	Below	66.25 (19.56)
	On/above	71.88 (17.12)
Stretch (51L to 100L above)	Below	58.76 (19.81)
	On/above	63.95 (17.35)
Difficult (101L to 250L above)	Below	53.63 (19.77)
	On/above	53.70 (18.01)

¹ From Mesmer and Hiebert (2011); used with permissions of the authors.

As illustrated in Figure 4.1, the patterns of response to text difficulty changes differed somewhat for below- and on-level readers. On- and below-level readers had nearly identical performance in difficult texts, which were about 200L above level, but unlike on-level students, below-level readers' performance across text difficulty categories reflected a curvilinear pattern. Their performance dipped more sharply beginning with the on-level texts and sloped more steeply than that of the on-level readers. In fact, performance for below-level readers did not reach the designated 70% for any texts except easy ones. This exploratory work suggested that no readers could be stretched to 200L above their targeted levels (about two grade levels). Because the levels of text were always relative to the student's own level, readers were not reading at the same text levels but were reading texts that were harder or easier for them.

Figure 4.1: Comprehension Rates for Below-Level vs. On- or Above-Level Third-Grade Students on Texts of Different Levels



What happened was that below-level readers, even in on-level texts, were still not performing well; therefore, in a sense, an on-level text was a stretch text for them. These preliminary results suggested that students’ reading proficiencies determined the upper limits of their performances. In particular, students who had been designated as performing below-grade level were not able to rise to the occasion to the same degree as did students designated as at-grade level or above.

Stretching Students Through Text Length

Although rarely mentioned in schemes about stretching, length definitely factors into text elements that make a text challenging. In fact, the Common Core does not mention text length other than to suggest to secondary teachers that short, dense texts are good exemplars for supporting students in close reading and answering text-based questions (NGA Center & CCCSO, 2010b). In the elementary grades, length is particularly important because it changes significantly throughout the grades. The average length of a passage, book, or text that students read in first grade is about 50–250 words. By fourth grade, the average length of a textbook passage, chapter, or worksheet passage is 2,000 words, at least an eightfold increase. Several authors have found text length (i.e., number of words) to predict text level in the Reading Recovery scheme, suggesting that length factors into challenge in the primary grades (Cunningham et al., 2005; Hatcher, 2000). Certainly as length shifts, so also do reading behaviors. Students in the primary grades orally read short blocks of text with supportive pictures in a matter of minutes; however, students in the intermediate grades must read extended texts silently without pictures for upwards of 20 minutes.

Few studies examine text length (Calfee & Hiebert, 2011; Hiebert, Wilson, & Trainin, 2010; Mesmer & Hiebert, 2013). Calfee and Hiebert (2011) found, for example, that length could be a variable that explained the different levels of achievement for California fourth graders on the NAEP and the California State Test (CST). The percentage of students who were proficient or higher on the CST was 38% higher than that of the NAEP results. Significant length differences characterize these two tests, with the NAEP passages being about 800–1,000 words and the CST passages running 350–400 words.

Hiebert et al. (2010) investigated how students of different proficiency levels performed at different points within a lengthy passage. Students in the two lower quartiles showed reasonable rates and levels of comprehension in the beginning portions of the passage but had significantly depreciated comprehension scores in the latter portions. The findings suggested that stamina effects were at play, as students tired during the lengthy passage.

Mesmer and Hiebert (2013) manipulated text length and difficulty to identify the degree to which these factors interacted and how students of differing proficiency levels (at level vs. below level) were affected by this combination. Three different sets of text passages were designed at three difficulty levels (400L, 600L, and 800L). Within each difficulty level, one passage was 200 words and one was 1,000 words. Topics were kept consistent across a difficulty level (e.g., schools and community helpers for 400L-level texts; budgets and money for 600L-level texts; and natural resources and oil for 800L-level texts). Using a within-subjects design, researchers required all students to read all passages, with comprehension being the outcome variable. As texts became more difficult and longer, comprehension decreased. At every difficulty level, students comprehended the short versions better than the long versions. The results suggested that length compounded the effects of difficulty, rendering texts of the same difficulty level harder.

As educators continue to explore how to stretch elementary students, attention to text length is warranted. Especially important will be understanding how to support students as they confront the length shifts at various developmental junctures. Clearly, one of these shifts is in the movement toward reading chapter books in late first grade or early second grade. Another shift comes in late third grade or early fourth grade, when lengthy expository texts begin to prevail. Of course, at this point, genre is also at play in presenting challenges to students. In a recent piece, Mesmer et al. (2012) asked, “How can research be designed to distinguish the

relative contributions to text length effects of reader fatigue and cumulative deficits in memory?” (p. 245). The key to understanding the impact of length is its influence on readers’ stamina and fatigue.

Stretching Students Through Text Genre

Although a great deal of debate exists in the field about exactly where to draw the genre lines, it appears that a convenient way to think about genre is to divide passages into narratives or expository pieces (Mesmer et al., 2012). What is known about genre is that expository texts are often dense, with new, unknown vocabulary, often the type of domain-specific Tier 3 words that represent complicated concepts or processes (Beck, McKeown, & Kucan, 2013; Fang, 2006). In narrative texts, students are introduced to new vocabulary, but frequently the words are Tier 2 words that enhance the meaning of the text, express degree, or modify the core of a sentence. Note the examples of texts in Table 4.3. The words in the excerpt from *Boy* are, in some cases, easily inferred, as they are compound words or ancillary to the passage. In the expository text sample, however, the concepts of *poaching* as either a process or an action and the noun *ivory* are essential to understanding the text. If one does not know the meaning of these words, getting the gist of the passage is not possible.

Table 4.3: Examples of Differing Vocabulary Patterns in Expository and Narrative Texts

Expository Text1	Narrative Text2
DOZENS OF AFRICAN ELEPHANTS SLAUGHTERED ³ That headline has become all too common. Last month, poachers killed at least 86 elephants in Chad and in Cameroon. Both countries are in a region of Africa that has lost more than 60% of its elephants to illegal hunters in the past decade, according to a recent study from the Wildlife Conservation Society. In 2012 alone, experts say, 30,000 elephants were killed in countries across Africa. “We’re seeing the highest levels of poaching since our record-keeping began,” Crawford Allan, of the World Wildlife Fund, told TFK. Why are so many elephants being killed?	My four friends and I had come across a loose floorboard at the back of the classroom, and when we pried it up with the blade of a pocketknife , we discovered a big hollow space underneath. This, we decided, would be our own secret hiding place for sweets and other small treasures such as conkers , and monkey-nuts , and birds’ eggs. Every afternoon, when the last lesson was over, the five of us would wait until the classroom had emptied, then we would lift up the floorboard and examine our secret hoard , perhaps adding to it or taking something away.

¹ The Price of Ivory (April 19, 2013). *Time for Kids*. Retrieved from <http://www.timeforkids.com/news/price-ivory/89921>

² From Roald Dahl (1984). *Boy*. New York, NY: Puffin Books.

³ Boldfaced words are ones that appear rarely in written English (as defined by Hiebert (2005).

Both Hiebert (2008) and Mesmer (2008) write about how readability formulas can artificially inflate the difficulty of expository texts due to the repetition of infrequent words. When readability formulas are used, they count each infrequent word, whether or not it is repeated elsewhere in the passage, as an occurrence of a “hard” word. Thus, in the expository passage example in Table 4.3, the word ivory would be counted as a difficult word each time it occurred, despite the fact that repetition of the word actually provides the student with support and practice. This artifact of the formulas especially should cause teachers to carefully review expository texts before completely trusting the estimates delivered by the formulas.

As Mesmer et al. (2012) concluded in their review, a great deal more research must be conducted to understand exactly how genre operates within text complexity models, and this is true of models for stretching or challenging students as well. Genre may be best represented by multivariate approaches that characterize the many text features that represent the label. In addition, the text features that present challenges in each genre may differentially apply to various outcomes. For instance, prior knowledge may operate more in the expository format than in the narrative format. Clearly, a second generation of research is needed to move the typical diet of text in elementary classrooms beyond simply including various genres to challenge students appropriately.

Stretching Students Through Text Cohesion

The estimation of text difficulty has only recently gone beyond evaluating the difficulty of individual words and individual sentences. The classic readability formulas, and even the second-generation formulas, theoretically treat each word and sentence separately, as if each word and sentence were derived from a separate source. Their frequencies and lengths are only joined when entered into the equations, and the newer formulas give no consideration to the ways that the words and sentences in a text relate to each other. But recent work on text cohesion using a tool called Coh-Metrix has changed all this (Graesser, McNamara, & Kulikowich, 2011).

Text cohesion is the degree to which the words and ideas are represented both within and across sentences (Givón, 1995). When a text is cohesive, there is a thread that runs through it that allows readers to construct a connected gist of the main ideas. Texts with coherence marking have ideas repeated and introduced at a pace that optimally mixes

new and previously stated information. Table 4.4 illustrates elementary-level texts with different levels of cohesion. Note that in the most cohesive example, words and phrases are repeated often across sentences. In the least cohesive example, there is almost no repetition of words, and in the medium-cohesion text, there are repetitions, but their spacing is across paragraphs more than across sentences.

Table 4.4: Examples of Expository Texts from CCSS Exemplars at Three Levels of Referential Cohesion¹

Low Cohesion	Medium Cohesion	High Cohesion
This island is covered with snow. No trees grow. Nothing has green leaves. The land is white as far as you can see. Then something small and round and black pokes up out of the snow. A black nose sniffs the air. Then a smooth white head appears. (From <i>Where Do Polar Bears Live?</i> Thomson, 2010)	Horses move in four natural ways, called gaits or paces. They walk, trot, canter, and gallop. The walk is the slowest gait and the gallop is the fastest. When a horse walks, each hoof leaves the ground at a different time. It moves one hind leg first, and then the front leg on the same side. (From <i>Horses</i>, Simon, 2006)	Most plants make seeds. A seed contains the beginning of a new plant. Seeds are different shapes, sizes, and colors. All seeds grow into the same kind of plant that made them. Many plants grow flowers. Flowers are where most seeds begin. A flower is made up of many parts. (From <i>Seed to Plant</i>, Gibbons, 1991)

¹ Boldfaced words illustrate repetition of words across sentences within texts.

Graesser et al. (2011) treated cohesion marking as a multivariate variable, including many factors in its calculation. They visually demonstrated how the levels of five text cohesion factors—narrativity, syntactic simplicity, word concreteness, causal cohesion, and referential cohesion—unfold across the grades in the texts of language arts, social science, and science. The findings suggest, among other things, that narrativity is highest in the language arts text of the earliest grades and that referential cohesion is highest in science texts. As one might expect, syntactic simplicity is highest in the earliest texts, particularly in the science genre.

The introduction of cohesion to the estimation of text difficulty contributes greatly to the theoretical foundation upon which a paradigm of challenge might be based. What has been established is that text cohesion interacts with prior knowledge and student ability (Graesser et al., 2011). But current tools are quite complicated, and there is not a great deal of information about how differences in cohesion marking affect elementary-grade students. Much more must be learned about how cohesion can be pragmatically applied in classrooms.

Conclusions: Programmatically Addressing Challenge

The Common Core text complexity standard and overall focus on challenging text (NGA Center & CCSSO, 2010a) and the need for students to “stretch their reading abilities,” as outlined in Appendix A to the standards (NGA Center & CCSSO, 2010b), have introduced a major shift in reader–text matching paradigms that promises to balance the intense focus on the avoidance of frustration with the importance of challenge. Nonetheless, this introduction raises some important issues. Shanahan (2011) expressed the following:

We have tended to overgeneralize from younger readers (for whom easier text allows a more systematic focus on decoding) to older readers (who may do better with more intellectually challenging texts). Now, I fear that the Common Core is over-generalizing in the other direction. Harder beginning reading books may stop many young readers in their tracks. (p. 21)

I have this same fear, especially in light of the fact that the rationale for increasing text difficulty is based on studies of secondary students. When carefully examined, patterns in the data that are frequently cited to support claims of textbook simplification do not actually hold true for elementary students (Chall, 1977; Hayes et al., 1996).

Existing research is scant and simply not sufficient to support the increases in text levels required in the elementary grades by the CCSS’s staircase of text complexity. There is not enough empirical data to suggest exactly how students should be stretched in text; however, in this piece I identify text and other factors that may be considered in future research. In the past, classroom reactions to inappropriate text standards have been extreme. Either teachers (or, more likely, district supervisors) knuckle down and insist that every student in a given grade reads texts of a certain level or teachers abandon ship altogether and default to reading aloud anything that might be considered challenging. But I caution schools and teachers to resist what I call the “read-aloud solution”; instead, a blend of scaffolded challenge reading with some reading aloud should characterize stretching students in the elementary school.

At a basic level, teachers must know the reading levels of their students and estimates of the difficulty of the texts they wish to use. Although this is a basic tenet of reader–text matching, frequently the obvious is overlooked. While the reader–text matching standards of Betts (1946) should indeed be questioned, I caution educators to remember that stretch text should not cause frustration. Stretch texts, whatever the research ultimately decides they may be, should represent optimal challenge, not gut-wrenching exasperation. Shanahan (2011) notes the

opposite response to challenge that might occur: “When the books get hard, the usual responses have been to move kids to easier books, to stop using textbooks, or to read the texts to the students” (p. 20). How very ironic it would be if the text standards designed to challenge students in actuality water down their exposure to challenging texts.

As identified in this chapter, additional factors that may affect students’ abilities to be stretched include text levels, text length, genre, and cohesion. All of these are malleable factors that can be manipulated and designed into text. In presenting a framework for texts in the early grades, Mesmer et al. (2012) proposed four elements: content (e.g., words, concepts, sentences, ideas, genre), sequence in which the content is presented, pace of presentation, and repetition of content. As researchers develop a theory of challenge that contributes to the important notion of stretching students, each of these elements of a text program must be addressed. A paradigm for understanding how to stretch students in text must move beyond an isolated, drive-by approach to a more consistent, programmatic one. Stretching students cannot and should not be relegated only to a Friday afternoon read-aloud and discussion. It must be infused into the text choices made over weeks, months, and years. Certainly, the arguments put forth for challenge in the Common Core suggest that it is the accumulated effects of text that resulted in lower ACT scores or grades in college (ACT, 2006). So then must the approach to stretching students in text also be longitudinal, across days, weeks, months, and years. How text length, difficulty, genre, cohesion, and text levels are balanced and introduced across a unit of study or a developmental period will support or inhibit fruitful “stretching.” Focused and consistent efforts at presenting students with challenging texts that stretch their capacity will ultimately have the kinds of effects intended by the Common Core writers.

Are we going to lower the fences or teach kids to climb? asks Shanahan (2011) in the title of a recent *Reading Today* article. The message is important. For too long we have been overly concerned about the height of the fences and not concerned enough about teaching kids to climb. I think that stretching students in texts might be like adjusting the uneven bars in the gym. When gymnasts are at a certain level in their training, they are expected to mount the bars using a springboard or other device to begin their routines. This means that the bar is typically above their head and several feet ahead of them. They must run and bounce on the springboard and reach for the bar to begin the routine. Sometimes they fall on the dense 12-inch mats beneath them, but eventually they can consistently make it. Throughout a meet or workout, you will see coaches

raise and lower the bars to accommodate different heights because, even though the mount is challenging—and, in fact, over the heads of the gymnasts—there are still limits placed on the gymnast by factors such as height and arm length. No one expects the bar to be set the same for a gymnast who is four feet and three inches tall as it is for a gymnast who is four feet and eight inches tall. The same is true with stretch students. We want them to leap and grab, but we should set the bar relative to their characteristics. As argued in another piece, stretching students in text is a dynamic activity that cannot be dictated by static text difficulty standards (Mesmer & Hiebert, 2013). The duty of researchers is to continue to create knowledge to support teachers as they work to develop stronger readers in elementary school.

References

- ACT, Inc. (2006). *Reading between the lines: What the ACT reveals about college readiness in reading*. Iowa City, IA: Author. Retrieved from <http://www.act.org/research/policymakers/pdf/reading-report.pdf>.
- Beck, I.L., McKeown, M.G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction* (2nd ed.). New York, NY: Guilford.
- Betts, E.A. (1946). *Foundations of reading instruction, with emphasis on differentiated guidance*. New York, NY: American Book Company.
- Calfee, R.C., & Hiebert, E.H. (2011). *Using cohort analyses to examine long-term effects of reading initiatives in California*. Paper presented at the annual conference of the Society for the Scientific Study of Reading, St. Petersburg, FL. Retrieved from https://www.academia.edu/11191211/Using_cohort_analyses_to_examine_long-term_effects_of_reading_initiatives_in_California
- Chall, J.S. (1977). *An analysis of textbooks in relation to declining SAT scores*. Princeton, NJ: College Entrance Examination Board.
- Clay, M.M. (1985). *The early detection of reading difficulties* (3rd ed.). Portsmouth, NH: Heinemann.
- Coleman, D., & Pimentel, S. (2012). *Revised publishers' criteria for the Common Core State Standards in English Language Arts and Literacy, Grades 3–12*. Washington, DC: National Governors Association Center for Best Practices & Council of Chief State School Officers. Retrieved from http://www.corestandards.org/assets/Publishers_Criteria_for_3-12.pdf.
- Cunningham, J.W., Spadorcia, S.A., Erickson, K.A., Koppenhaver, D.A., Sturm, J.M., & Yoder, D.E. (2005). Investigating the instructional supportiveness of leveled texts. *Reading Research Quarterly*, 40(4), 410–427.
- Ehri, L.C., Dreyer, L.G., Flugman, B., & Gross, A. (2007). Reading rescue: An effective tutoring intervention model for language-minority students who are struggling readers in first grade. *American Educational Research Journal*, 44(2), 414–448.
- Ekwall, E.E., Solis, J.K.E., & Solis, E.S., Jr. (1973). Investigating informal reading inventory scoring criteria. *Elementary English*, 50(2), 271–274, 323.
- Fang, Z. (2006). The language demands of science reading in middle school. *International Journal of Science Education*, 28(5), 491–520.
- Gamson, D. A., Lu, X., & Eckert, S. A. (2013). Challenging the research base of the Common Core state standards: A historical reanalysis of text complexity. *Educational Researcher*, 42(7), 381–391.
- Givón, T. (1995). *Functionalism and grammar*. Philadelphia, PA: John Benjamins.
- Graesser, A.C., McNamara, D.S., & Kulikowich, J.M. (2011). Coh-Metrix providing multilevel analyses of text characteristics. *Educational Researcher*, 40(5), 223–234.
- Halladay, J.L. (2012). Revisiting key assumptions of the reading level framework. *The Reading Teacher*, 66(1), 53–62.
- Harrison, C. (1980). *Readability in the classroom*. Cambridge, UK: Cambridge University Press.
- Hatcher, P.J. (2000). Predictors of reading recovery book levels. *Journal of Research in Reading*, 23(1), 67–77.
- Hayes, D.P., Wolfer, L.T., & Wolfe, M.F. (1996). Sourcebook simplification and its relation to the decline in SAT-Verbal scores. *American Educational Research Journal*, 33(2), 489–508.
- Hiebert, E.H. (2005). In pursuit of an effective, efficient vocabulary curriculum for the elementary grades. In E.H. Hiebert & M. Kamil (Eds.), *The teaching and learning of vocabulary: Bringing scientific*

research to practice (pp. 243–263). Mahwah, NJ: LEA.

Hiebert, E.H. (2008). The word zone fluency curriculum: An alternative approach. In M.R. Kuhn & P.J. Schwanenflugel (Eds.), *Fluency in the classroom* (pp. 154–170). New York, NY: Guilford.

Hiebert, E.H., & Martin, L.A. (2015). Changes in the texts of reading instruction during the past fifty years. In P.D. Pearson & E.H. Hiebert (Eds.), *Research-based practices for teaching Common Core literacy*. New York, NY: Teachers College Press.

Hiebert, E.H., & Mesmer, H.A.E. (2013). Upping the ante of text complexity in the Common Core state standards: Examining its potential impact on young readers. *Educational Researcher*, 42(1), 44–51.

Hiebert, E.H., Wilson, K.M., & Trainin, G. (2010). Are students really reading in independent reading contexts? An examination of comprehension-based silent reading rate. In E. H. Hiebert & D. R. Reutzel (Eds.), *Revisiting silent reading: New directions for teachers and researchers* (pp. 151–167). Newark, DE: International Reading Association.

Martin, M.O., Mullis, I.V.S., & Kennedy, A.M. (Eds.). (2007). *Progress in international reading literacy study (PIRLS): PIRLS 2006 Technical Report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center.

Mesmer, H.A.E. (2008). *Tools for matching readers to texts: Research-based practices*. New York, NY: Guilford.

Mesmer, H.A., Cunningham, J.W., & Hiebert, E.H. (2012). Toward a theoretical model of text complexity for the early grades: Learning from the past, anticipating the future. *Reading Research Quarterly*, 47(3), 235–258.

Mesmer, H.A.E., & Hiebert, E.H. (April 9, 2011). *An examination of the effects of discrepancies in reader-text match on comprehension*. Paper presented at the annual meeting of the American

Educational Research Association, New Orleans, LA.

Mesmer, H.A.E., & Hiebert, E.H. (2013). *How far can third graders be “stretched”? Exploring the influence of text difficulty and length*. Unpublished manuscript.

Morgan, A., Wilcox, B.R., & Eldredge, J.L. (2000). Effect of difficulty levels on second-grade delayed readers using dyad reading. *The Journal of Educational Research*, 94(2), 113–119.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010a). *Common Core state standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Author. Retrieved from www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010b). *Common Core state standards for English language arts & literacy in history/social studies, science, and technical subjects: Appendix A: Research supporting key elements of the standards and glossary of key terms*. Washington, DC: Author. Retrieved from http://www.corestandards.org/assets/Appendix_A.pdf

Nelson, J., Perfetti, C., Liben, D., & Liben, M. (2012). *Measures of text difficulty: Testing their predictive value for grade levels and student performance*. Washington, DC: Council of Chief State School Officers.

Pikulski, J.J., & Shanahan, T. (1982). Informal reading inventories: A critical analysis. In J.J. Pikulski & T. Shanahan (Eds.), *Approaches to the informal evaluation of reading* (pp. 94–116). Newark, DE: International Reading Association.

RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.

Roskos, K., & Neuman, S.B. (2013). Common Core, commonplaces, and

community in teaching reading. *The Reading Teacher*, 66(6), 469–473.

Shanahan, T. (2011). Common Core standards: Are we going to lower the fences or teach kids to climb? *Reading Today*, 29(1), 20–21.

Stahl, S.A., & Heubach, K.M. (2005). Fluency-oriented reading instruction. *Journal of Literacy Research*, 37(1), 25–60.

Stenner, A.J., Burdick, H., Sanford, E.E., & Burdick, D.S. (2007). *The Lexile Framework for reading* (Technical Report). Durham, NC: MetaMetrics.

Williamson, G.L. (2006). *Aligning the journey with a destination: A model for K–16 reading standards*. Durham, NC: MetaMetrics.

CHAPTER 5

The Relationship Between a Silent Reading Fluency Instructional Protocol on Students' Reading Comprehension and Achievement in an Urban School Setting¹

Timothy V. Rasinski
Kent State University

S. Jay Samuels
University of Minnesota

Elfrieda H. Hiebert
TextProject & University of California, Santa Cruz

Yaacov Petscher
Florida Center for Reading Research

Karen Feller
Reading Plus

Reading fluency has been defined as the ability to simultaneously process written texts accurately, automatically, and with appropriate prosody and comprehension (NICHD, 2000; Rasinski, 2006, 2010). Although it has been relatively neglected in reading curricula and instruction for years (Allington, 1983; Rasinski & Zutell, 1996), recent reviews of empirical research have identified reading fluency as a critical element in successful literacy instruction (Chard, Vaughn, & Tyler, 2002; Kuhn & Stahl, 2003; NICHD, 2000; Rasinski & Hoffman, 2003).

Chall's (1996) model of reading development posits reading fluency as a task to be mastered in the primary grades, and indeed most research on fluency to date has focused on the primary grades. For example, several

¹ This chapter was previously published in *Reading Psychology*, (v32, n1, p75-97). The definitive publisher-authenticated version published in 2011: <http://www.tandfonline.com/doi/abs/10.1080/02702710903346873#.VYLxEOf7LDE>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

studies report significant correlations between and predictive ability of measures of oral reading fluency and third-grade student performance on the reading portion of the Florida Comprehensive Assessment Test (FCAT), a criterion-referenced test (CRT) of reading achievement that aligns with Florida's Sunshine State Standards (Buck & Torgesen, 2003; Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2008). Similarly, instructional research into fluency has generally focused on the primary grades (e.g., Rasinski, Padak, Linek, & Sturtevant, 1994; Rasinski & Stevenson, 2005; Stahl & Heubach, 2005). This research has consistently found positive effects for fluency instruction not only on students' reading fluency but also on their word recognition, comprehension, and overall reading achievement.

More recently, scholars have suggested that reading fluency also may be an important concern for students beyond the elementary grades (Schatschneider et al., 2004; Torgesen, Nettles, Howard, & Winterbottom, 2004). Rasinski, Padak, McKeon, Wilfong, Friedauer, and Heim (2005), for example, report a robust and significant correlation between a measure of high school students' reading fluency (automaticity) and a measure of silent reading comprehension. Moreover, significant numbers of high school students in the study were found to be substantially below norms of acceptable performance in reading fluency and also comprehension. Rasinski, Rikli, and Johnston (2009) reported significant and substantial correlations between measures of fluency (prosody) among upper elementary and middle school students and a standardized test of silent reading comprehension. Additionally, the magnitude of the correlation is roughly the same at the three grade levels studied—grades 3, 5, and 8. Reading fluency, it appears, is not an issue solely for the primary grades, as fluency is associated with reading achievement beyond the primary grades. Furthermore, significant numbers of students beyond the primary grades have yet to achieve appropriate levels of fluency in their reading, and, as a result, these students also experience difficulties in comprehension and general reading achievement.

Most definitions of reading fluency tend to associate it with oral reading. Prosodic or expressive reading, for example, one aspect of fluency, is most often associated with and observed in oral reading. When an individual reads orally, prosody, or a lack of prosody, becomes clearly apparent. However, prosody is not observable during silent reading. Moreover, most instructional methods for fostering fluency in students involve some form of oral reading.

Despite the overt focus on oral reading for fluency development, all

fluency instruction presupposes a link to silent reading and silent reading comprehension (Rasinski, 2006, 2010). More to the point, oral fluency instruction presumes that improvements in oral fluency and oral reading comprehension will also be manifested in silent reading fluency and silent reading comprehension. Because silent reading is such a ubiquitous form of reading beyond the elementary grades, instruction in oral reading is worthwhile primarily to the extent that it can positively impact readers' silent reading comprehension.

Oral reading instruction does pose some serious practical limitations, however. Since oral reading is not as common a form of reading as silent reading beyond the primary grades, oral reading may not have the same degree of face validity or authenticity as silent reading. In group instructional settings, oral reading is most often conducted with one student at a time; other students in the group usually do not read while another student is reading. Efficiency in the use of time for reading is, thus, diminished. In addition, oral reading by one student may cause disruptions for other students. Listening to classmates reading orally may cause students in a classroom to become distracted and devote less attention to their own reading or learning task. Finally, by the middle grades, fear and embarrassment as a result of miscues made while reading orally can further diminish the effectiveness of oral reading activities and students' confidence in their own reading.

These limitations beg the question, then: Is it possible to promote fluency in reading, and thereby improve comprehension, through silent reading instruction? In an initial study into this question, Reutzel, Jones, Fawson, and Smith (2008) reported positive results with silent reading fluency instruction. Using an instructional method called Scaffolded Silent Reading (ScSR) with third-grade students, Reutzel and his colleagues found gains in word recognition, reading rate, prosody, and comprehension that were essentially equal to gains typically found through repeated guided oral reading instruction. ScSR was designed to counter concerns and limitations that have been raised about independent or sustained silent reading (NICHD, 2000). Embedded in the ScSR instructional framework are the following:

- Teacher guidance in selecting appropriately challenging materials
- High levels of engagement in reading during time allotted for reading
- Teacher interaction with students after reading
- Feedback given to students about the quality and quantity of their reading

- Student accountability for the time spent in silent reading

The present retrospective study extends the work of Reutzel and colleagues (2008) by testing the effects of a program designed to teach and improve silent reading fluency on the reading comprehension and overall reading achievement of elementary, middle school, and high school students (grades 4 through 10) in a large urban school district.

Background

This study was conducted in cooperation with Miami-Dade County (Florida) Public Schools to determine the relationship between student participation in a silent reading instructional program and overall student reading achievement in grades 4 through 10, as measured by the FCAT with selected schools in Regions II and III of the Miami-Dade County Public Schools. The experimental treatment employed in the study was Reading Plus (RP), a computer-based reading fluency and comprehension intervention system that is designed to develop silent reading fluency and overall reading proficiency.

Method

Participants

A total of 16,143 students from grades 4 through 10 in 23 schools in Regions II and III in the Miami-Dade County Public School System participated in the study; 5,758 students made up the treatment group, while the remaining 10,385 students constituted the control group. Both regions of the district had significant populations of minority students with 34% African American and 56% Latino American.

Subpopulations in the sample included the following:

- Learning disabled (LD; 6% of total; 541 participating, and 491 nonparticipating)
- English language learners (ELLs; 3% of total; 176 participating, and 286 nonparticipating)

The 23 schools were distributed across elementary (11) and middle and high schools (12). In a number of schools, only those students who scored achievement level 1 or 2 (nonproficient) on the 2006 Reading portion of the FCAT were assigned to RP. In other schools, students from specific grade levels or subpopulations were assigned. Most nonparticipating students who engaged in alternative interventions were assigned to Scholastic's Read 180 and/or Renaissance Learning's

Accelerated Reader. Elementary-level students (grades 4 and 5) received reading and language arts reading instruction in their regular curriculum.

In all cases, treatment students were those who had: (a) completed one or more RP lessons during the 2006 to 2007 school year and (b) had valid 2006 and 2007 FCAT Reading scores as recorded in the Miami-Dade County Student Information System. As the data in Table 5.1 indicate, students who were chosen for the RP intervention were performing at significantly lower levels than their classmates in the control condition.

Procedures

At the beginning of the 2006-2007 school year, teachers in the two regions of the school district were trained on the intent and use of the RP program and were guided in identifying appropriate students from their classes to participate in the intervention. Implementation began soon after and continued until administration of the 2007 FCAT in early March of 2007.

Prior to the implementation of the intervention, students completed the Reading Placement Appraisal assessment to establish their initial placement level in RP. The 20-minute placement test assessed independent reading rate, comprehension, and vocabulary to determine the most appropriate starting level. The placement assessment consisted of three parts. Part I presented students with 100-word selections followed by a set of literal-recall questions. Content difficulty was automatically adjusted by the program according to a student's reading rate and comprehension to ascertain the independent reading level. Part II presented 300-word selections followed by a set of diverse comprehension questions to confirm the independent reading level. Part III assessed a student's vocabulary. From these, an instructional reading level was established, and students were placed at appropriate levels within each component of the program. Students continued to be assessed on similar tasks throughout the program with appropriate adjustments made to the level of activities as a result of their performances on these formative assessments.

The RP intervention involved students in a series of lessons that were provided on a digitized network platform in individual computer environments. A specific sequence of activities was followed during the lesson period, and the difficulty level of the activities was adjusted as a function of a student's progress. Each RP lesson required approximately 30 minutes to complete. Treatment schedules varied within the 23 schools, but most schools followed a schedule of either two 45-minute sessions per

week or three 30-minute sessions per week for approximately six months. Students who were part of the 45-minute session schedule generally completed more than one guided reading lesson per session.

Each lesson began with a perceptual accuracy and visual efficiency (PAVE) warmup. This activity consisted of two parts, Scan and Flash. In the Scan portion of the activity, students scanned the computer screen to count the number of times a target letter or number appeared on the screen. The target and other letters or numbers were flashed in a left-to-right presentation. The presentation speed increased in accordance with the student's proficiency. In the second part of the activity, Flash, a series of letters or numbers ranging in length from 2 to 12, depending on the student's placement level, was flashed across the screen (at 1/6 of a second per flash). The length of the flash increased in response to the student's ability to correctly recreate the sequence. This warm-up activity aimed to increase students' visual perception, attention skills, and automaticity in the recognition of print. Research (e.g., Torgesen & Hudson, 2006) suggests that one of the defining characteristics of a proficient reader is the ability to sustain attention. According to Pikulski (2006), "instant, accurate, and automatic access to all these dimensions of a printed word is the needed fluency that will allow readers to focus their attention on comprehension rather than on decoding" (p. 90).

The next RP activity, Guided Reading, provided students with extensive structured silent reading practice in order to build fluency within an authentic reading experience where students read for meaning. During Guided Reading sessions, students read texts selected from a diverse collection of narrative and expository stories at their instructional reading level, a practice that research has supported. The work of O'Connor, Harty, Larkin, Sackor, and Zigmond (2002) showed that providing daily intervention lessons using grade-level texts was not nearly as successful as providing daily lessons using texts matched to the instructional reading levels of the individual students. O'Connor et al. argue that selecting texts of appropriate complexity should be a first step in the design of effective instruction and intervention.

RP selections were leveled using several readability formulas (e.g., Spache, Dale-Chall, and Fry). The RP program was designed to automatically, continually, and dynamically monitor student performance and progress, adjusting the reading content level to match each student's achievement. Once students were able to read passages at their current levels with grade-appropriate rates and good comprehension, they would be able to advance to subsequent levels. In addition, the program used a

mix of instructional formats and scaffolds to further match individualized needs and rates of progress. These included variation of the length of reading segments, number of comprehension questions, use of repeated readings, and assignment of prereading techniques, as research on fluency development has demonstrated that struggling and developing readers are the least likely to engage in the effective practice that would provide them with the opportunity to integrate the varied reading instruction they receive (Allington, 2006; Chinn, Waggoner, Anderson, Schommer & Wilkinson, 1993; Hiebert, 1983).

The RP program contained approximately 600 reading selections, ranging from preprimer to adult-level texts, including high-content, low-readability selections for older struggling students. A wide range of genres was featured, including selections such as “The Lighthouse Visitor,” a mystery on a third-grade level; a fifth-grade nonfiction selection about, “How Basketball Was Born”; and a tenth-grade nonfiction selection on “Peer Counseling.” As students progressed through the levels, the content became increasingly informational. Lesson texts were presented in either a guided or independent manner, each within controlled presentation formats and rate parameters.

Following each reading selection were comprehension questions coded for specific comprehension skills, including literal understanding, interpretation, analysis, evaluation, and appreciation. The rate at which the text was presented was incrementally increased as a function of students’ comprehension performance on these questions. As students progressed through the levels, the texts became progressively more challenging. The intent of the Guided Reading activities was to provide students with authentic reading experiences that would build comprehension and fluency and that would be presented at a level of difficulty that would provide the maximum acceleration of progress. Additionally, given that the difficulty of texts was established using the Spache (for primary grade-level texts), Dale-Chall (for middle grade-level texts), and Fry (for primary- and middle-level texts) readability formulas, all of which rely on high-frequency word lists, students had considerable opportunities to develop fluency with a core group of high-frequency words. This is an essential skill, as Torgesen and colleagues (Rashotte, MacPhee, & Torgesen, 2001; Torgesen & Hudson, 2006) argue that limited sight vocabularies are a principle characteristic of students who continue to have reading disabilities beyond the initial phase of learning to read.

The Guided Reading component was followed by a cloze comprehension activity that used structured context-analysis tasks to

develop comprehension competency. It employed a dual approach that combined foci uses on both improving students' comprehension as well as vocabularies. Each cloze activity required students to use context to complete the meaning of sentences and passages, thus enhancing comprehension. Students also were required to derive the meaning of difficult or unfamiliar words by analyzing the information in the surrounding context, thus enhancing vocabulary.

The vocabulary component of the RP lesson focused on 240 vocabulary words per grade level. Students completed contextual word-meaning activities on words that they missed on the pretest. Each word was first presented in a sentence that the computer program provided orally. Next, the word was used in the context of a paragraph. Finally, students were asked to select from choices provided the sentences that demonstrated proper usage and meaning of the target word. The passage was available for rereading, with clues from the passage highlighted after an incorrect response.

Assessments

The FCAT was part of a statewide initiative to raise academic standards for students in the state of Florida. The FCAT consisted of two kinds of tests. The first was a CRT, which measured how well students were meeting the Sunshine State Standards in reading, writing, mathematics, and science. The second was a norm-referenced test (NRT), which permitted a comparison of Florida student performance on reading and mathematics with the performance of students nationwide. The NRT used during the time of this study was the Stanford Achievement Test-10 (SAT-10). The reading section evaluated students' ability to understand the meaning of informational and literary passages. Both portions of the FCAT were administered to all students in grades 3 through 10, and results were reported publicly in summary form. Pretesting occurred during the spring 2006 administration of the FCAT. Posttesting occurred during the spring 2007 administration of the FCAT.

Results

Data Analysis

A 3 x 7 x 3 x 2 x 2 (Group x Grade x Minority x ELL x LD) analysis of variance (ANOVA) was used to test if differences existed in the simple difference score of the posttest minus the pretest among the groups

receiving different levels of treatment. Contrasts were conducted in the ANOVA pertaining to the main effects of grade level, minority status, and ELL and LD identification to examine if groups differed in their mean gain score across levels of the intervention. To control for multiple statistical tests being employed on the FCAT CRT and NRT on the same sets of students, Benjamini and Hochberg's (1995) Linear Step-Up procedure was employed. This procedure differs slightly from other type 1 error control procedures in that, in its simplest form, it attempts to control the false discovery rate by aiming to keep the ratio of false rejections to total rejections at 5%. Specifically, when all null hypotheses are true, the Linear Step-Up procedure will control the experiment-wise error rate at .05 (just as other traditional approaches attempt). However, when some of the null hypotheses are false, the Linear Step-Up will ensure that the false rejection rate does not go above 5%. The benefit to this approach is that it appears to be more powerful than traditional approaches, such as the Bonferroni correction (Maxwell & Delaney, 2004). In addition to hypothesis testing of means among groups, a standardized effect size (i.e., Cohen's *d*) was used to express the distributional differences in standard deviation units. Cohen (1988) has provided guidelines that suggest that an effect size of 0.20 is small, 0.50 is medium, and 0.80 is large; however, he is quick to note that the qualitative designation for the magnitude of the effect is largely contextual. This has been echoed more recently by Hill, Bloom, Black, and Lipsey (2008), who argued that these guidelines are somewhat inefficient for interpreting achievement or intervention effects in education.

It is important to note that in instances where random assignment does not occur, covarying preexisting differences on the pretest is not necessarily the most appropriate procedure, since variability on baseline scores may be attributed to the lack of random assignment and reflect meaningful initial values (Maxwell & Delaney, 2004). While some opt to use a posttest-only approach to the analyses of group differences, doing so ignores the value of the baseline score. An alternative strategy is to utilize initial performance to calculate a gain score that allows a meaningful comparison of change between two time points. Though the difference score has been often maligned as a poor index of change (e.g., Cronbach & Furby, 1970), Rogosa (1995) has shown that gain scores are as reliable as a covariance adjusted score and are more appropriate than posttest scores only for use in quasi-experimental studies. Moreover, it has been well established that results from an ANOVA of gain scores are identical to results from a repeated measures ANOVA with two time points and two groups (Huck & McLean, 1975; Maxwell & Delaney, 2004).

A summary of the ANOVA results for the FCAT CRT and NRT are reported in Table 5.1, with subsequent post-hoc data reported for subgroups in Tables 5.2 through 5.7. Results indicated that significant main effects existed for grade level, ELL status, and LD identification; with interactions between grade and group, ELL status and group, and LD identification and group also statistically significant for the FCAT Reading CRT measure. Somewhat similar findings were observed for the NRT analyses, whereby significant effects occurred for grade, ELL status, grade X group, and ELL X group.

Table 5.1: ANOVA Results for Florida CRT and NRT Outcomes

Measure	Source	df	F	p-value
CRT	Grade	6	68.94	<0.001
	Minority	2	3.35	0.035
	ELL	1	88.31	<0.001
	LD	1	3.89	0.032
	Group	2	4.14	0.160
	Grade X Group	12	3.29	<0.001
	Minority X Group	4	0.62	0.649
	ELL X Group	2	8.92	<0.001
	LD X Group	2	3.11	0.044
	Error	7538		
NRT	Grade	6	133.79	<0.001
	Minority	2	0.69	0.503
	ELL	1	6.61	0.010
	LD	1	3.22	0.079
	Group	2	0.33	0.721
	Grade X Group	12	2.07	0.016
	Minority X Group	4	1.55	0.184
	ELL X Group	2	4.50	0.011
	LD X Group	2	2.54	0.095
	Error	7897		

Note. *p*-values reflect Linear Step-Up adjustments.

Table 5.2 presents FCAT Reading (CRT) Developmental Scale gain scores and SAT–10 (NRT) gain scores by grade level for all students who participated in 1 to 39 RP lessons, students who received 40 or more RP lessons, and students who received no RP lessons. RP students had significantly greater gains than non-RP students in grades 5, 6, 7, 8, and 9 on the CRT and in grades 4, 5, 6, 7, 8, and 10 on the NRT. Students receiving the RP intervention experienced significantly greater reading

Table 5.2: Gain Scores on the FCAT Reading (CRT) Developmental Scale Scores and SAT-10 (NRT) for All Students

Measure	Grade	No Lessons			1–39 Lessons			40+ Lessons			Contrast		Effect Size		
		N	M	SD	N	M	SD	N	M	SD	F	p-value	d1	d2	d3
CRT	4	529	158.75	224.69	461	162.18	220.81	340	181.42	200.85	2.05	0.160	0.02	0.09	0.10
	5	449	71.43	216.53	393	78.37	200.07	364	117.46	209.16	9.32	0.006	0.03	0.20	0.21
	6	1423	48.03	216.19	563	80.45	237.77	217	130.06	240.08	28.60	0.002	0.15	0.21	0.38
	7	1256	46.27	199.35	508	109.19	212.73	307	157.78	212.40	88.25	0.002	0.32	0.23	0.56
	8	1546	44.76	180.50	502	128.45	195.77	403	137.20	185.51	113.58	0.002	0.46	0.04	0.51
	9	2803	66.48	190.09	406	84.31	202.17	328	107.23	203.30	14.85	0.002	0.09	0.11	0.21
	10	2379	33.78	215.55	521	22.70	207.14	445	20.39	182.12	2.16	0.160	-0.05	-0.01	-0.06
NRT	4	528	5.05	26.17	459	7.04	24.07	337	11.74	21.82	14.76	0.002	0.08	0.19	0.26
	5	445	13.60	22.10	391	20.33	25.03	360	21.19	23.40	21.94	0.002	0.30	0.03	0.34
	6	1416	11.36	24.35	560	11.77	23.05	217	17.60	23.62	8.78	0.006	0.02	0.25	0.26
	7	1239	5.06	23.21	497	5.64	22.60	303	9.22	22.37	6.66	0.024	0.02	0.16	0.18
	8	1530	7.46	25.13	482	10.20	25.68	393	11.97	22.57	12.25	0.002	0.11	0.07	0.18
	9	2719	13.06	28.12	383	7.12	31.47	324	14.17	27.01	0.86	0.363	-0.21	0.22	0.04
	10	2267	0.45	29.29	465	6.60	28.15	415	8.24	24.16	35.95	0.002	0.21	0.06	0.27

Note. p-values reflect Linear Step-Up adjustments.

Table 5.3: Gain Scores on the FCAT Reading (CRT) Developmental Scale Scores and SAT-10 (NRT) for African American Students Receiving 40+ Lessons of the RP Intervention Versus Students Receiving No RP Lessons

Measure	Grade	No Lessons			1–39 Lessons			40+ Lessons			ANOVA		Effect Size		
		N	M	SD	N	M	SD	N	M	SD	F	p-value	d1	d2	d3
CRT	4	236	147.01	243.59	234	133.40	224.47	162	176.31	211.79	1.19	0.310	-0.06	0.19	0.12
	5	158	69.93	229.94	193	60.77	194.51	235	90.20	204.82	1.14	0.310	-0.04	0.15	0.09
	6	480	12.77	203.35	267	38.50	223.58	113	89.80	229.47	11.87	0.003	0.13	0.23	0.38
	7	310	34.55	160.09	234	100.80	199.30	167	143.91	211.69	40.30	0.003	0.41	0.22	0.68
NRT	8	447	28.85	172.17	211	95.86	201.78	208	126.10	180.07	45.38	0.003	0.39	0.15	0.56
	9	760	52.08	182.64	110	50.77	235.27	113	85.08	217.77	2.22	0.200	-0.01	0.15	0.18
	10	465	16.62	227.29	195	13.69	221.68	226	-4.59	186.62	1.33	0.310	-0.01	-0.08	-0.09
	4	236	7.47	27.75	232	4.78	24.69	161	13.39	21.77	3.96	0.092	-0.10	0.35	0.21
	5	155	14.92	22.40	193	20.77	25.26	232	19.43	24.11	1.62	0.172	0.26	-0.05	0.20
	6	475	10.46	24.18	266	9.25	21.70	113	17.12	22.76	3.69	0.093	-0.05	0.36	0.28
	7	311	1.64	22.63	228	6.25	22.11	165	10.20	21.58	16.84	0.030	0.20	0.18	0.38
	8	439	7.37	24.22	200	7.70	23.96	205	14.61	21.35	11.42	0.003	0.01	0.29	0.30
	9	740	13.53	25.38	108	8.58	30.08	110	14.54	24.47	0.10	>.500	-0.19	0.20	0.04
	10	436	1.56	27.89	170	10.72	27.63	210	8.13	22.40	11.67	0.003	0.33	-0.09	0.24

Note. *p*-values reflect Linear Step-Up adjustments.

Table 5.4: Gain Scores on the FCAT Reading (CRT) Developmental Scale Scores and SAT-10 (NRT) for Latino American Students Receiving 40+ Lessons of the RP Intervention Versus Students Receiving No RP Lessons

Measure	Grade	No Lessons			1–39 Lessons			40+ Lessons			ANOVA		Effect Size		
		N	M	SD	N	M	SD	N	M	SD	F	p-value	d1	d2	d3
CRT	4	261	172.09	208.66	206	194.28	206.21	160	183.09	192.18	0.46	0.500	0.11	-0.05	0.05
	5	257	63.47	204.15	189	93.24	207.88	110	181.68	215.96	22.63	0.002	0.15	0.43	0.58
	6	842	70.66	217.61	271	124.06	240.62	94	167.21	248.43	23.34	0.002	0.25	0.18	0.44
	7	802	49.42	217.78	218	118.21	226.99	120	171.46	223.43	42.11	0.002	0.32	0.23	0.56
NRT	8	935	58.07	186.99	245	165.63	184.04	175	151.53	195.10	66.74	0.002	0.58	-0.08	0.50
	9	1686	70.17	198.09	259	95.46	188.81	178	120.52	202.56	12.93	0.002	0.13	0.13	0.25
	10	1553	35.92	218.24	288	21.57	193.29	190	45.94	176.45	0.01	0.500	-0.07	0.13	0.05
	4	260	3.89	24.78	206	9.68	22.80	158	10.46	22.11	8.87	0.007	0.23	0.03	0.26
	5	256	12.34	22.45	187	19.80	24.74	109	25.79	20.77	28.89	0.002	0.33	0.24	0.60
	6	839	11.69	24.49	269	14.45	23.09	94	17.13	24.54	6.01	0.061	0.11	0.12	0.22
	7	786	6.80	23.64	215	5.17	23.54	119	8.66	23.44	0.07	0.500	-0.07	0.15	0.08
	8	927	7.52	25.64	237	11.96	26.98	168	8.93	23.04	2.18	0.163	0.17	-0.11	0.06
	9	1626	11.95	29.26	238	6.82	32.10	177	13.59	28.49	0.20	0.651	-0.18	0.21	0.06
	10	1484	1.42	29.70	259	5.31	28.53	178	8.08	25.68	10.84	0.002	0.13	0.10	0.22

Note. *p*-values reflect Linear Step-Up adjustments.

Table 5.5: Gain Scores on the FCAT Reading (CRT) Developmental Scale Scores and SAT-10 (NRT) for Caucasian Students Receiving 40+ Lessons of the RP Intervention Versus Students Receiving No RP Lessons

Measure	Grade	No Lessons			1–39 Lessons			40+ Lessons			ANOVA		Effect Size		
		N	M	SD	N	M	SD	N	M	SD	F	p-value	d1	d2	d3
CRT	4	20	120.05	220.68	15	180.47	299.45	9	216.89	201.07	1.11	0.500	0.27	0.12	0.44
	5	25	101.68	200.66	9	116.78	129.20	11	118.27	115.22	0.09	0.500	0.08	0.01	0.08
	6	65	53.82	249.75	21	11.76	273.90	6	307.33	165.52	1.70	0.500	-0.17	1.08	1.02
	7	131	49.97	166.95	35	67.26	212.11	18	189.72	141.61	8.08	0.032	0.10	0.58	0.84
	8	118	6.92	156.06	33	81.88	214.94	16	113.19	148.35	8.90	0.032	0.48	0.15	0.68
	9	302	78.11	158.26	31	107.10	178.94	32	132.28	149.32	3.98	0.185	0.18	0.14	0.34
	10	300	38.62	177.59	30	39.43	176.71	22	50.50	172.80	0.07	0.500	0.00	0.06	0.07
NRT	4	20	-3.65	22.59	15	4.53	26.40	9	5.11	22.03	1.07	0.500	0.36	0.02	0.39
	5	25	14.88	16.64	9	17.22	25.34	11	19.91	19.19	0.55	0.500	0.14	0.11	0.30
	6	66	14.15	24.81	21	7.24	30.40	6	29.17	27.62	0.35	0.500	-0.28	0.72	0.61
	7	130	2.93	21.40	33	3.61	21.67	17	1.82	22.28	0.01	0.500	0.03	-0.08	-0.05
	8	119	7.01	23.67	32	6.78	23.02	16	7.69	25.38	0.01	0.500	-0.01	0.04	0.03
	9	298	16.92	28.46	30	0.80	33.94	32	16.69	29.11	1.20	0.500	-0.57	0.47	-0.01
	10	288	-4.87	28.03	27	-4.11	22.11	20	11.80	23.36	5.41	0.032	0.03	0.72	0.59

Note. *p*-values reflect Linear Step-Up adjustments.

Table 5.6: Gain Scores on the FCAT Reading (CRT) Developmental Scale Scores and SAT-10 (NRT) for Learning Disabled Students Receiving 40+ Lessons of the RP Intervention Versus Students Receiving No RP Lessons

Measure	Grade	No Lessons			1-39 Lessons			40+ Lessons			ANOVA		Effect Size		
		N	M	SD	N	M	SD	N	M	SD	F	p-value	d1	d2	d3
CRT	4	32	275.44	242.57	39	134.69	383.83	24	166.25	287.98	1.93	0.500	-0.58	0.08	-0.45
	5	23	60.52	399.07	19	53.84	260.44	29	100.72	204.83	0.26	0.500	-0.02	0.18	0.10
	6	67	109.82	243.18	78	-10.95	298.95	20	148.70	288.44	0.48	0.500	-0.50	0.53	0.16
	7	51	131.02	317.02	17	43.82	272.88	18	127.44	276.30	0.08	0.500	-0.28	0.31	-0.01
	8	80	92.93	297.24	74	157.54	206.27	31	117.03	249.77	0.76	0.500	0.22	-0.20	0.08
	9	149	48.01	256.80	62	42.37	279.07	22	75.91	130.70	0.09	0.500	-0.02	0.12	0.11
	10	89	-29.31	276.77	85	-18.89	284.61	23	-47.74	217.10	0.01	0.500	0.04	-0.10	-0.07
NRT	4	90	-3.52	27.86	44	-1.20	26.27	7	-1.43	17.16	0.30	0.500	0.08	-0.01	0.08
	5	69	12.39	25.22	51	18.12	27.19	28	28.71	26.90	1.21	0.500	0.23	0.39	0.65
	6	282	11.83	28.00	40	8.05	24.37	7	19.71	31.92	0.11	0.500	-0.14	0.48	0.28
	7	270	5.45	23.36	115	1.80	21.94	11	17.18	16.35	0.57	0.500	-0.16	0.70	0.50
	8	384	7.49	25.39	58	4.31	32.33	12	7.75	21.35	6.35	0.083	-0.13	0.11	0.01
	9	414	20.30	30.30	20	9.85	22.93	5	17.20	28.01	0.02	0.500	-0.35	0.32	-0.10
	10	445	-9.16	25.75	22	13.55	28.10	19	2.63	26.49	0.67	0.500	0.88	-0.39	0.46

Note. *p*-values reflect Linear Step-Up adjustments.

Table 5.7: Gain Scores on the FCAT Reading (CRT) Developmental Scale Scores and SAT-10 (NRT) for English Language Learners Receiving 40+ Lessons of the RP Intervention Versus Students Receiving No RP Lessons

Measure	Grade	No Lessons			1–39 Lessons			40+ Lessons			ANOVA		Effect Size		
		N	M	SD	N	M	SD	N	M	SD	F	p-value	d1	d2	d3
CRT	4	25	466.72	388.73	27	296.81	284.58	14	276.79	162.04	4.18	0.043	-0.44	-0.07	-0.49
	5	16	284.38	504.17	37	137.76	316.82	23	247.74	278.86	0.01	0.500	-0.29	0.35	-0.07
	6	65	308.68	315.45	18	184.39	364.18	11	164.18	214.91	3.22	0.102	-0.39	-0.06	-0.46
	7	89	263.81	286.30	7	247.00	390.93	15	253.00	261.51	0.03	0.500	-0.06	0.02	-0.04
NRT	8	91	198.20	274.16	7	129.57	250.81	17	252.71	176.50	0.36	0.500	-0.25	0.49	0.20
	9	26	24.12	24.97	27	23.81	15.30	14	17.00	20.70	0.01	0.500	-0.01	-0.45	-0.28
	10	16	8.69	24.48	37	22.70	27.43	22	32.68	21.70	0.77	0.500	0.57	0.36	0.98
	4	64	25.59	28.55	18	21.61	19.48	11	24.18	30.11	0.24	0.500	-0.14	0.13	-0.05
	5	89	19.87	24.03	7	33.57	27.57	15	12.07	15.99	2.53	0.500	0.57	-0.78	-0.32
	6	92	15.20	31.13	6	-6.33	20.87	16	10.06	28.65	0.23	0.500	-0.69	0.79	-0.16
	7	270	5.45	23.36	115	1.80	21.94	11	17.18	16.35	0.57	0.500	-0.16	0.70	0.50
	8	384	7.49	25.39	58	4.31	32.33	12	7.75	21.35	6.35	0.083	-0.13	0.11	0.01
	9	414	20.30	30.30	20	9.85	22.93	5	17.20	28.01	0.02	0.500	-0.35	0.32	-0.10
	10	445	-9.16	25.75	22	13.55	28.10	19	2.63	26.49	0.67	0.500	0.88	-0.39	0.46

Note. *p*-values reflect Linear Step-Up adjustments.

achievement gains than non-RP students at all grade levels on at least one reading achievement measurement (and at grades 5, 6, 7, and 8 significantly greater achievement gains were found on both tests). Effect sizes by grade level ranged from .03 to .34 (small to moderate in magnitude). None of the gain score comparisons of all students (Table 5.2) demonstrated significantly greater gain scores in favor of the non-RP students. Moreover, the trends in gain scores are worth noting. Students receiving the intermediate number of RP lessons (1 to 39) tended to have gains that were greater than students receiving no lessons but had gains that were less than students receiving 40 or more lessons. This suggests that the effects of the RP lessons are cumulative—more instruction using RP led to greater gains in reading achievement.

Tables 5.3 through 5.7 report FCAT Reading (CRT) Developmental Scale gain scores and SAT-10 (NRT) gain scores by grade level for students who were African American (Table 5.3), Latino American (Table 5.4), Caucasian (Table 5.5), LD (Table 5.6), and ELLs (Table 5.7). Aside from the ELL students, the data indicate that students receiving RP instruction made generally greater gains on the FCAT CRT and the NRT than students not receiving RP.

Table 5.8 presents statewide and district mean developmental scale scores for the CRT for grades 4 through 10 statewide and for the individual school district from which the RP schools were drawn. Mean gain scores for the statewide and district-level CRT are also presented. The mean gain scores for students engaged in the RP intervention for 40 or more lessons (Table 5.2) were greater than the statewide and district-level gains (Table 5.8) at every grade level for which a comparison was possible. Moreover, mean gain scores for students engaged in the RP intervention for 1 to 39 lessons (Table 5.2) also were greater than the statewide and district-level gains (Table 5.8) at every grade level except for grade 5.

Table 5.8: Dade County Reading CRT and Statewide Mean Development Scale Scores (DSS)

Grade	Mean 2006 DSS	Mean 2007 DSS	Mean DSS Gain
4	1554 (1573)	1393 (1420)	161 (154)
5	1618 (1659)	1537 (1557)	81 (101)
6	1644 (1694)	1583 (1624)	61 (70)
7	1773 (1801)	1694 (1722)	79 (78)
8	1814 (1862)	1730 (1786)	84 (76)
9	1851 (1912)	1789 (1844)	62 (68)
10	1881 (1947)	1864 (1931)	17 (16)

Note. Values in parentheses are statewide mean reading developmental scale scores.

Discussion

The present retrospective study examined the effects of a silent reading fluency and proficiency intervention system on the comprehension and overall reading achievement of students in grades 4 through 10 in a large urban school district. Results indicated that students participating in the program for a minimum of 40 lessons (20 hours of instruction) over approximately six months made significantly greater gains on both the reading CRT and NRT that were part of the FCAT than students who did not participate in the program. Students participating in the program also demonstrated gains on the CRT that were greater than the mean gains for the state and district level. The gains were found generally in all grade levels studied and in all subpopulations except for ELLs. Moreover, greater involvement in the RP intervention was associated with greater gains for students.

In many cases, the gains were not only statistically significant with substantive effect sizes, but also the contrasts between RP and non-RP groups provided interesting information regarding the magnitude of performance differences. For example, in grades 6, 7, and 8, the mean gains on the CRT portion of the FCAT were more than double the gains of nonparticipating students. For the same grade levels, gains on the NRT (SAT-10) by the RP intervention students were 55%, 82%, and 60%, respectively, greater than nonparticipating students.

Comments made by principals, teachers, and other educators in the schools that participated in the study were close to universally positive in support of the intervention system. Teachers and administrators using RP noticed the positive impact the program had on student achievement and attitudes toward learning.

The results of the study suggest that reading programs such as RP that are aimed at improving silent reading fluency and proficiency through extensive, focused, wide, and repeated reading in which students are held accountable for their work can have a significant and substantial positive effect on student reading comprehension and overall reading achievement.

Positive results were also demonstrated for various subpopulations that are often considered to be at risk for reading difficulties. African American, Latino American, special education, and learning disabled students who participated in the RP intervention generally demonstrated significantly and substantially greater gains in measures of reading achievement on both the CRT and NRT portions of the FCAT than students not participating in the intervention.

The only students who did not appear to benefit from the RP intervention were ELL students in grades 4, 5, 6, 7, and 8 (see Table 5.7). The best explanation for this lack of positive effects may lie in the fact that ELL students more than any other subpopulation of students are in the process of learning a new language, particularly the sounds of the language. Until the oral form of English becomes familiar and word decoding skills are mastered, ELL students may find oral reading where they hear and decode the written language into its oral form most beneficial. It is also worth noting that in the present study the sample size of ELL students was relatively small.

Aside from ELL students, however, the RP intervention, and, we assume, similar silent reading fluency and comprehension programs, hold great potential for significantly improving student reading achievement at a variety of grade levels.

The results of the study also suggest that although fluency is normally considered to fall within the domain of oral reading, silent reading fluency is a salient concept in reading. Moreover, the study further suggests that instruction aimed at improving silent reading fluency can have positive effects on reading achievement that are similar to those found with oral reading instruction, without some of the limitations that are associated with oral reading.

An additional finding from the study supports previous work by Rasinski and colleagues (Rasinski et al., 2005; Rasinski, Rikli, & Johnston, 2009) indicating that reading fluency is an important goal for reading instruction beyond the primary grades. In the previous work cited, Rasinski and his colleagues note that reading fluency continues to be an important predictor of reading achievement in the upper elementary through secondary grade levels and that significant numbers of students have not attained sufficient levels of fluency in their reading. The present study demonstrates that instruction in fluency, albeit silent reading fluency, for students beyond the primary grades can result in positive outcomes in reading comprehension and overall reading achievement. While current interest in reading seems to be shifting to helping middle and secondary school students improve their reading comprehension and achievement, the present study suggests that fluency-oriented instruction has great potential for making this goal a reality.

References

- Allington, R.L. (1983). Fluency: The neglected reading goal. *The Reading Teacher*, 36(6), 556–561.
- Allington, R.L. (2006). *What really matters for struggling readers: Designing research-based programs* (2nd ed.). Boston: Allyn & Bacon.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, 57(1), 289–300.
- Hill, C.J., Bloom, H.S., Black, A.R., & Lipsey, M.W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, 2(3), 172–177.
- Buck, J. & Torgesen, J. (2003). *The relationship between performance on a measure of oral reading fluency and performance on the Florida Comprehensive Assessment Test* (Technical Report #1). Tallahassee, FL: Florida Center for Reading Research.
- Chall, J.S. (1996). *Stages of reading development* (2nd ed.). Fort Worth, TX: Harcourt-Brace.
- 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(5), 386–406.
- Chinn, C.A., Waggoner, M.A., Anderson, R.C., Schommer, M., & Wilkinson, I.A.G. (1993). Situated actions during reading lessons: A microanalysis of oral reading error episodes. *American Educational Research Journal*, 30(2), 361–392.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cronbach, L.J., & Furby, L. (1970). How we should measure “change”: Or should we? *Psychological Bulletin*, 74(1), 68–80.
- Hiebert, E.H. (1983). An examination of ability grouping for reading instruction. *Reading Research Quarterly*, 18(2), 231–255.
- Huck, S.W., & McLean, R.A. (1975). Using a repeated measures ANOVA to analyze the data from a pretest-posttest design: A potentially confusing task. *Psychological Bulletin*, 82(4), 511–518.
- Kuhn, M.R., & Stahl, S.A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, 95(1), 3–21.
- Maxwell, S.E., & Delaney, H.D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). New York, NY: Taylor & Francis.
- 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. 00-4769). Washington, DC: U.S. Government Printing Office.
- O'Connor, R.E., Bell, K.M., Harty, K.R., Larkin, L.K., Sackor, S.M., & Zigmond, N. (2002). Teaching reading to poor readers in the intermediate grades: A comparison of text difficulty. *Journal of Educational Psychology*, 94(3), 474–485.
- Pikulski, J.J. (2006). Fluency: A developmental and language perspective. In S.J. Samuels & A.E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 70–93). Newark, DE: International Reading Association.
- 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(2), 119–134.
- Rasinski, T. (2006). Reading fluency instruction: Moving beyond accuracy, automaticity, and prosody. *The Reading Teacher*, 59(7), 704–706.

- Rasinski, T.V. (2010). *The fluent reader: Oral and silent reading strategies for building fluency, word recognition, and comprehension*. New York, NY: Scholastic.
- Rasinski, T.V., & Hoffman, J.V. (2003). Oral reading in the school literacy curriculum. *Reading Research Quarterly*, 38(4), 510–522.
- Rasinski, T.V., & Zutell, J.B. (1996). Is fluency yet a goal of the reading curriculum? In E.G. Sturtevant & W.M. Linek (Eds.) *Growing literacy: 18th yearbook of the College Reading Association* (pp. 237–246).
- Rasinski, T.V., Padak, N., Linek, W., & Sturtevant, E. (1994). Effects of fluency development on urban second-grade readers. *The Journal of Educational Research*, 87(3), 158–165.
- Rasinski, T.V., Padak, N.D., McKeon, C.A., Wilfong, L.G., Friedauer, J.A., & Heim, P. (2005). Is reading fluency a key for successful high school reading? *Journal of Adolescent and Adult Literacy*, 49(1), 22–27.
- Rasinski, T., Rikli, A., & Johnston, S. (2009). Reading fluency: More than automaticity? More than a concern for the primary grades? *Literacy Research and Instruction*, 48(4), 350–361.
- Rasinski, T., & Stevenson, B. (2005). The effects of Fast Start Reading: A fluency-based home involvement reading program, on the reading achievement of beginning readers. *Reading Psychology: An International Quarterly*, 26(2), 109–125.
- Reutzel, D.R., Jones, C.D., Fawson, P.C., & Smith, J.A. (2008). Scaffolded silent reading: A complement to guided repeated oral reading that works! *The Reading Teacher*, 62(3), 194–207.
- Roehrig, A.D., Petscher, Y., Nettles, S.M., Hudson, R.F., & Torgesen, J.K. (2008). Accuracy of the DIBELS oral reading fluency measure for predicting third grade reading comprehension outcomes. *Journal of School Psychology*, 46(3), 343–366.
- Rogosa, D. (1995). Myths and methods: “Myths about longitudinal research” plus supplemental questions. In J.M. Gottman (Ed.), *The analysis of change* (pp. 3–66). Mahwah, NJ: Lawrence Erlbaum Associates.
- Schatschneider, C., Buck, J., Torgesen, J.K., Wagner, R.K., Hassler, L., Hecht, S., & Powell-Smith, K. (2004). *A multivariate study of factors that contribute to individual differences in performance on the Florida Comprehensive Reading Assessment Test: A Preliminary Report* (Technical Report #5). Tallahassee, FL: Florida Center for Reading Research.
- Stahl, S.A., & Heubach, K.M. (2005). Fluency-oriented reading instruction. *Journal of Literacy Research*, 37(1), 25–60.
- Torgesen, J.K., & Hudson, R.F. (2006). Reading fluency: Critical issues for struggling readers. In S.J. Samuels & A.E. Farstrup (Eds.) *What research has to say about fluency instruction* (pp. 130–158). Newark, DE: International Reading Association.
- Torgesen, J., Nettles, S., Howard, P., & Winterbottom, R. (2004). *Brief report of a study to investigate the relationship between several brief measures of reading fluency and performance on the Florida Comprehensive Assessment Test-Reading in 4th, 6th, 8th, and 10th grades* (Technical Report #6). Tallahassee, FL: Florida Center for Reading Research.

CHAPTER 6

Exploring the Added Value of a Guided Silent Reading Intervention: Effects on Struggling Third-Grade Readers' Achievement¹

D. Ray Reutzel
Utah State University

Yaacov Petscher
Florida Center for Reading Research &
Florida State University

Alexandra N. Spichtig
Reading Plus

Reading research has produced an emerging consensus on several essential elements of beginning reading instruction, and fluency is widely agreed to be one of the key components, as reading fluency creates the bridge between word recognition and reading comprehension processes (National Institute of Child Health, & Human Development (NICHD), 2000; Rasinski, 1989; Reutzel & Hollingsworth, 1993; Samuels & Farstrup, 2006). The initial stages of reading fluency occur when students are able to automatically recognize words. As fluency develops, automatic word recognition eventually leads to the achievement of the ultimate goal of reading: comprehension (Torgesen & Hudson, 2006; Samuels, 2007). Topping (2006) described this later stage of fluency development, when word recognition bridges comprehension processes, as “the extraction of maximum meaning at maximum speed in a relatively continuous flow, leaving spare, simultaneous processing capacity for other higher order processes” (p. 107).

In its report, the National Reading Panel (NRP; NICHD, 2000) reviewed 77 studies of guided repeated oral reading (GROR)

¹ This chapter was previously published in *The Journal of Educational Research* (v105, n6, p404-415, 2012). The definitive publisher-authenticated version published is available online at: <http://www.tandfonline.com/doi/abs/10.1080/00220671.2011.629693>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

with feedback. This approach to reading fluency practice includes oral repeated readings of a single grade- or instructional-level text, while the reader simultaneously receives feedback from a teacher or from other, more proficient readers. The NRP found substantial scientific evidence to support the efficacy of GROR with feedback for increasing students' reading fluency.

However, the NRP report (NICHD, 2000) sparked considerable controversy when it reported a lack of research supporting independent silent reading practice in the classroom as an effective means for developing students' reading fluency, referring to the types of silent reading that are included in activities like Sustained Silent Reading (SSR) or Drop Everything and Read (Allington, 2002; Coles, 2000; Cunningham, 2001; Edmondson & Shannon, 2002; Krashen, 2002). Thus, one of the unintended consequences of the NRP's report was to suppress the previously prevalent use of silent independent reading practice in the classroom to develop students' reading fluency. Although silent reading practices such as SSR had been already generally criticized, sharp critiques of independent silent reading increased significantly after the report was released in 2000. In today's world of high-stakes accountability, it is increasingly difficult for classroom teachers to justify the use of instructional practices that do not have the imprimatur of the NRP or are not sanctioned as evidence-based by the federal government.

Consequently, since the turn of the new millennia, GROR with feedback has become the dominant way in which teachers encourage students to practice their reading in classrooms in order to develop fluency. Yet once GROR with feedback became the dominant method of classroom instruction used to develop fluency, it became more and more apparent to teachers, administrators, and researchers that this particular mode of reading instruction used in school for fluency practice—oral and guided—was unrelated to the most common way in which most accomplished adolescent and adult readers *actually* read—independently and silently. Although guided repeated oral reading may be useful as a tool for reading fluency practice in school, the long-term goal of any fluency development should be, in fact, to help students become avid, competent, and independent silent readers. There is little question that the opportunity to read is strongly associated with gains in students' reading achievement, regardless of whether that reading is conducted silently or orally (Allington, 2002; Guthrie, Schafer, & Huang, 2001; NICHD, 2000). Also, in addition to the strong correlation of evidence between time spent reading and reading achievement, causal evidence for the efficacy of engaging

students in independent silent reading practice has been steadily growing.

In a large-scale experiment that provided students with an additional 20 minutes of independent silent reading of trade books per day, Block, Cleveland, and Reed (2006) found that the additional practice led to significant annual gains in students' vocabulary, comprehension, and fluency. Similarly, Samuels and Wu (2004) reported results of a quasi-experiment that compared the reading comprehension and reading achievement gains of third- and fifth-grade students who read independently and silently for either 15 or 40 minutes. Students in the 40-minute group evidenced significantly better reading achievement and comprehension than the group that read for 15 minutes. These and other studies over the past decade (e.g., Hiebert & Reutzel, 2010; Kamil, 2008; Kelley & Clausen-Grace, 2006, 2010; Kuhn & Schwanenflugel, 2008; Reutzel, Fawson, & Smith, 2008; Reutzel, Jones, Fawson, & Smith, 2008; Reutzel, Jones, & Newman, 2010) have begun to provide a causal research base showing that students' reading comprehension, fluency, and achievement can benefit from the opportunity to read independently and silently when specific conditions of reading fluency practice are implemented. It also appears that using independent silent reading as the means to practice fluency makes more sense both developmentally and empirically for students who are older than age 8, or who are at least in second grade (Wright, Sherman, & Jones, 2004, 2010; Kuhn, 2004, 2005).

Today's educational culture of increased accountability has compounded teacher concerns regarding the use of independent silent reading to practice fluency, especially when this method is used with low-achieving, struggling students. A persistent fear among classroom teachers is that some students may not keep their eyes on their text when they are assigned silent independent reading tasks (Donovan, Smolkin, & Lomax, 2000; Fresch, 1995; Hiebert, Wilson, & Trainin, 2010). Guidance within silent reading contexts is key, as students achieving in the bottom quartile of their class frequently attend less well when they read silently in an unguided context as compared to a guided context (Hiebert et al., 2010). As reported in several studies, when the challenge level of texts and the task of reading independently and silently were carefully scaffolded and guided by the teacher, even primary-level, struggling readers were able to engage in the practice successfully (Bryan, Fawson, & Reutzel, 2003; Kamil, 2008; Kelley & Clausen-Grace, 2006, 2010; Reutzel, Fawson, & Smith, 2008; Reutzel, Jones, Fawson, & Smith, 2008; Reutzel, Jones, & Newman, 2010).

In addition, Brenner and Hiebert (2010) recently reported research related to a professional development program intended to help teachers

increase the amount of time students' eyes were on text during silent reading. These researchers, among others who have recently examined the independent silent reading process, discovered that previous explanations of the "eyes on text" phenomenon had seemingly overlooked a fundamental contributor to that process—*the eyes* (Samuels, Hiebert, & Rasinski, Chapter 2, this volume).

Rasinski, Samuels, Hiebert, Petscher, and Feller (Chapter 5, this volume) reported on research conducted with students in grades 4 through 10 using the computer-based guided silent reading fluency intervention Reading Plus, the same intervention used in this study. Their study, conducted with students in grades 4 through 10, showed a strong relationship between the use of the Reading Plus program for silent reading practice and subsequent gains in reading comprehension. Use of the program was also shown to improve general literacy achievement on a state criterion-referenced and normative-referenced national reading achievement test.

Although these researchers found a relationship between students' guided silent reading practice with the Reading Plus program and the students' gains in reading comprehension and general reading achievement, much less is known about how such a supplementary intervention program may influence the reading behaviors and achievement of younger struggling readers. In today's environment of high accountability, there is a need for carefully constructed evaluations of commercially available supplementary intervention programs by credible organizations such as the U.S. Department of Education's What Works Clearinghouse. Thus, studies such as the one reported by Rasinski et al. (Chapter 5, this volume) are of evaluative and practical importance for classroom teachers and administrators who are seeking guidance and evidence to support the selection and use of available reading interventions for challenged readers.

Although previous studies have shown silent reading to be an effective way to improve reading skills, more recent studies have shown that the conditions for silent reading practice in school often result in students acting like they are reading when they are not (Hiebert & Reutzel, 2010). As a result, this study sought to determine whether a computer-based guided silent reading fluency intervention, using a combination of scaffolded reading passages and comprehension questions, could reliably increase struggling students' reading achievement and comprehension by helping these readers keep their eyes on the text during silent reading. Specifically, the research question addressed by this study was the

following: How does the guided silent reading fluency intervention Reading Plus affect struggling third-grade students' performance on criterion-referenced and norm-referenced tests of reading comprehension and reading achievement?

Method

Research Design

This study used a matched quasi-experimental research design. The study's quasi-experimental control and treatment groups were constructed by the use of a propensity score sampling and matching process. A propensity score, as defined by Rosenbaum and Rubin (1983), has a conditional probability of assignment to a particular treatment given a vector of observed covariates. Put simply, a propensity score is the probability of being in the treatment group derived from a logistic regression when accounting for important matching variables. The primary objective for researchers using propensity scores is to select a series of variables that would be considered important for matching students. In traditional reading research, these variables might include race/ethnicity, socioeconomic status, English learner status, primary exceptionality status, gender, and some type of baseline measure of achievement (e.g., a pretest). The main effects and interactions among these and other variables are then included in a logistic regression to determine the probability of being in the treatment when controlling for these important matching covariates (Shadish, Cook, & Campbell, 2002). The probabilities resulting from the logistic regression may then be used to match students who actually received the intervention with those who did not, creating matched treatment and control groups. In this way, students are more probabilistically matched at the pretest, allowing for stronger causal inferences regarding differences on the posttest or on gain scores than a simple comparison of all available students in a sample.

There are several limitations that should be noted in regard to using propensity scores to construct an experimental sample such as the one used in this study: (a) propensity scores tend to be most practically used with larger samples; (b) missing data can be problematic for propensity analyses, as the techniques are still relatively new; and (c) propensity scores assume that no further confounds exist that may predict the propensity. Nevertheless, despite these acknowledged limitations, propensity scores are now viewed as one of the strongest quasi-experimental methods for assessing relationships between treatments and outcomes (Shadish et al., 2002).

Participants

Three criteria were used to select students for the control and treatment group samples. First, students who were selected were age 9 or older. Second, students were also selected if they were identified as being struggling readers if the results from their end-of-year, high-stakes third-grade achievement test (the FCAT) identified that they were at risk for not being promoted to fourth grade. Finally, retained third-grade students who were re-enrolled in third-grade classrooms during the implementation period, which ran from the beginning of the school year through the administration of the FCAT and SAT-10 in early March, were selected.

We selected struggling third-grade readers for participation in this study for two reasons. First, age 9 (or third grade) has been shown to be an age and stage of reading development where independent silent reading becomes both possible and advisable based upon recent research findings (Hiebert & Reutzel, 2010). Second, the stage of reading development that occurs as third graders transition into fourth grade has been long associated with what Chall, Jacobs, and Baldwin (1990) described as the fourth-grade reading slump.

The archival full student sample file available to researchers consisted of records on 1,253 third-grade students enrolled in a large, urban public school system in the state of Florida. Of these 1,253 students, 158 represented the required special case of students at age 9 retained at the end of third grade. Thus, all 158 third-grade students in the study's sample population were not promoted to fourth grade due to their performance on the FCAT. These 158 retained third-grade struggling readers attended 11 different elementary schools within this Florida urban school district. The final propensity score sample constructed for this study's matched, quasi-experimental research design consisted of 40 students in the control and 40 students in the experimental treatment group, for a total of $n=80$ retained third-grade students.

Instrumentation

At the time of this study, the FCAT was a major component of Florida's testing effort to assess student achievement in reading, writing, math, and science as represented in Florida's Sunshine State Standards (SSS) (Florida Department of Education [FDE], 2007). The SSS reading portion of the FCAT is a group-administered, criterion-referenced test consisting of 6 to 8 narrative or informational reading passages, wherein students respond to between 6 and 11 multiple-choice items per passage.

Embedded within these 6 to 11 multiple-choice questions are four content clusters: (a) reference and research, (b) words and phrases in context, (c) the main idea, and (d) comparison/cause and effect.

Based on their scores, students are placed into one of five performance levels on a scaled score ranging from 100 to 500. Levels 1 and 2 reflect below grade-level performance in reading, with Level 1 being the lowest indication of reading performance. Levels 3 and above represent proficiency in reading comprehension at or above grade-level standards.

Students who score below Level 1 proficiency on the FCAT in third grade must be retained for another year, according to Florida law. If they can demonstrate the required reading level or proficiency through the approved alternate test (the SAT-10) or through a student portfolio, they can be granted an exemption and be promoted to fourth grade. Thus, the students selected for this study represented the highest-risk segment of the overall third-grade population. The internal-consistency reliability for the FCAT-SSS has been shown to be 0.90 (Cronbach's alpha); moreover, test score content and concurrent validity have been established through a series of expert panel reviews and data analyses (Florida State Department of Education, 2007). The construct validity of the FCAT-SSS as a comprehensive assessment of reading outcomes recently received strong support in an empirical analysis of its relationships with a variety of other reading comprehension, language, and basic reading measures (Schatschneider, Fletcher, Francis, Carlson, & Foorman, 2004).

The SAT-10 is approved for use by the U.S. Department of Education and is constructed to determine if students in kindergarten through grade 12 are meeting national or state standards in reading, mathematics, and language. The reading section of the SAT-10 has an alpha reliability coefficient of 0.87, the math section 0.80 to 0.87, and the language section 0.78 to 0.84. Alternate forms of reliability coefficients ranged in the low 0.90s for the total reading section. The SAT-10, by design, evidences content and criterion-related validity since its development is tied very closely to assessing progress toward meeting state and national standards in reading, mathematics, and language (Berk, 1998; Carney & Morse, 2005).

Control and Treatment Groups

All 80 retained third-grade students in both the control (n=40) and treatment (n=40) groups followed the state-approved Comprehensive Core Reading Program (CCRP) adopted by this large, urban Florida

school district. The CCRP delineated specific protocols unique to the third-grade retained students requiring schools to provide a dedicated and uninterrupted two-hour block of classroom instructional time for reading instruction for all students. Whole-group explicit reading instruction was provided daily for the first 30 to 40 minutes using Houghton Mifflin's Reading Treasures comprehensive core reading program. Thirty minutes of the 2-hour block were dedicated to writing instruction. For the remainder of the 2-hour reading instructional block, teachers differentiated instruction using small groups and center rotations, during which time students practiced, demonstrated, and extended skills that were previously taught during the teacher-led explicit reading instruction. Approved supplemental reading intervention programs could be used at this time. Some of these included QuickReads (repeated oral readings of the same passage), Elements of Reading: Vocabulary (an oral vocabulary instructional program), and the supplemental activities provided with Houghton Mifflin's Reading Treasures. Retained students were required to receive intensive intervention in areas of their demonstrated deficiencies during the mandated 2-hour reading instructional block.

In addition to the dedicated 2-hour block instructional time that all control and treatment students received, all of the participating students received an additional 30 minutes of supplemental reading instruction every day. Supplemental reading programs included: Soar to Success, Essential Elements of Reading and Voyager Passport (Essential Elements), Earobics, and Reading Plus. Treatment group students engaged in Reading Plus, while the control group used one of the other three supplementary reading interventions. The alternative interventions differed from Reading Plus as well as from one another in their curricular emphasis. According to program developers, Soar to Success provides instruction in four of the essential components of reading outlined by the NRP report (NICHD, 2000) and Reading First (phonics, fluency, vocabulary, and comprehension), whereas Essential Elements and Earobics address five of the components (phonemic awareness, phonics, fluency, vocabulary, and comprehension). Reading Plus focuses on three of the components: fluency, vocabulary, and comprehension. Another essential difference between Reading Plus and the alternative interventions is the mode of the reading experience. Within Reading Plus, the emphasis is on guided, silent reading, while the three alternative interventions emphasize guided oral reading.

Reading Acceleration Programs (Control Group)

Selected struggling third-grade students (n=40) received one of three accelerated reading treatments during three weekly 30-minute sessions. The key elements of each of these interventions are summarized below, drawing on a report from the Florida Center for Reading Research (2007).

- Soar to Success is designed to accelerate students' reading ability and help them quickly and easily apply comprehension and decoding strategies to other content area texts through the use of reciprocal teaching, an instructional technique that uses teacher–student dialogue to teach students to use cognitive strategies of summarizing, clarifying, questioning, and predicting. Each 30-minute lesson consists of five parts: (a) Revisiting (students reread self-selected Soar to Success books for fluency development), (b) Reviewing (students review strategies and summarize what was read using graphic organizers), (c) Rehearsing (a teacher-guided or independent preview of the daily reading is presented), (d) Reading and reciprocal teaching—students read silently and then engage in four reciprocal teaching strategies (summarizing, clarification, questioning, and predicting), and (e) Responding/reflecting—students complete written reflections and engage in discussions to bring closure to the daily activity.
- The Essential Elements program is designed to accelerate reading growth and assist students in reaching grade-level expectations through the use of teacher modeling, guided and independent practice, and immediate corrective feedback. The program consists of daily lessons that are taught in small groups. A typical 30-minute lesson for third-grade students consists of advanced vocabulary word analysis, fluency-building passage reading, and comprehension strategies. Third-grade students in need of additional support in word study may engage in an optional Targeted Word Study component.
- Earobics is designed to help striving readers develop foundational skills through the use of software, teacher-directed activities, manipulatives and books. The program consists of two parts: Part one is designed for first- and second-grade students, and part two is designed for second-grade students and older who are struggling with fluency. Students may engage in software games that target phonemic awareness and phonics skills, or teachers may provide

explicit instruction in language enrichment, phonemic awareness, letter-sound correspondences, decoding, and early reading and writing.

Guided Silent Reading Fluency Intervention (Treatment Group)

An equal number of selected struggling third-grade students (n=40) received the comparison treatment: Reading Plus, a supplementary guided silent reading intervention. Students involved in this guided silent reading intervention participated in a series of online, computer-based sessions that included a specific sequence of daily activities. As struggling students participated in this guided silent reading intervention, the difficulty level of the reading material was adjusted automatically as a function of a student's progress based upon reading comprehension and reading rate analyses.

Students began the intervention by completing the Reading Placement Appraisal (RPA) assessment to establish their initial placement level within the supplementary guided silent reading intervention program. This 20-minute placement test assessed independent reading level, reading rate, comprehension, and vocabulary to determine the most appropriate practice starting level. The RPA consisted of three parts. Part I presented students with several 100-word selections, each followed by a set of literal-recall questions. Content difficulty was adjusted according to a student's comprehension performance and reading rate mastery to ascertain a student's tentative independent reading level. Part II, with its 300-word selections and diverse comprehension questions, served to confirm the independent reading level. Part III assessed a student's vocabulary level. From the three-part RPA assessment, an instructional reading level was established for individual students, who were then placed at appropriate levels within each component of the program. Students continued to be assessed on similar tasks throughout the intervention period, with appropriate adjustments being made to the level of the reading selections as a result of students' performance on these formative assessments. As students participated in this supplementary silent reading fluency intervention, they were provided with reading lessons and continuous feedback about their silent reading in an individual, computer-based, online environment.

Each lesson began with a perceptual accuracy and visual efficiency warm-up. This activity consisted of two parts, Scan and Flash. In the Scan activity, students scanned the computer screen to count the number of

times a target letter or number appeared on the screen. The target letter and other letters or numbers were flashed in a left-to-right presentation. The presentation speed increased in accordance with the student's proficiency. In the second activity, Flash, a series of letters or numbers—ranging in length from 2 to 12 characters, depending on the students' placement level—were flashed (1/6 of a second per flash, which does not permit moving of the eyes and thus provides single fixation training). The amount of numbers or letters increased in response to the students' ability to correctly recreate the sequence. This warm-up activity aimed to increase students' visual perception, attention, and automaticity in the discrimination and recognition of print. Studies conducted by numerous researchers (e.g., Brenner & Hiebert, 2010; Mirsky, 1996; Torgesen & Hudson, 2006) suggested that one of the defining characteristics of proficient readers is the ability to sustain attention and keep their eyes on the text. According to Pikulski (2006), "instant, accurate, and automatic access to all these dimensions of a printed word is the needed elements of fluency that will allow readers to focus their attention on comprehension rather than on decoding" (p. 75).

The next part of the guided silent reading session provided students with extensive structured silent reading practice to build fluency within an authentic reading experience where students read for meaning. This activity involved timed, guided, left-to-right reading practice, in which students read texts selected from a diverse collection of narrative and expository texts at each student's independent instructional reading level. This is noteworthy because the work of O'Connor and colleagues (2002), as reported by Allington (2006), showed that providing daily intervention lessons using grade-level texts was not nearly as successful as providing daily lessons using texts matched to the instructional reading levels of struggling readers. O'Connor and colleagues argued that selecting texts of appropriate challenge should be a first step in the design of effective supplementary reading instruction and intervention. In fact, this is no less true when designing effective silent reading practice for regular education students in elementary classrooms (Reutzel, Jones, Fawson, & Smith, 2008).

Lesson text selections were matched to struggling readers' independent reading levels using Spache, Dale-Chall, and Fry readability formulas. The supplementary guided silent reading intervention computer environment was programmed to continuously and dynamically monitor students' performance using both reading rate measures and responses to comprehension questions, adjusting the reading content level to match each student's progress. In addition, the guided silent reading

intervention program used a mix of instructional formats and scaffolds to further match individualized needs and rates of progress. These included variation in the presentation of text, the length of reading segments, the location and number of comprehension questions, and the use of repeated readings. Thus, students were able to progress through increasingly challenging levels of readings in this intervention based on several factors. Students had to be able to read passages at their current levels with grade-appropriate rates and good comprehension before they were advanced to subsequent levels.

This supplementary guided silent reading intervention provided approximately 600 reading selections, ranging from pre-primer to adult-level texts, including high-interest/low-readability selections for older struggling students. Selections represented a wide range of genres, such as “Miguel’s Big Day,” a family life story; “The Lighthouse Visitor,” a mystery; and “Looking at Clouds,” a science/nature story. As students progressed through the varied guided silent reading levels, the texts became longer and more challenging, and content choices became more informational. Lesson texts were presented within both a guided silent reading format (a moving window guided students’ eyes across lines of print from left to right) and an independent reading format without any left-to-right guidance. Regardless of the nature of the lesson or activity, text was presented within a controlled format and rate parameter for each student in the online environment. Dynamically controlled by individual student performance, comprehension questions were either interspersed among individual reading segments or were found at the conclusion of the story. All comprehension questions were electronically coded by the system to continuously track student performance with 25 comprehension skills based on Bloom’s (1956) taxonomy, including literal understanding, interpretation, analysis, evaluation, and appreciation. The format (wide vs. repeated readings) and rate at which text was presented on screen was then incrementally increased as a function of students’ performance on these comprehension questions and reading rate performances during the reading events.

The lessons provided students with authentic reading experiences that build comprehension, fluency, and stamina at a level of difficulty that supports an increase in reading capacity. Additionally, given that the difficulty of texts was established using the Spache (for primary-level texts) and Dale-Chall (middle-grade-level texts) formulas—both of which rely on high-frequency word lists—students had considerable opportunity to develop fluency with a core group of high-frequency words while reading

these texts. This is important because Torgesen and colleagues (Rashotte, MacPhee, & Torgesen, 2001; Torgesen & Hudson, 2006) argue that limited sight vocabularies are a principal characteristic of students with reading disabilities beyond the initial phase of learning to read.

The guided silent reading component of the intervention was also followed by the cloze vocabulary component, which used structured contextual analysis activities to assist struggling students in developing comprehension competency. These cloze exercises were intended to encourage students to use context clues to complete the meaning of sentences as well as longer passages. Students also practiced deriving the meanings of difficult or unfamiliar words by analyzing the surrounding context in these cloze activities, thus potentially enhancing wide-reading vocabulary-learning strategies and skills.

Several factors informed just-in-time instructional decisions that were sensitive to student characteristics, such as age, reading level, performance, progress, and instructional trajectory: performance scores within each practice module, the interconnectedness of the various practice modules, integrated formative assessments following each lesson, and a highly sophisticated operating system. The system not only dynamically adjusted each student's differentiated lesson format within each practice module, but it also provided unique adjustments for daily practice sessions. The integration of these modules allowed the system to provide each student with a practice environment that uniquely addressed his or her individual silent reading development needs at any moment in time during the implementation period.

Data Analysis

In order to assess the added value of this silent reading fluency intervention with third-grade struggling readers, a propensity score analysis was used in this study to match the 40 students from the sample of 158 who did not receive this supplementary silent reading fluency intervention to a group of 40 students who were similar with regard to demographics, prior FCAT achievement, and performance on the SAT-10. The 40 struggling students completed an average of 71 lesson units during the study. The logistic regression used in this study to construct the propensity scores predicted group membership with race/ethnicity, limited English proficiency status, primary exceptionality status, and reading performance on the previous year's FCAT-SSS and the SAT-10. Prior technical reports have indicated that the correlation between FCAT

scores from year to year is approximately 0.75 in elementary schools (FDE, 2005); moreover, the correlation between the FCAT-SSS and the SAT-10 in grade 3 is 0.78. As such, while a strong correlation exists between the two assessments of reading comprehension, it was important to capture the unexplained covariance in scores. By using both measures in the propensity score matching, greater specificity could be attained. Resulting propensity scores were used in a secondary analysis to match students based on their designation as having received treatment or not. Once students were appropriately matched, they were designated to receive or not receive the supplementary silent reading fluency intervention (control vs. treatment group). After a full year, the resulting student scores were analyzed using a one-way analysis of variance (ANOVA) with a linear step-up to control for the false-discovery rate (FDR) (Benjamini & Hochberg, 1995).

Results

Summaries of the demographics and descriptive statistics for the FCAT and SAT-10 scores for the treatment and matched controls groups are reported in Tables 6.1 and 6.2. As can be seen by the reported indices, the two groups were reasonably matched from the propensity analysis. The mean pretest score for the matched control group on the FCAT-SSS was 814.90 (SD=217.92) compared to the treatment groups' mean of 845.50 (SD=117.69), corresponding to a standardized coefficient of $g=0.17$. Similarly, the mean pretest score on the SAT-10 for the treatment group was 575.75 (SD=16.04), compared to the control groups' mean of 570.73 (SD=18.90), and corresponded to a standardized coefficient of $g=0.28$. Because students who participated in the program were from different classes and schools, and the analysis was based on available archival data, the ratio of students to classes was small, precluding a mixed-effects modeling of the data to account for clustering at the classroom and school levels.

Table 6.1: Demographic Comparison of Treatment and Matched Control Students

Demographics	Treatment (n=40)	Control (n=40)
% Black	65	58
% White	0	8
% Latino	35	35
% ELL	15	13
% ESE	5	25

Table 6.2: Descriptive Statistics for the Control and Treatment Groups on the FCAT and SAT-10 Pre- and Posttest Scores

Measure	Control Group		Treatment Group	
	Mean	SD	Mean	SD
FCAT Pre	845.50	117.69	819.90	217.92
FCAT Post	1,012.33	357.46	1,322.63	171.24
SAT-10 Pre	570.03	18.90	575.75	16.04
SAT-10 Post	597.93	34.95	608.53	24.43

ANOVA was used to assess the extent to which the treatment and matched control students were statistically differentiated on the posttest scores for both the FCAT-SSS and the SAT-10. In order to control for the FDR, a linear step-up procedure was used for any statistically significant finding.

FCAT-SSS results indicated that a significant effect existed for treatment ($F [1,79]=24.52, p < 0.001$), suggesting that treatment students' scores on the posttest were significantly higher than the matched control. The mean posttest score for the silent reading fluency intervention students was 1,322.63 ($SD=171.24$) compared to the matched control's mean of 1,012.33 ($SD=357.46$). A more appropriate way to contextualize these results is to calculate an effect size, which communicates, in standard deviation units, how large the differences were between the means of the two groups, regardless of sample size. A standardized effect size value $g=1.09$ was estimated, indicating that the mean for the students who were receiving the silent reading fluency intervention were performing one full standard deviation above the mean for the matched controls. In context, Cohen (1988) provided guidelines stating that an effect size of 0.80 would be considered large. In practical terms, 80% of the treatment students who received the supplementary guided, silent reading fluency intervention in this study achieved reading proficiency as measured by the FCAT (achievement level of 3 or higher) and were promoted to the next grade level, as compared to 32% of the of the matched control students.

Conversely, no statistically important findings were observed for the SAT-10 differences in the ANOVA ($F [1,79]=2.59, p=0.11$), despite a higher posttest SAT-10 score for students receiving the supplementary guided, silent reading intervention ($M=608.53, SD=23.43$) compared to the matched controls ($M=597.83, SD=34.95$). Two important components to consider in these seemingly conflicting findings are the issues of power and baseline equivalence. Given the present total sample size in the design ($n=80$), a potential reason for the lack of statistical significance in

the SAT-10 analysis is due to a small effect size that could be observed, and not a sufficient sample size to detect it. Indeed, a power analysis with $n=80$, $\alpha=0.05$, and $\text{power}=0.80$ indicates that the minimum detectable effect size would be 0.63. As such, in the case of the FCAT, a statistically significant finding was observed with an estimated effect size difference of over 1.0. With the SAT-10 data, a quick calculation of the posttest mean differences would yield a standardized coefficient of 0.30, yet with minimum detectable effect size of only 0.67, it would not be possible to obtain a statistically meaningful finding with this group. This does not imply, however, that if the sample size were larger or the baseline effect were smaller, a statistically significant effect would be obtained, as the pretest differences suggest that a more diverse sample could be used to provide a more accurate match. Notwithstanding this limitation, these results represent preliminary evidence that a moderate to strong relationship between the added value of the guided silent reading fluency intervention (Reading Plus) and student performance in reading exists for retained third-grade students in Florida, given the measured outcomes.

Discussion

Providing the highest quality of reading instruction for all students is a central focus of current educational reforms and practices. Such an emphasis is particularly critical in the era of the Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Much of the past research on silent reading has focused upon comparing the results obtained from silent independent reading versus oral, guided reading practice, and such studies have typically found that guided oral reading practice is more effective for students and is also preferred by teachers. However, we believe this is largely due to the fact that guided oral reading provides a check on whether students are actually reading and how well they do so when that check is not possible with silent reading conducted by students independently. Therefore, prior to the turn of the millennia, these studies comparing guided, oral reading versus independent silent reading practice contributed little to an understanding of how silent reading practice might become more effective.

Instead of providing yet one more comparison of independent silent reading versus a largely guided oral approach to reading practice, this study examined how changing silent reading practice conditions from silent, independent reading to silent, *guided* reading affected the reading

comprehension and reading achievement of struggling third-grade readers. This study compared the reading comprehension and achievement of third-grade struggling readers who received guided practice during silent reading using Reading Plus with a group of matched control students who received a combination of other school district-approved supplementary reading interventions—Early Success, Soar to Success, Earobics, and a combined program of Essential Elements of Reading: Vocabulary and Voyager Passport—using largely guided oral reading practice. The goal of using the treatment intervention was to provide struggling third-grade students with sufficient guidance, intensity, consistency, and appropriateness of silent reading practice in an online, computer-based environment to substantially increase their reading comprehension and reading rate achievement. Guided silent reading practice as provided in the treatment intervention was continuously adapted for format of reading practice (repeated vs. wide readings, short or long passages), pace and level of structure during reading (guided vs. independent reading), and level of reading challenge (readability and genre types) via the use of reading efficiency measures and comprehension assessments during online silent reading practice with leveled texts. Using guided silent reading practice not only frees the teacher to provide more instruction and assistance to targeted students during reading practice sessions but also assures that when the teacher is not present, struggling readers who read silently are on task and have their eyes on the page.

Statistically significant differences were identified in favor of the guided silent reading treatment group on struggling third graders' reading comprehension and reading achievement scores on the FCAT test. The effect sizes were large, slightly greater than a full standard deviation, favoring the supplementary guided silent reading intervention as compared to other school district-approved supplementary reading interventions for use with these retained third-grade struggling readers.

The statistically significant findings and large effect sizes favoring the guided silent reading practice provided to struggling third-grade readers can be at least partially explained by turning to other research on effective approaches for providing silent reading practice to students in schools. First, one possible reason this type of guided silent reading intervention was successful because the intervention increased this sample of third-grade struggling readers' opportunities to read. It did so in a number of ways. Second, in past research on silent reading, struggling readers often selected books that were too difficult for them to read fluently. The supplementary guided silent reading intervention's

computer environment monitored students' comprehension of texts and then automatically and continuously adjusted the format of practice, genre, and level of challenge to match the students' abilities to comprehend the texts they were reading silently. This is essential to success because when students cannot read the texts they have selected for silent reading, they do not read much. In the title of a classic article, Allington (1977) reminded us that "if they don't read much, how are they ever gonna get good?" When students do not read much during silent or oral reading practice time, they do not benefit in terms of achievement from the time allocated. Further, when struggling readers cannot or do not read silently, they find it difficult to keep their eyes on the text and focus their attention (Hiebert et al., 2010). The guided silent reading intervention used in this study assured that students' eyes were on the text by providing visual and perceptual modeling practice, monitoring their comprehension responses to the reading of increasingly challenging and longer text selections, and continuously adjusting the level of text and question challenge based upon these indicators.

The supplementary guided silent reading intervention used in this study also promoted student motivation because students were provided with a selection of appropriately leveled texts from which they could choose stories that most interested them (Fawson, Reutzel, Read, Smith, & Moore, 2009; Swan, Coddington, & Guthrie, 2010). Studies have shown that students who read widely, as was the case in this intervention, learn more vocabulary word meanings through their reading and increase their abilities to manage and comprehend a variety of text structures and genres (Cunningham & Allington, 2010; Pressley, 2002). In this study, providing struggling students with continuous feedback on their reading performance in terms of rates and comprehension was also helpful to students as a part of designing effective guided silent reading practice conditions. Adjusting passage and lesson difficulty also seemed to help struggling students make significant progress. Holding students accountable for their time spent reading by measuring students' reading rates as well as their responses to comprehension questions and cloze passages lets them know they are going to be monitored for the time spent in reading practice. In our study, accountability assured that students' eyes were on the text, which has been shown to predict students' reading achievement (Brenner & Hiebert, 2010). Finally, the supplementary guided silent reading intervention tested in this study focused more time and practice on developing students' fluency, vocabulary, and comprehension skills than did the control programs that gave considerable time and

practice to increasing students' word-recognition automaticity through decoding practice. Focusing students' practice on fluency, vocabulary, and comprehension may have transported students more efficiently over the fluency bridge from decoding to comprehension than did more decoding practice.

No significant differences were found between the control and treatment groups of retained struggling third-grade readers on the SAT-10 nationally norm-referenced reading achievement text and reading comprehension subtest. With respect to the SAT-10 findings, locating a sample size of retained third-grade struggling readers sufficiently large enough to power the analysis of SAT-10 reading scores proved to be daunting, even with an initial sample of more than 1,200 third-grade students. A post hoc power analysis of the sample size for this study, $n=80$, determined that the obtained sample size was too small to provide sufficient statistical power to detect a difference in third-grade students' SAT-10 reading comprehension and achievement scores. As a result, we cannot be sure that this guided silent reading intervention was any more or less effective than other supplementary reading interventions provided to this sample of struggling third-grade readers as measured by the SAT-10, although posttest SAT-10 reading comprehension mean scores trended higher for the supplementary guided silent reading intervention treatment group than for the control group. However, these results can also be used to argue that the guided silent reading intervention used in this study was at least as useful as were the other school district-approved supplementary reading interventions provided to this group of struggling readers.

As a result, the evidence presented in this study demonstrates that providing struggling third-grade readers with a guided silent reading intervention in an online, computer-based environment via the Reading Plus supplementary reading intervention yielded large effects on reading achievement and comprehension scores on a high-stakes, state-administered test, the FCAT, which is used by Florida schools to determine both individual student progress and school progress toward meeting the requirements of adequate yearly progress (AYP).

Limitations

The results of this study comparing a matched sample of struggling third-grade readers who were retained in grade level for poor reading performance were limited by the total sample size ($n=80$). The criteria used to select struggling readers for this study was poor performance

on the FCAT test at end of third grade, resulting in retention in that grade for another year. These criteria were fairly narrow as compared to those used in other research focused on struggling readers. Very often, struggling readers are selected based upon performance that is at least one standard deviation below the mean on traditional reading or achievement measures. The study was also limited to a comparison of a single guided silent reading intervention, Reading Plus, with a variety of other nationally marketed supplementary reading programs. It was not the purpose of this study to compare Reading Plus with any other specific interventions. Therefore, nothing can be said about this individual intervention's efficacy in comparison to other interventions not evaluated in this study.

The study was also limited by its geographical location and demographics. The study took place in a large, predominantly Hispanic, southeastern, urban school district environment. Therefore, the results of this study may not be generalizable to other regions, types, or sizes of school districts, or to other ethnic groups across the nation. The tests used in this study were also a limitation. Although the use of criterion-referenced state reading tests has become the standard by which most schools are judged as achieving AYP, the FCAT represents only one of many such tests used nationally and may be more or less technically and psychometrically sound in comparison to other such tests used across the nation. Similarly, the SAT-10 is only one of many psychometrically sound, nationally distributed, and norm-referenced reading achievement tests available and sold nationally. Finally, the design of the study was a limitation as well. Even though the use of propensity scores provides a more exacting approach for matching student characteristics to form experimental groups, it is nevertheless limited by the characteristics selected by the researchers for doing so. It is not as strong a research design for making inferences as would be a true, randomized, controlled experimental study.

Implications

This study provided emerging evidence supporting the use of a guided silent reading intervention known as Reading Plus for improving the reading comprehension and achievement scores of struggling third-grade readers on the FCAT. It did not provide similar evidence for the use of this guided silent reading intervention for improving the reading comprehension and achievement scores of struggling third-grade readers on the SAT-10. Future researchers may want to broaden the criteria used

to select struggling readers in order to enlarge the sample size. To increase the ability to generalize findings to other groups in other schools and classes across the nation, struggling readers should be selected from more than a single grade level, a single school district, and a single region of the country. In future research, a randomized, controlled trial would provide stronger evidence for making inferences about the potential efficacy of this guided silent reading intervention for struggling readers. Other supplementary intervention programs could also be used in future comparisons of the efficacy of the Reading Plus guided silent reading intervention used in this study. Additionally, future evaluations of this intervention's efficacy could also be assessed with varied reading and achievement assessment instruments that would provide a more sensitive measurement as well as multiple, converging data points. Future research may also investigate the use of wave or growth modeling to examine the build-up effects for this intervention in order to determine optimal length of use to achieve maximum improvements in reading comprehension and achievement.

Despite these improvements and the previously noted limitations, this study provides important evidence supporting the efficacy of a supplementary guided silent reading intervention with a sample of matched third-grade struggling readers who were retained at grade level. The guided silent reading intervention not only afforded this group of struggling third-grade students with appropriately challenging and varied reading genres to be both motivating and within their reach, but it also resulted in the great majority of these students making sufficient enough progress to be promoted to the next grade level. The guided silent reading intervention in this study provided students with guidance in terms of visual and perceptual modeling and rate management during silent reading; formatted their reading practice individually; adapted the text to be read by type, genre, and level of challenge; and continuously monitored their performance during silent reading practice. This combination of guided silent reading intervention elements nested within an adaptive online presentation environment was effective with this group of struggling third-grade readers on the FCAT test after a full year trial. Thus, the results of this study indicate that a guided silent reading intervention employing a suite of instructional elements as described in this study can offer classroom teachers a potentially useful and efficacious tool for providing struggling third-grade students with effective supplementary guided silent reading practice at school.

References

- Allington, R.L. (1977). If they don't read much, how they ever gonna get good? *Journal of Reading*, 21(1), 57–61.
- Allington, R.L. (2002). *Big brother and the national reading curriculum: How ideology trumped evidence*. Portsmouth, NH: Heinemann.
- Allington, R.L. (2006). Fluency: Still waiting after all these years. In S.J. Samuels & A.E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 94–105). Newark, DE: International Reading Association.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, 57(1), 289–300.
- Berk, R.A. (1998). Review of the Stanford Achievement Test, Ninth Edition. In J. C. Impara & B. S. Plake (Eds.), *The thirteenth mental measurements yearbook* (pp. 925–928). Lincoln, NE: University of Nebraska Press.
- Block, C.C., Cleveland, M.D., & Reed, K.M. (2006). When twenty minutes of literacy instruction is added to the day: Which learning environments increase students' overall achievement, vocabulary, comprehension, fluency, and affective development? In *Scholastic classroom books compendium of research* (pp. 18–36). New York, NY: Scholastic.
- Bloom, B.S. (1956). *Taxonomy of educational objectives* (Vol. 1). New York, NY: McKay.
- Brenner, D., & Hiebert, E.H. (2010). The impact of professional development on students' opportunity to read. In E.H. Hiebert & D.R. Reutzel (Eds.), *Revisiting silent reading: New directions for teachers and researchers* (pp. 198–220). Newark, DE: International Reading Association.
- Bryan, G., Fawson, P.C., & Reutzel, D.R. (2003). Sustained silent reading: Exploring the value of literature discussion with three non-engaged readers. *Reading Research and Instruction*, 43(1), 47–73.
- Carney, R.N., & Morse, D.T. (2005). Review of the Stanford Achievement Test, Tenth Edition. In R.A. Spies & B.S. Plake (Eds.), *The sixteenth mental measurements yearbook* (pp. 969–975). Lincoln, NE: University of Nebraska Press.
- Chall, J.S., Jacobs, V.A., & Baldwin, L.E. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Harvard University Press.
- Chinn, C.A., Waggoner, M.A., Anderson, R.C., Schommer, M., & Wilkinson, I.A.G. (1993). Situated actions during reading lessons: A microanalysis of oral reading error episodes. *American Educational Research Journal*, 30(2), 361–392.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Coles, G. (2000). *Misreading reading: The bad science that hurts children*. Portsmouth, NH: Heinemann.
- Cunningham, J.W. (2001). The National Reading Panel report. *Reading Research Quarterly*, 36(3), 326–335.
- Cunningham, P.M., & Allington, R.L. (2010). *Classrooms that work: They can all read and write* (5th ed.). Boston, MA: Pearson.
- Donovan, C.A., Smolkin, L.B., & Lomax, R.G. (2000). Beyond the independent-level text: Considering the reader–text match in first graders' self-selections during recreational reading. *Reading Psychology*, 21(4), 309–333.
- Eder, D., & Felmlee, D. (1984). The development of attention norms in ability groups. In P.L. Peterson, L.C. Wilkinson, & M. Hallinan (Eds.), *The social context of instruction: Group organization and group processes* (pp. 189–207). New York, NY: Academic Press.
- Edmondson, J., & Shannon, P. (2002). The will of the people. In R.L. Allington (Ed.), *Big brother and the national reading*

curriculum: How ideology trumped evidence (pp. 224–231). Portsmouth, NH: Heinemann.

Fawson, P.C., Reutzel, D.R., Read, S., Smith, J.A., & Moore, S.A. (2009). The influence of differing the paths to an incentive on third graders' reading achievement and attitudes. *Reading Psychology, 30*(6), 564–583.

Florida Center for Reading Research (2007). *Interventions for struggling readers*. Retrieved from <http://www.fcrr.org/interventions/Interventions.shtm>

Florida Department of Education. (2005). *FCAT handbook—A resource for educators*. Tallahassee, FL: Author. Retrieved from <http://fcat.fldoe.org/handbk/complete.pdf>

Florida Department of Education (2007). *K-12 Assessment*. Retrieved from http://fcat.fldoe.org/pdf/ufr_07_content.pdf

Fresch, M.J. (1995). Self-selection of early literacy learners. *The Reading Teacher, 49*(3), 220–227.

Guthrie, J.T., Schafer, W.D., & Huang, C.W. (2001). Benefits of opportunity to read and balanced instruction on the NAEP. *The Journal of Educational Research, 94*(3), 145–162.

Hiebert, E.H. (1983). An examination of ability grouping for reading instruction. *Reading Research Quarterly, 18*(2), 231–255.

Hiebert, E.H., & Reutzel, D.R. (2010). Revisiting silent reading in 2020 and beyond. In E.H. Hiebert & D.R. Reutzel (Eds.), *Revisiting silent reading: New directions for teachers and researchers* (pp. 290–299). Newark, DE: International Reading Association.

Hiebert, E.H., Wilson, K.M., & Trainin, G. (2010). Are students really reading in independent reading contexts? An examination of comprehension-based silent reading rate. In E.H. Hiebert & D.R. Reutzel (Eds.), *Revisiting silent reading: New directions for teachers and researchers* (pp. 151–167). Newark, DE: International Reading Association.

Hoffman, J.V., O'Neal, S.F., Kastler, L.A., Clements, R.O., Segel, K.W., & Nash, M.F. (1984). Guided oral reading and miscue focused verbal feedback in second-grade classrooms. *Reading Research Quarterly, 19*(3), 367–384.

Kamil, M.L. (2008). How to get recreational reading to increase reading achievement. In Y. Kim, V.J. Risko, D.L. Compton, D.K. Dickinson, M.K. Hundley, R.T. Jimenez, & D. Well Rowe (Eds.), *57th yearbook of the National Reading Conference* (pp. 31–40). Oak Creek, WI: National Reading Conference.

Kelley, M., & Clausen-Grace, N. (2006). R5: The sustained silent reading makeover that transformed readers. *The Reading Teacher, 60*(2), 148–156.

Kelley, M., & Clausen-Grace, N. (2010). R5: A Sustained Silent Reading makeover that works. In E.H. Hiebert & D.R. Reutzel (Eds.), *Revisiting silent reading: New directions for teachers and researchers* (pp. 168–180). Newark, DE: International Reading Association.

Krashen, S. (2002). More smoke and mirrors: A critique of the National Reading Panel Report on fluency. In R.L. Allington (Ed.), *Big brother and the national reading curriculum: How ideology trumped evidence* (pp. 112–124). Portsmouth, NH: Heinemann.

Kuhn, M.R. (2004). Helping students become accurate, expressive readers: Fluency instruction for small groups. *The Reading Teacher, 58*(4), 338–344.

Kuhn, M.R. (2005). A comparative study of small group fluency instruction. *Reading Psychology, 26*(2), 127–146.

Kuhn, M.R., & Schwanenflugel, P.J. (Eds.). (2008). *Fluency in the classroom*. New York: Guilford Press.

Mirsky, A.F. (1996). Disorders of attention: A neuropsychological perspective. In G.R. Lyon & N.A. Krasnegor (Eds.), *Attention, memory, and executive function* (pp. 71–95). Baltimore, MD: Paul H. Brookes Publishing.

National Governors Association Center for Best Practices & Council of Chief State

School Officers (NGACBP & CCSSO). (2010). *Common Core State Standards for English language arts & literacy in history/social studies, science, and technical subjects, Appendix A*. Washington, DC: Author. Retrieved from <http://www.corestandards.org/the-standards>

National Institute of Child Health, & Human Development. (2000). *Report of the national reading panel: Teaching children to read*. Washington, DC: National Institute of Child Health and Human Development & Institute of Education Sciences.

O'Connor, R.E., Bell, K.M., Harty, K.R., Larkin, L.K., Sackor, S.M., & Zigmond, N. (2002). Teaching reading to poor readers in the intermediate grades: A comparison of text difficulty. *Journal of Educational Psychology*, 94(3), 474–485.

Pikulski, J.J. (2006). Fluency: A developmental and language perspective. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 70–93). Newark, DE: International Reading Association.

Pressley, M. (2002). Comprehension strategies instruction: A turn-of-the-century status report. In C.C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 11–27). New York, NY: Guilford.

Rashotte, C.A., MacPhee, K., & Torgesen, J.K. (2001). The effectiveness of a group reading instruction program with poor readers in multiple grades. *Learning Disabilities Quarterly*, 24(2), 119–134.

Rasinski, T. . (1989). Fluency for everyone: Incorporating fluency instruction in the classroom. *The Reading Teacher*, 42(9), 690–693.

Rasinski, T.V., Samuels, S.J., Hiebert, E.H., Petscher, Y., & Feller, K. (2011). The relationship between a silent reading fluency instructional protocol on students' reading comprehension and achievement in an urban school setting. *Reading Psychology*, 32(1), 75–97.

Reutzel, D.R., & Hollingsworth, P.M. (1993). Effects of fluency training on

second graders' reading comprehension. *The Journal of Educational Research*, 86(6), 325–331.

Reutzel, D.R., Fawson, P.C., & Smith, J.A. (2008). Reconsidering silent sustained reading: An exploratory study of scaffolded silent reading. *The Journal of Educational Research*, 102(1), 37–50.

Reutzel, D.R., Jones, C.D., Fawson, P.C., & Smith, J.A. (2008). Scaffolded silent reading: A complement to guided repeated oral reading that works! *The Reading Teacher*, 62(3), 194–207.

Reutzel, D.R., Jones, C.D., & Newman, T.H. (2010). Scaffolded silent reading: Improving the conditions of silent reading practice in classrooms. In E.H. Hiebert & D.R. Reutzel (Eds.), *Revisiting silent reading: New directions for teachers and researchers* (pp. 129–150). Newark, DE: International Reading Association.

Rosenbaum, P.R., & Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.

Samuels, S.J. (2007). The DIBELS tests: Is speed of barking at print what we mean by reading fluency? *Reading Research Quarterly*, 42(4), 563–566.

Samuels, S.J. & Farstrup, A.E. (2006). *What research has to say about fluency instruction*. Newark, DE: International Reading Association.

Samuels, S.J., & Wu, Y. (2004). *How the amount of time spent on independent reading affects reading achievement: A response to the National Reading Panel*. Minneapolis, MN: University of Minnesota, Department of Educational Psychology.

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(2), 265–282.

Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference* (2nd ed.). Boston, MA:

Houghton Mifflin.

Swan, E.A., Coddington, C.S., & Guthrie, J.T. (2010). Engaged silent reading. In E.H. Hiebert & D.R. Reutzel (Eds.), *Revisiting silent reading: New directions for teachers and researchers* (pp. 95–111). Newark, DE: International Reading Association.

Topping, K.J. (2006). Building reading fluency: Cognitive, behavioral, and socioemotional factors and the role of peer-mediated learning. In S.J. Samuels & A.E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 106–129). Newark, DE: International Reading Association.

Torgesen, J.K., & Hudson, R.F. (2006). Reading fluency: Critical issues for struggling readers. In S.J. Samuels & A.E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 130–158). Newark, DE: International Reading Association.

Wonder-McDowell, C., Reutzel, D.R., & Smith, J.A. (2011). Does instructional alignment matter? Effects on struggling second graders' reading achievement. *The Elementary School Journal*, 112(2), 259–279.

Wright, G., Sherman, R., & Jones, T.B. (2004). Are silent reading behaviors of first graders really silent? *The Reading Teacher*, 57(6), 546–553.

Wright, G., Sherman, R., & Jones, T.B. (2010). Developmental considerations in transferring oral reading skills to silent reading. In E.H. Hiebert & D.R. Reutzel (Eds.), *Revisiting silent reading: New directions for teachers and researchers* (pp. 57–66). Newark, DE: International Reading Association.

III: LOOKING AHEAD

CHAPTER 7

Comprehension-Based Silent Reading Rates: What Do We Know? What Do We Need To Know?¹

Elfrieda H. Hiebert

TextProject, Inc. & University of California, Santa Cruz

S. Jay Samuels

University of Minnesota

Timothy V. Rasinski

Kent State University

As has been the case with many aspects of reading instruction, an emphasis on oral versus silent reading activities has varied in particular educational eras (Allington, 1984). During the whole language period of the 1990s, silent reading experiences were emphasized (Hagerty, 1999). Some oral reading occurred during guided reading and for obtaining running records, but occasions for monitored, repeated oral reading were few, even for beginning and struggling readers. However, in 2000, when the National Reading Panel (NRP; NICHD, 2000) concluded that guided, repeated oral reading but not sustained silent reading (SSR) facilitated fluency, comprehension, and vocabulary, the pendulum swung to an almost-exclusive emphasis on oral reading. An emphasis on oral reading went beyond the primary grades since the NRP had concluded that the fluency of all students through the fourth grade and struggling readers through high school was enhanced with guided, repeated oral reading. Evidence of the dominant role of oral reading can be seen in the prominence of the Dynamic Indicators of Basic Essential Literacy Skills (DIBELS)—a test of oral reading tasks—in the implementation of Reading First (Gamse, Jacob, Horst, Boulay, & Unlu, 2009).

Oral reading serves many critical functions, especially during the early stages of reading development. However, when the reading diet is no longer a balanced one, with oral reading dominating the menu, as

¹ This chapter was previously published in *Literacy Research and Instruction* (v51, n2, p110-124, 2012). The definitive publisher-authenticated version published is available online at: <http://www.tandfonline.com/doi/abs/10.1080/19388071.2010.531887>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

we believe is now the case, the prospects of “the poor getting poorer” (Stanovich, 1986) are likely. When the emphasis is on oral reading speed without attention to comprehension, as has been the case with DIBELS (Good & Kaminski, 2002) and similar oral reading assessments that have been prominent in Reading First implementations, beginning and struggling readers may come to perceive reading as nothing more than word calling (Samuels, 2007). Especially for the students whose reading experiences occur primarily in school, a diet heavy on oral reading with an emphasis on speed is unlikely to lead to the levels of meaningful, silent reading that are required for full participation in the workplace and communities of the digital-global age.

Proficient silent reading is the means whereby individuals access the ever-increasing stores of knowledge within texts that are required for the workplace and community. Silent reading rates and processing are limited by capacities such as eye movements (Samuels, Hiebert, & Rasinski, 2010), but oral reading rates are even more constrained by the speed of speech production. Very early in the acquisition of reading proficiency, silent reading rates exceed oral reading rates. Data on oral and silent reading norms (Hasbrouck & Tindal, 2006; Taylor, Frackenpohl, & Pettee, 1960) show that silent reading rates exceed oral reading rates by at least 30%, even for students who are at the 50th percentile in the primary grades. Once speech production becomes stable in early adolescence, the amount of words that can be read silently becomes substantially greater than can be read orally. As the findings of the NRP (NICHD, 2000) indicate, simply creating silent reading venues will not guarantee that students’ time will be used productively. However, under the right conditions where students read texts at appropriate difficulty levels, they process many more words in silent than in oral reading. In this context, our focus is on a construct that has been described as comprehension-based silent reading rate (Hiebert, Wilson, & Trainin, 2010). As this term implies, the emphasis on silent reading rate is always a function of appropriate levels of comprehension. Therefore, in considering comprehension-based silent reading rate, comprehension and rate are inseparable.

In this chapter, we review existing research on the silent reading performances of students and the nature of opportunities to read in classrooms that support meaningful, silent reading. In addition to a review of descriptive research on levels of performance and opportunities to read, we highlight several empirical studies that illustrate a new era of research on meaningful, silent reading where the effects on comprehension are

established. In the final section of the chapter, conclusions are summarized and questions are raised that require the attention of researchers if students are to be provided with the experiences that engender the kinds and levels of silent reading proficiencies that are needed in the 21st century.

Current Evidence on Comprehension-Based Silent Reading Rates

In this section, we examine three areas of descriptive research related to comprehension-based silent reading rate: (a) typical patterns of development and performance, (b) the relationship between oral and silent reading, and (c) attention to comprehension-based silent reading rates within most current instructional programs.

Typical patterns of development and performance

Whereas there are several sets of oral reading norms (e.g., Good & Kaminski, 2002; Hasbrouck & Tindal, 2006), information on silent reading norms is limited. Carver (1989) identified only two sets of norms that extended across the school years: those he had developed (Carver, 1982) and those of Taylor (1965). Carver chose the Taylor data to establish grade-equivalent norms because they were adequate on the following dimensions: sample size, sampling technique, range of grades represented, rates estimated in words per minute (wpm), and reliability. According to Carver's extrapolations of the Taylor norms, rate in wpm ranges from 0 to 81 at grade 1 to 333+ in grade 18 (graduate-level/proficient adult). There is a gain with each successive grade in school of about 10 to 20 wpm.

Even with Carver's (1989) additions, it should be remembered that the original data for these norms were gathered in the late 1950s (Taylor et al., 1960). The data that Carver (1983) gathered approximately 20 years after Taylor (1965) would suggest that reading rates stay fairly constant, but the technological changes even since Carver gathered his data have been substantial. Further, the silent reading norms that are currently available are provided for only the 50th percentile. How students do at the 10th, 25th, 75th, and 90th percentiles is also important information.

Despite these limitations, the silent reading norms have a significant component that the various oral reading norms that have proliferated over the past 20 years do not: The silent reading norms (Taylor et al., 1960) are based on comprehension. This distinction is an important one. We do not know how today's American students' comprehension-based silent reading rates compare to those of their counterparts 50 years

ago. There is data available, however, that can serve as a baseline for comparison.

While there are no data to indicate how current students' development of comprehension-based silent reading rates changes over time, there are data on how well students at particular levels perform on silent reading tests. The National Assessment of Educational Progress (NAEP; National Center for Education Statistics, 2009) shows that approximately 35% of a fourth-grade cohort fails to attain the basic level, while an additional 32% fails to reach the proficient level that is the goal for all students. Our analyses of the texts that have been used on these assessments indicate that they are at approximately a 3.5 grade-level according to conventional readability formulas. The text on the NAEP, then, is not the complex text that is emphasized within the Common Core Standards (National Governors Association for Best Practices & Council of Chief State School Officers, 2010).

When follow-up studies have been conducted on the NAEP where students read aloud texts that they have read silently, data indicate that it is the speed with which students read, not their word accuracy, that distinguishes students who achieve different standards. While there are differences in word accuracy across levels, these differences are not statistically significant (Pinnell et al., 1995). These differences in speed have been used as added justification for an emphasis on oral reading in instruction and assessment, as will be discussed shortly. However, it appears that many students, even those in the bottom quartile, can recognize almost 90% of the words in grade-level texts (Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003; Wise, Ring, & Olson, 1999). This word recognition is slow, however. In the Jenkins and colleagues study, fourth graders who had been classified as reading disabled read approximately 100 fewer words in a one-minute reading of a text than their same-age, skilled-reading peers.

The relationship between oral and silent reading

There has been a large amount of literature documenting the relationship between oral reading performances (without comprehension and measured as words correct per minute [wcpm]) and performances on silent reading tests, which have included standardized reading tests and state standards-based tests (Good, Simmons, & Kame'enui, 2001; McGlinchey & Hixson, 2004; Schatschneider et al., 2004). In Marston's (1989) review of such studies, the correlations were between .63 and

.90, with the most clustering around .80. In Good and Jefferson's (1998) review of the correlations within a single grade (grade 3), the correlations ranged from .60 to .80. Wiley and Deno (2005) and Pressley, Hildren, and Shankland (2005) have reported lower correlations (.40 to .50).

The finding of a high correlation between wcpm in oral reading and comprehension has had a strong influence on policy and practice in reading education. Underlying the mandates for instructional practice, assessment, and curricular materials regarding oral reading, there appears to be an implicit assumption that practice in oral reading will carry over to proficient silent reading rates and comprehension. Even with struggling adolescent readers, the fluency interventions reviewed by Wexler, Vaughn, Edmonds, and Reutebuch (2008) concentrated on oral reading. This emphasis, even when it has produced higher reading oral rates (which was not always the case), has typically not been reflected in improved comprehension on silent reading tasks.

Critical questions need to be asked, as Valencia and colleagues (2010) point out, about the use of wcpm as the basis for instructional assignments (e.g., who gets particular tiers of an intervention and who doesn't) and for instructional practices (i.e., an emphasis on oral reading activities). As Valencia and colleagues (2010) have shown, even including a measure of prosody within a wcpm assessment or varying the length of time produced stronger predictors of comprehension. The findings of Valencia and colleagues also point to the need for considering the changing relationship between wcpm and comprehension with developmental and proficiency levels. Within their study of second, fourth, and sixth graders, they found that the correlation between wcpm and comprehension decreased as students' proficiency increased.

In addition, the limitations of correlations need to be remembered. A high correlation does not necessarily impute a causal relationship between two variables. Further, when a variable (as is the case with wcpm) has a deviant range, the magnitude of correlations is affected (Valencia et al., 2010). In particular, a high correlation does not mean that two processes are identical. While there are shared processes, such as automatic recognition of words, there are significant and not-so-subtle differences between oral and silent reading processes. The most obvious is the role of vocalization. Whereas overt vocalization can be an impediment in silent reading, it is the outcome in oral reading. Every word needs to be read in an oral reading context, whereas readers can use context to grasp the meaning of words that they cannot pronounce while reading silently (Nagy, Anderson, & Herman, 1987).

Another significant difference lies in support for staying with the task. Oral reading involves external supports such as the teacher or a recording device. These external supports mean that students' attention to the text is ensured. They can't move their eyes away from the text for several minutes and daydream. They cannot skip over a page or two that looks disinteresting or too difficult. In silent reading, individuals do not have external supports. Students need to learn to persevere as well as monitor what they are reading if they are to comprehend texts that they are reading silently. Monitoring strategies become particularly critical when background knowledge is limited, a circumstance that is often the case for less proficient readers (McKeown, Beck, & Blake, 2009).

As scholars have observed (e.g., Pearson, 2006; Pressley et al., 2005; Rasinski, 2006; Samuels, 2007), instruction that aims at increasing students' wcpm without attention to comprehension has the potential to adversely affect comprehension and knowledge acquisition. The development of proficient silent reading strategies and habits, including comprehension-based silent reading rates, likely require unique experiences and instruction. We move next to studies of the nature of current reading instruction to determine attention given to comprehension-based silent reading rates.

Attention to comprehension-based silent reading rates within most current instructional programs

The presence (or lack thereof) of opportunities to read silently in school predicts reading achievement (Foorman et al., 2006; Guthrie, Schafer, & Huang, 2001; Taylor, Frye, & Maruyama, 1990). Most students, however, do not spend substantial periods of their school day reading. In the 1980s, Gambrell (1984) reported that students read approximately 14 minutes daily. Close to a decade later, Foertsch (1992) documented similar amounts of reading. A decade after that, a survey by Donahue, Finnegan, Lutkus, Allen, and Campbell (2001) showed that fourth graders reported, on average, reading 10 or fewer pages per day in school and for homework, which translates into approximately 8 to 12 minutes of daily reading.

In a recent analysis of how much time during a 90-minute reading block was spent reading texts, Brenner, Hiebert, and Tompkins (2009) found that students spent an average of 18 minutes reading text—20% of the reading period. Half of this time was spent reading orally, and half was spent reading silently. The 90-minute reading periods were spent on a variety of activities, such as lessons on elements of texts or words,

playing word games, and completing workbook pages. Students' time in engaged reading, however, was limited. Even within the time that students spent reading in a single school day, the time devoted to any one text was typically short. Students might read a story about a boy writing to his grandmother in Korea in the large-group, teacher-directed period, an excerpt of text about a basketball player in the small group directed by their teacher, and a fantasy about woodland creatures during partner reading.

Since the teachers in the Brenner and colleagues study (2009) were expected to specifically follow the guidance in the teachers' guides of the core reading programs, Brenner and Hiebert (2010) examined recommendations for opportunities to read within these guides. The third-grade editions of six leading core reading programs provided an average of 15 minutes of reading volume per day, ranging from approximately 10 to 24 minutes. The findings of the observational study, then, reflected the recommendations in the teachers' guides.

When students do have opportunities to read silently, there appears to be little scaffolding of the task. Consequently, as the report of the NRP (NICHD, 2000) concluded, opportunities to read that lack structure and support often fail to produce the hoped-for outcomes. Without appropriate structure and support, students often engage in what some teachers have called "fake reading" during SSR (Griffith & Rasinski, 2004). While the structuring of recreational reading is critical to its success, scaffolding the processes of proficient silent reading would be expected to occur during reading periods, not recreational reading times.

In their examination of the opportunities to read that are recommended in these programs, Brenner and Hiebert (2010) also analyzed teachers' guides for differences in types or amounts of reading opportunities for students of different proficiency levels. We might expect that at the third-grade level—the focus of that analysis—particular forms of scaffolding might be provided for struggling readers. For example, students who are not adept at reading on their own might be assigned accessible texts and monitored more frequently by the teacher. However, that was not the case in the core reading programs. Low-performing students were given the same texts for the same periods of time as their higher-performing peers.

When differentiations are made within interventions, it appears that practices can discourage attention to comprehension-based silent reading rate. Rather, an underlying assumption appears to be that it is decoding skills, especially as represented by phonological deficits, that

require attention (see, e.g., Wexler et al., 2008). The assumption that low-performing students lack decoding abilities leads to mandates such as that of the California State Board of Education (2007). Curricular materials adopted by that state for interventions aimed at struggling readers in grades four through eight must provide approximately 9,000 words of decodable text, including two decodable reading selections for each of the 44 sound-spelling correspondences. Such texts are typically short, and the instructional routines emphasize oral reading (Wexler et al., 2008). For example, in one of the currently approved programs for struggling middle-school readers in California, the texts in the comprehension component are typically 60 to 80 words long (Engelmann, Osborn, & Hanner, 2002). Participation with such texts is unlikely to develop the strategies of proficient, independent reading. When reading is tedious, students are less likely to read outside of school as well (Anderson, Wilson, & Fielding, 1988). At the same time, their more proficient peers are enriching their knowledge and vocabularies. As Stanovich (1986) describes it, the rich get richer, and the poor get poorer.

Experimental Evidence on the Instruction of Comprehension-Based Silent Reading Rates

There are increasing indications of the kinds of scaffolds that can support the development of effective silent reading habits among readers. This research literature is not extensive by any means, but over the last several years a handful of studies have been conducted. We have chosen three of these studies—one from each of three developmental levels: primary, middle and high school, and young adult/college—to illustrate the emerging evidence of how silent reading proficiencies can be guided through instruction.

Study with primary-level students

A study by Reutzel, Fawson, and Smith (2008) indicates that guidance in silent reading in the primary grades can have efficacious effects on fluency, vocabulary, and comprehension. Conducted in the wake of the NRP (NICHD, 2000), Reutzel and colleagues were interested in whether a well-designed silent reading treatment could produce comparable results to the guided repeated oral reading (GROR) that the NRP had identified as the gold standard for promoting fluency, comprehension, and vocabulary. Because the studies within the meta-analysis had primarily used oral reading measures in establishing effects

on students’ rate, comprehension, and vocabulary, Reutzel and colleagues used oral reading measures as outcomes. This emphasis on oral reading measures seems like an appropriate one at this transition point in the primary grades when students move from predominantly oral reading to silent reading. Further, this study is the only experiment with a focus on reading modes that has been conducted since the NRP’s report.

The study involved 4 third-grade teachers and their 72 third-grade students. The schools had approximately 35 to 50% African, Asian, and Latino American students, and more than half of the students in the schools qualified for free or reduced lunch. Students were randomly assigned to one of two treatment conditions: scaffolded silent reading (ScSR) with monitoring and a wide reading of different genres at students’ independent reading levels, or GROR of grade-level texts with feedback from teachers and peers.

The ScSR and GROR treatments were similar in four ways. First, an equivalent amount of time was spent on core reading instruction and in the experimental conditions. Second, teachers in all four classrooms used the same instructional materials and procedures. Third, all four classrooms used a take-home reading library, records for which were reviewed by teachers weekly to ensure that students were reading 15 minutes daily outside of school. Finally, teachers in both conditions began daily sessions by modeling fluent reading of a text and discussing with students various characteristics of fluent reading.

The two treatments in this study, ScSR and GROR, were differentiated on six dimensions, as illustrated in Table 7.1.

Table 7.1: Differences between ScSR and GROR Treatments on Six Dimensions

Dimension	ScSR	GROR
Mode of reading	silent	oral
Nature of reading	wide	repeated
Frequency of feedback/monitoring	weekly	daily
Social nature	isolated	collaborative
Source of texts	student-selected	teacher assigned
Text Difficulty	independent	grade level

Quantitative results indicated no significant differences between these two forms of reading practice on third-grade students’ fluency, accuracy, comprehension, or expression with the exception of one significant difference favoring ScSR on expression of a single passage. For these primary-level students, then, silent reading experiences that had been

carefully designed and executed could produce results as efficacious as the guided, repeated oral reading that the NRP (NICHD, 2000) recommended as the means for increasing reading proficiency.

These effects cannot be attributed only to the mode of reading in that five additional variables distinguished the two interventions. Other variables, such as self-selection versus teacher-assignment of texts, have been shown to influence students' interest in reading and their sense of agency as readers (Guthrie et al., 2006) and wide reading has been shown to be more efficacious than repeated reading (Kuhn, 2005; Kuhn et al., 2006). At the same time, the silent reading condition had components that could be argued to be potential challenges for primary-level students, such as less teacher monitoring and less opportunity for social interaction. Overall, the components of the silent reading condition describe the typical act of silent reading where readers choose texts to read that are appropriate for them (including text difficulty) and then read them a single time in a solitary fashion and without continual monitoring. In contexts where teachers worked to ensure that the expectations and conditions were appropriate, primary-level students were able to benefit from time devoted to typical silent reading venues to the same degree as students who spent equivalent amounts of time in highly prescribed reading contexts.

Study with middle- and high-school students

The Reutzel and colleagues (2008) study provides information on the efficacy of silent reading scaffolds that are provided at an appropriate developmental time. This support can be provided in classrooms. As the review of the NAEP data indicated, there are many students who have passed this transition point and have less-than-efficacious patterns. Research indicates a variety of programs in digital contexts that have been offered as providing scaffolding for struggling middle- and high-school readers. Features of digital contexts, such as the ability to instantaneously change the difficulty of the text in response to comprehension performances, allow for precision in scaffolding that is difficult to achieve in a classroom setting or even a tutoring one. While the claims of these programs, by and large, have not been validated, studies are beginning to be conducted, as illustrated by Rasinski, Samuels, Hiebert, Petscher, and Feller (2011). This Rasinski and colleagues study indicates that consistent participation in a digital context over a school year can result in improved performances on high-stakes assessments—both a norm-referenced test (NRT) and a criterion-referenced test (CRT).

The study was conducted in a large, urban school system. To deal with the historically poor performances of students in this district, schools were offered several supplementary reading programs. The study was conducted in 23 schools with students in grades 4 through 10 where a web-based reading intervention, Reading Plus (RP), was implemented. In some of these schools, only the low-performing students were assigned to RP. Other schools chose to use RP with specific sub-populations or grade levels. Students not assigned to RP may have used Scholastic's Read 180 and/or Renaissance Learning's Accelerated Reader. The study had a significant portion of African American (46%) and Latino American (50%) students. Sub-populations in the sample included learning disabled (6%) and English language learning students (3%).

Over the six months of the study, students participated in either two 45-minute sessions or three 30-minute sessions weekly. Since students moved through lessons in individual computer environments, differences in length of the individual sessions did not influence content coverage. During their first RP session, students completed an assessment with texts at varying difficulty levels that determined their independent silent reading rates, comprehension, and vocabulary. Performances on these assessments formed the basis for the instructional paths that students then followed in the intervention.

A typical lesson contained two warm-up activities that were intended to build foundational skills such as attention, left-to-right tracking, perceptual accuracy, and visual memory. The heart of each lesson was a structured silent reading activity where students read texts at their instructional reading level from a database of 600 selections ranging from preprimer to adult-level texts. Each reading of a text was followed by comprehension questions (focusing on literal understanding, interpretation, analysis, evaluation, and appreciation). The digital environment ensured that adjustments in instructional experiences were made continually based on student performances. For example, the lengths of segments within texts were increased or decreased based on a student's comprehension and silent reading rate.

Since the aim was for students to participate in the program for a minimum of 30 hours (approximately 40 45-minute sessions), students were divided into two groups: those who received 40 or more lessons over the course of the school year, and those who received 39 or fewer. While there are serious limitations to employing gain scores to test for differences between groups, Rogosa (1995) has shown that the gain score is as reliable

as a covariance-adjusted score, and it is more appropriate to use in quasi-experimental studies than posttest only. For all grades but grade 9, students with the 40+ lesson interventions had significantly higher performances on the NRT, and for all grades but 4 and 10, significantly higher performances on the CRT. Effect sizes were established in relation to performances of students within the same schools who did not participate in this particular intervention. Effect sizes by grade level ranged from .03 to .34 (small to moderate in magnitude [Cohen, 1988]).

In relation to the CRT, mean gain scores for students who received 40+ RP lessons were greater than the statewide and district-level gains at each of the seven grade levels. In several cases, the gains were substantial in their magnitude, as was the case in grades 6, 7, and 8, where mean gains on the CRT were more than double the gains of nonparticipating students. Typically, low-achieving students' growth for a year's worth of instruction is less than what is expected for average and high-achieving students. Middle- and high-school students who are struggling readers are often inconsistent in their performances on transfer measures after an intervention (Torgesen et al., 2007; Wexler et al., 2008). However, that was not the case with this intervention, where students performed substantially better on the high-stake tests of their state that were conducted in typical paper-and-pencil, large-group contexts. It is noteworthy that the yearlong gains made by the primarily low-achieving students in this silent reading intervention were substantially larger than the mean overall gains at the state and district levels.

Study with college students

The final study that we present as evidence for the thesis that comprehension-based silent reading skills can be supported was conducted with college students. Radach, Vorstius, and Reilly (2010) initiated this study after identifying a proliferation of speed-reading programs promoted on the internet but finding that the few available studies on these programs typically had arrived at unfavorable conclusions (e.g., McNamara, 2000). The proliferation of programs, Radach and colleagues argue, illustrate a fascination by the general populace with improving their reading rates. Of claims within 12 programs promoted on the Internet, Radach and colleagues identified two that have some substantiation in research on reading processes and designed an intervention around these processes: decreasing inter-word regressions and attending to meaning units.

The experimental group received acoustic feedback following inter-word regressions. To support attending to meaning units, a word group (typically a phrase) of two to three words was highlighted by alternating the color of the text on the computer screen. A comparison group was simply told to increase their reading speed. After a session of establishing baseline comprehension-based silent reading rates, the two groups participated in four training sessions. Each session consisted of two cycles of learning and practice. In the learning part of the cycle, students received feedback (acoustic feedback and phrasing support in the experimental group, and encouragement to slow down or speed up in the comparison group). The feedback was intended to support a 20% speed increase during each of the four sessions. This feedback was provided during the learning cycle at the end of each page of text when comprehension questions were also provided. In the practice phase of a cycle, feedback was not provided.

The pre- and posttest consisted of eight passages of texts, each about 400 words long (3,200 words in all). All were nonfiction pieces on popular science topics. To assess comprehension, eight statements per passage (64 in all) were presented for verification of three levels of text representation: verbatim, paraphrases, and inferences.

On average, participants had baseline reading rates of approximately 198 wpm (comparison) and 185 wpm (experimental). In the posttest, wpms were 365 (comparison) and 350 (experimental). The comprehension of the experimental group increased approximately 3%, while the comprehension of the control group decreased about 6%. In the specific training group, regressive saccades back to earlier words dropped by 50%, indicating that this aspect of the training was very effective. However, the specific speed training techniques produced no advantage over the group with the unspecified fluency training (i.e., “read faster”).

A point to be emphasized is that reading speed of even the fastest reader in the sample did not exceed 700 words. Approximately one in five of the participants (22%) read at 400 wpm or higher. But overall, the students read at approximately 350 wpm. Another observation is that the college students’ baseline rates of approximately 190 wpm was considerably lower than the average rate reported by Taylor, Frankenhohl, and Pettee (1960) for college students—280 wpm—and considerably below the optimal rate that Taylor and colleagues claimed could be attained with training—480 wpm. The training brought the college students in 2010 to the range of college students *before* training in 1960.

Comprehension-Based Silent Reading Rates: Emerging Answers and Remaining Questions

Comprehension-based silent reading rates, Radach and colleagues (2010) argue, represent a nexus in understanding the roles of and relationships between word recognition and comprehension. Knowledge about this topic is of more than theoretical interest. This knowledge is also critical in understanding how students can be brought to the levels that are required for careers and community participation. The review of research shows that answers are emerging on this topic. Even so, information is limited on how comprehension-based silent reading rates develop and also on the nature of comprehension-based silent reading performances of students of different proficiency levels. The only consistent form of evidence on the latter comes from the NAEP (National Center for Education Statistics, 2009). Since that assessment provides only a single text level and does not distinguish across different kinds of tasks in the report of results, we know what it is that below-basic and basic students cannot do, but we do not know what it is that they can do. To create viable instructional programs that support readers' silent reading capacities, descriptions of the tasks and texts with which students are successful also require attention.

Another consistent finding from the research is a consistently high correlation between wcpm from oral reading assessments and performances on silent reading tests. This finding has been the justification for a heavy emphasis on oral reading activities and assessments in instruction. While oral and silent reading share processes, there are also distinctions in the two processes. Further, the heavy emphasis on oral reading has not resulted in the needed increases in comprehension-based silent reading rates.

Systematic instruction that supports students' capacities as readers who comprehend texts at optimal reading rates is not evident in typical classrooms or in the guidance for teachers that is found within the core reading programs that have been promoted as part of state and federal mandates. A recent, albeit small, group of studies is beginning to show how such systematic instruction can be provided. While the evidence is limited in scope and size at this point, these studies indicate that there are instructional mechanisms that can support students in developing the comprehension-based silent reading rates needed for the 21st century.

While the handful of studies that have emerged over the past several years point to potential solutions, the questions regarding

comprehension-based silent reading rates far outweigh the answers. These questions need to be addressed before widespread changes in practice can occur. While there is clearly a serious gap in current levels of comprehension-based silent reading rates relative to the demands of the digital-global age, moving swiftly to implement solutions such as increased silent reading time or interventions that encourage students to “read faster” are unlikely to make the changes that are needed. While most students would likely benefit from higher allocations of time devoted to silent reading during the school day, silent reading events require careful design. In the Reutzel and colleagues (2008) study, students were not simply told to sit at their desks and to read silently for extended periods of time. Their teachers made numerous shifts in the design of reading events to support students’ silent reading. Further, the intervention in the Reutzel and colleagues study took a very different form than that of the Rasinski and colleagues study (2011) with middle- and high-school students. The struggling readers in the Rasinski and colleagues study participated in a program with an underlying platform that made it possible to change levels of text within a single session based on students’ comprehension and reading rate. This pattern is quite unlike the “one size fits all” perspective that is often promoted by mandates or guidelines, as was the case with the SSR model of Hunt (1970), where students, regardless of developmental or proficiency level, were given the same treatment.

While the research to date is insufficient to provide guidelines for practices, the initial findings suggest questions on which research programs can be built. We describe the basis and potential directions for two critical questions: (a) what are optimal comprehension-based silent reading rates that are fairly consistent across tasks and texts? and (b) when and in what contexts should comprehension-based silent reading rates be developed?

What are optimal comprehension-based silent reading rates?

There appears to be a strong tendency among people to want to “break the barrier” in tasks where speed is involved, as is the case with silent reading. Evidence from well-designed studies by reputable researchers can help to guard against this tendency in instruction on silent reading proficiencies. In the digital-global age where the amount of information has increased incredibly, the ability to sustain comprehension-based silent reading rates over extended text is a necessity. But the emphasis needs to be on sustaining meaningful comprehension at

appropriate rates across numerous settings, not simply the rate at which students are reading. Carver (1990) emphasized the need for considering how readers adapted processes and rates to different kinds of texts and in different tasks. However, few definitive descriptions are available, particularly in the form of norms.

Levels of comprehension with different texts and tasks also require attention. The level of acceptable comprehension in the Rasinski and colleagues (2011) study was 70%. If students fell below that level, the architecture of the digital platform was such that they were moved to somewhat easier text or text was presented at a slower speed. In the Radach and colleagues (2010) study, the students in the “read faster” group fell below this percentage. Comprehension levels of 70%, for some tasks, may be insufficient, while for other texts and tasks entirely adequate. The sacrifices in comprehension as a function of rate is, as Radach et al. have noted, a critical area that requires further study.

When and in what contexts should comprehension-based silent reading rates be developed?

In a domain such as playing the piano, which, like reading words, involves both a cognitive and physical component, practice and development occurs over an extended period of time and as a result of substantial experience. Similarly, support for optimal comprehension-based silent reading rates needs to be viewed as a long-term endeavor with different emphases at different points. In this section, we speculate about when and where scaffolds might be put in place for such a long-term endeavor. We use Chall’s (1983) six reading stages, provided in Table 7.2, as the basis to distinguish between different phases in reading development.

Table 7.2: Chall’s (1983) Reading Stages

Stage	Primary Task	Grade Span
0	Prereading	Through kindergarten
1	Initial reading or decoding	Grades 1–2
2	Confirmation, fluency, ungluing from print	Grades 2–3
3	Reading for learning the new	Grades 4–8
4	Reading for multiple viewpoints	High school
5	Construction and reconstruction: A world view	College

Source: Chall, J.S. (1983). *Stages of reading development*. New York, NY: McGraw-Hill.

During Chall's stage zero, frequent conversations between adults (e.g., parents, kindergarten and preschool teachers) and children about reading likely occur, including the reading of literacy-focused books, such as *I Can Read with My Eyes Shut!* (Dr. Seuss, 1978). However, systematic experiences in attending to silent reading strategies and rates would not be expected to be a focus until children begin formal instruction. As students move into the first of Chall's stages, short periods of time would be devoted to reading "just with your eyes." These periods would be carefully paced and monitored and would not be expected to stretch for the long periods of time that advocates of readers' workshop (Hagerty, 1999) or SSR (Hunt, 1970) recommend. As students' reading proficiency increases in stage two, silent reading episodes would be expected to increase somewhat (although, again, not to the extremes recommended in the readers' workshop and SSR literature). As Reutzel and colleagues (2008) demonstrate, there are numerous elements of these events that require attention, including (but not limited to) the length of time, teacher monitoring, appropriate texts, and clarity about the anticipated outcomes of the event.

It is in stages 3 and 4 where the careful orchestration of silent reading events in classrooms is likely to have the greatest pay-off in terms of supporting optimal reading rates. Third to fourth grade has been described as the point where silent reading processes have developed sufficiently to be more efficient than oral reading (Juel & Holmes, 1981). For students who are vulnerable as readers, earlier in this span of time is likely to support their development of appropriate comprehension-based silent reading rates rather than later. In addition to the careful crafting of texts and tasks in classroom settings, the digital contexts that have added demands on literacy proficiencies may be one of the primary means for supporting more efficacious silent reading proficiencies. The architecture of digital programs can be designed so that the length of time, the accessibility of text, and the tasks can be carefully adjusted to students' growing capacity as readers.

While, to date, it is frequently the struggling readers who participate in digital contexts, the effects of such participation for students who are proficient readers require attention, particularly when silent reading rates are beginning to stabilize. Once individuals reach stage 5, making changes to baseline comprehension-based silent reading rates is likely challenging and difficult. Consequently, projects that determine different configurations of experiences in such contexts, especially at stages 3 and 4, should be a priority in research on reading for understanding. Such an emphasis is particularly needed during a time when a theme

within the private and public sector is to increase the preparedness of individuals for the marketplace and communities of the digital-global economy (National Governors Association for Best Practices & Council of Chief State School Officers, 2010).

Conclusion

The need for efficient silent reading habits for success in the digital-global age is unarguable. There is emerging evidence that these habits can be enhanced through scaffolding, both on the part of teachers and from digital supports. These supports look quite different than the SSR that Hunt (1970) advocated. This structuring can begin when students are in the early stages of reading (Reutzel et al., 2008). Further, it is highly likely that the process is an ongoing endeavor, extending through the elementary grades and into middle and high school as students encounter new genres and content. At least for the students who depend on schools to become literate, good silent reading does not just happen as a result of an emphasis on oral reading fluency training. For many students, good silent reading habits require that they participate in structured silent reading experiences that model efficient reading.

The authors acknowledge their role as members of an advisory board to the publisher of the program described in the second study of the empirical section of this review. For their role on the advisory board, the authors receive an annual, flat honorarium that is not tied to sales of the product.

References

- Allington, R.L. (1984). Content coverage and contextual reading in reading groups. *Journal of Literacy Research*, 16(2), 85–96.
- Anderson, R.C., Wilson, P.T., & Fielding, L.G. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly*, 23(3), 285–303.
- Brenner, D., & Hiebert, E.H. (2010). If I follow the teachers' editions, isn't that enough? Analyzing reading volume in six core reading programs. *The Elementary School Journal*, 110(3), 347–363.
- Brenner, D., Hiebert, E.H., & Tompkins, R. (2009). How much and what are third graders reading? In E.H. Hiebert (Ed.), *Reading more, reading better* (pp. 118–140). New York, NY: Guilford.
- California State Board of Education. (2007). *Reading/language arts framework for California public schools: Kindergarten through grade twelve*. Sacramento, CA: California Department of Education.
- Carver, R.P. (1982). Optimal rate of reading prose. *Reading Research Quarterly*, 18(1), 56–88.
- Carver, R.P. (1983). Is reading rate constant or flexible? *Reading Research Quarterly*, 18(2), 190–215.
- Carver, R.P. (1989). Silent reading rates in grade equivalents. *Journal of Literacy Research*, 21(2), 155–66.
- Carver, R.P. (1990). *Reading rate: A review of research and theory*. San Diego, CA: Academic Press.
- Chall, J.S. (1983). *Stages of reading development*. New York, NY: McGraw-Hill.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Engelmann, S., Osborn, S., & Hanner, S. (2002). *Comprehension skills (BI)*. Columbus, OH: McGraw-Hill SRA.
- Foertsch, M.A. (1992). *Reading in and out of school: Factors influencing the literacy achievement of American students in grades, 4, 8, and 12, in 1988 and 1990*. Washington, DC: National Center for Educational Statistics, U.S. Department of Education.
- Foorman, B.R., Schatschneider, C., Eakin, M.N., Fletcher, J.M., Moats, L.C., & Francis, D.J. (2006). The impact of instructional practices in grades 1 and 2 on reading and spelling achievement in high poverty schools. *Contemporary Educational Psychology*, 31(1), 1–29.
- Gambrell, L. (1984). How much time do children spend reading during teacher-directed reading instruction? In J.A. Niles & L.A. Harris (Eds.), *Changing perspectives on research in reading/language processing and instruction, Thirty-third yearbook of the National Reading Conference* (pp. 193–198). Rochester, NY: National Reading Conference.
- Gamse, B.C., Jacob, R.T., Horst, M., Boulay, B., & Unlu, F. (2009). *Reading first impact study final report: Executive summary* (NCEE 2009-4039). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Good, R.H., III, Jefferson, G. (1998). Contemporary perspectives on curriculum-based measurement validity. In M.R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 61–88). New York, NY: Guilford.
- Good, R.H., III, & Kaminski, R.A. (2002). *DIBELS oral reading fluency passages for first through third grades* (Technical Report No. 10). Eugene, OR: University of Oregon.
- Good, R.H., III, Simmons, D.C., & Kame'enui, E.J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-

grade high-stakes outcomes. *Scientific Studies of Reading*, 5(3), 257–288.

Griffith, L.W., & Rasinski, T.V. (2004). A focus on fluency: How one teacher incorporated fluency with her reading curriculum. *The Reading Teacher*, 58(2), 126–137.

Guthrie, J.T., Schafer, W.D., & Huang, C.W. (2001). Benefits of opportunity to read and balanced instruction on the NAEP. *The Journal of Educational Research*, 94(3), 145–162.

Guthrie, J.T., Wigfield, A., Humenick, N.M., Perencevich, K.C., Taboada, A., & Barbosa, P. (2006). Influences of stimulating tasks on reading motivation and comprehension. *The Journal of Educational Research*, 99(4), 232–246.

Hagerty, P.J. (1999). *Readers' workshop: Real reading*. New York, NY: Scholastic.

Hasbrouck, J., & Tindal, G.A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher*, 59(7), 636–644.

Hiebert, E.H., Wilson, K.M., & Trainin, G. (2014). Are students really reading in independent reading contexts? An examination of comprehension-based silent reading rate. In E.H. Hiebert & D.R. Reutzel (Eds.), *Revisiting silent reading: New directions for teachers and researchers* (pp. 151–167). Newark, DE: International Reading Association.

Hunt, L.C., Jr. (1970). The effect of self-selection, interest, and motivation upon independent, instructional, and frustration levels. *The Reading Teacher*, 24(2), 146–151, 158.

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 & Practice*, 18(4), 237–245.

Juel, C. & Holmes, B. (1981). Oral and silent reading of sentences. *Reading Research Quarterly*, 16(4), 545–568.

Kuhn, M.R. (2005). A comparative study

of small group fluency instruction. *Reading Psychology*, 26(2), 127–146.

Kuhn, M.R., Schwanenflugel, P.J., Morris, R.D., Morrow, L.M., Woo, D.G., Meisinger, B., ... Stahl, S.A. (2006). Teaching children to become fluent and automatic readers. *Journal of Literacy Research*, 38(4), 357–389.

Marston, D.B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it. In M.R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18–78). New York, NY: Guilford.

McGlinchey, M.T., & Hixson, M.D. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review*, 33(2), 193–203.

McKeown, M.G., Beck, I.L., & Blake, R.G.K. (2009). Rethinking reading comprehension instruction: A comparison of instruction for strategies and content approaches. *Reading Research Quarterly*, 44(3), 218–253

McNamara, D.S. (2000). *Preliminary analysis of Photoreading* (Final Report). Moffett Field, CA: NASA Ames Research Center.

Nagy, W.E., Anderson, R.C., & Herman, P.A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, 24(2), 237–270.

National Center for Education Statistics. (2001). *The nation's report card. Fourth-grade reading 2000* (NCES 2001–499). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics. (2009). *The nation's report card. Reading 2009: National Assessment of Educational Progress at grades 4 and 8* (NCES 2010–458). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common*

Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects. Washington, DC: Authors.

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. 00-4769). Washington, DC: U.S. Government Printing Office.

Pearson, P.D. (2006). Foreword. In K.S. Goodman (Ed.), *The truth about DIBELS: What it is, what it does* (pp. v–xix). Portsmouth, NH: Heinemann.

Pinnell, G.S., Pikulski, J.J., Wixson, K.K., Campbell, J.R., Gough, P.B., & Beatty, A.S. (1995). *Listening to children read aloud: Data from NAEP's integrated reading performance record (IRPR) at grade 4* (Rep. No. 23-FR-04). Washington, DC: Office of Education Research and Improvement, U.S. Department of Education.

Pressley, M., Hildren, K., & Shankland, R. (2005). *An evaluation of end-grade-3 Dynamic Indicators of Basic Early Literacy Skills (DIBELS): Speed reading without comprehension, predicting little*. East Lansing, MI: Michigan State University, College of Education, Literacy Achievement Research Center.

Radach, R., Vorstius, C., & Reilly, R. (July 10, 2010). *The science of speed reading: Exploring the impact of speed on visuomotor control and comprehension*. Paper presented at the annual meeting of the Society for the Scientific Study of Reading, Berlin, Germany.

Rasinski, T. (2006). Reading fluency instruction: Moving beyond accuracy, automaticity, and prosody. *The Reading Teacher*, 59(7), 704–706.

Rasinski, T., Samuels, S.J., Hiebert, E., Petscher, Y., & Feller, K. (2011). The relationship between a silent reading fluency instructional protocol on students' reading comprehension and achievement in an urban school setting.

Reading Psychology, 32(1), 75–97

Reutzel, D.R., Fawson, P.C., & Smith, J.A. (2008). Reconsidering silent sustained reading: An exploratory study of scaffolded silent reading. *The Journal of Educational Research*, 102(1), 37–50.

Rogosa, D.R. (1995). Myths and methods: "Myths about longitudinal research" plus supplemental questions. In J.M. Gottman (Ed.), *The analysis of change* (pp. 3–65). Hillsdale, NJ: Lawrence Erlbaum Associates.

Samuels, S.J. (2007). The DIBELS tests: Is speed of barking at print what we mean by reading fluency? *Reading Research Quarterly*, 42(4), 563–566.

Samuels, S.J., Hiebert, E.H., & Rasinski, T.V. (2010). Eye movements make reading possible. In E.H. Hiebert & D.R. Reutzel (Eds.), *Revisiting silent reading: New directions for teachers and researchers* (pp. 24–44). Newark, DE: International Reading Association.

Schatschneider, C., Buck, J., Torgesen, J.K., Wagner, R.K., Hassler, L., Hecht, S., & Powell-Smith, K. (2004). *A multivariate study of factors that contribute to individual differences in performance on the Florida Comprehensive Reading Assessment Test* (Technical Report # 5). Tallahassee, FL: Florida Center for Reading Research.

Seuss, Dr. (1978). *I can read with my eyes shut!* New York, NY: Random House.

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360–407.

Taylor, B.M., Frye, B.J., & Maruyama, G.M. (1990). Time spent reading and reading growth. *American Educational Research Journal*, 27(2), 351–362.

Taylor, S.E., Frackenpohl, H., & Pettee, J.L. (1960). *Grade level norms for the components of the fundamental reading skills* (EDL Research and Information Bulletin #3). New York, NY: Educational Developmental Laboratories/McGraw Hill.

- Taylor, S.E. (1965). Eye movements in reading: Facts and fallacies. *American Educational Research Journal*, 2(4), 187–202.
- Torgesen, J.K., Myers, D., Schirm, A., Stuart, E., Vartivarian S., Mansfield, W., ... Haan, C. (2007). *Closing the achievement gap: Second year findings from a randomized trial of four reading interventions for striving readers*. Washington, DC: The Corporation for the Advancement of Policy Evaluation.
- Valencia, S.W., Smith, A.T., Reece, A.M., Li, M., Wixson, K.K., & Newman, H. (2010). Oral reading fluency assessment: Issues of construct, criterion, and consequential validity. *Reading Research Quarterly*, 45(3), 270–291
- Wexler, J., Vaughn, S., Edmonds, M., & Reutebuch, C.K. (2008). A synthesis of fluency interventions for secondary struggling readers. *Reading and Writing*, 21(4), 317–347.
- Wiley, H.I., & Deno, S.L. (2005). Oral reading and maze measures as predictors of success for English learners on a state standards assessment. *Remedial and Special Education*, 26(4), 207–214.
- 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(4), 271–304.

CHAPTER 8

Revisiting Silent Reading in 2020 and Beyond¹

Elfrieda H. Hiebert

TextProject, & University of California, Santa Cruz

D. Ray Reutzel

Utah State University

As the title of this book indicates, our interest lies in addressing how the current knowledge base about silent reading practices can provide a foundation for future instruction and research. Based on what we know in 2015, we might ask the question, what changes in reading instruction and practice need to be made now to positively influence students' literacy proficiencies five years from now? We have chosen the year 2020 not only because it directs us into the future but also because it is the year that President Obama (Dillon, 2010) has targeted as the point when the majority of high school graduates should have the literacy skills that successfully prepare them for college and a later career.

This goal is ambitious, but if even modest movement is to be made toward achieving it, increased attention needs to be directed toward the use of effective silent reading in classrooms. In the digital-global world of the 21st century, accessing, organizing, creating, sharing, and using knowledge are critical commodities. The acquisition and use of knowledge requires that students and employees develop the ability to read silently with skill and stamina in a variety of texts for a variety of purposes, because these texts are increasingly presented to the reader using a variety of traditional and digital media. For the necessary shift from oral repeated reading with feedback to effective silent reading to occur, literacy educators need to be reflective and strategic going forward. If the researchers who revisit the topic of silent reading in 2020 are to see movement toward greater literacy capacity among elementary students and high school graduates, literacy educators will need to recognize the unique contributions and roles of both oral and silent reading in developing proficient lifelong readers.

¹ This chapter was previously published in *Revisiting Silent Reading: New Directions for Teachers and Researchers*. The definitive publisher-authenticated version published in 2010 and in 2014 is available online at: <http://www.reading.org/general/Publications/Books.aspx> & <http://textproject.org/library/books/revisiting-silent-reading/>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In this conclusion, we summarize and synthesize themes from this volume to describe the roles of both oral and silent reading in balanced and thoughtful reading instruction. We close with descriptions of three sources for effective silent reading practices that are offered in this book: (a) instructional techniques and practices, (b) teacher support, and (c) digital contexts.

Clarifying the Role of Oral Reading

In the decade since the publication of the National Reading Panel (NRP) report (National Institute of Child Health and Human Development [NICHD], 2000), reading practice has been largely confined to repeated oral reading. An overemphasis on either oral or silent reading at particular points in time is not a unique phenomenon, as described by Hiebert, Wilson, and Trainin in Chapter 3 of this volume. When viewed from the perspective of the digital age in which the selection, evaluation, and interpretation of information is paramount, however, an overemphasis on oral reading seems particularly out of sync with the needs of individuals who are prepared to participate fully in the communities and marketplace of the 21st century.

We hasten to emphasize that the near-singular attention given to repeated oral reading practice with feedback has reflected an underlying misinterpretation of the findings of research related to the role and use of silent reading in classrooms. That is, the near-exclusive emphasis on oral reading seen in many of today's classrooms resulted from several inappropriate practice conditions associated with or embedded within past silent reading and self-selected reading practice routines.

An appropriate response to the observed overemphasis on oral reading practice in the past decade in classrooms is not to overreact by moving in the opposite direction and eliminating oral reading in favor of silent reading. Oral and silent are not competing forms of reading. Rather, they are complementary forms of reading that reflect students' developmental growth as readers. When reading educators revisit the topic of silent reading in a decade or two, we would expect to see particular kinds of oral and silent reading practices used in classrooms in developmentally responsive ways (e.g., oral repeated reading with younger, less proficient readers and silent wide reading with older, more proficient readers). It is clear that most adults read silently, whereas younger readers initially enjoy reading aloud to show off their new and emerging abilities as readers.

Among several important roles that oral reading can play in the initial and later stages of reading instruction is the teacher's use of oral reading to model fluent reading and guide younger readers toward increasingly fluent oral reading. By reading aloud to younger students, teachers make an otherwise mysterious and largely invisible process more concrete and accessible. Oral reading also provides a means whereby teachers can assess and monitor students' silent reading and give them timely feedback. A well-balanced reading program offers students numerous opportunities for reading, both oral and silent. As students evidence the ability to remain involved with reading for long periods of time (i.e., stamina) and increase their reading fluency, oral reading skills can be scaffolded through gradual release by knowledgeable teachers who help students move successfully into silent reading. In the study by Reutzel, Fawson, and Smith (2008), teachers carefully scaffolded third-grade students' silent reading to ensure an effective silent reading experience.

Similarly, we would hope that upper primary and middle-grade students and beyond are not spending sizable chunks of their school days in oral round robin reading or listening to their teachers read portions of a textbook aloud. Although these uses of oral reading are typically aimed at compensating for some students' struggles in reading, such practices tend to constrain individual students' reading practice time in ways that undermine long-term reading progress. If a sufficient number of students in a class cannot read a textbook, teachers would do well to access alternative texts that struggling students can read. Neither teacher read-alouds nor oral round robin reading of textbooks is likely to lift middle school and high school students' reading achievement and prepare them for college or a career.

Oral reading is also considered by many classroom teachers to be an essential part of monitoring student progress for the purpose of designing effective instruction and interventions to increase student reading fluency and achievement. Oral reading provides teachers with a window to understand struggling students' knowledge and use of underlying systems of written language (Goodman, 1969). Classroom teachers have for decades kept running records of their students' oral reading (Clay, 1985). However, in the past decade, the assessment of oral reading has been largely founded on curriculum-based measurement models (CBM; Deno, 1985, 2003).

For oral reading, a CBM assessment consists of one-minute samples of students' rate (i.e., how many words they read) and accuracy in reading a passage. For silent reading, the task involves reading a text in

which words have been systematically deleted and several choices given for the deleted word. This latter design is intended to establish students' rate of reading with comprehension. For classroom teachers faced with many students and limited time, CBM assessments are an efficient way of gathering information on students' oral and silent reading.

Unfortunately, the data drawn from the CBM assessments have been used in inappropriate ways, such as to designate students in different risk status levels—benchmarked, strategic, or intensive. Sadly, these one-minute CBM assessments have led to an overemphasis on reading speed at the expense of developing expression and comprehension with both native English speakers and English learners.

To ensure good student performance on such assessments, teachers may have students spend an excessive amount of instructional time reading short paragraphs and texts to increase reading speed. This inevitably leads to students who lack reading stamina because they are used to practicing their reading in a sprint-like fashion for short periods. Of course, oral reading norms obtained from these one-minute samples are also likely to overestimate real sustained reading speeds orally or silently, because long-distance runners pace themselves differently compared with sprinters (see Chapter 3 by Hiebert et al.). Emphasizing sprinting over long-distance reading can have particularly devastating consequences on the reading development of those students who have reading disabilities.

Even more disappointing is the fact that these quick CBM assessments have also displaced more intensive and comprehensive examinations of struggling readers' oral reading miscues and behaviors. We argue that high-quality reading assessment should not be dismissed or displaced because of inappropriate applications of CBM and overuses of oral reading fluency measures during the past decade. Whether the displaced assessment was an informal reading inventory with leveled texts or a running record taken while students read everyday texts in the classroom, sampling students' oral reading for insight into their linguistic knowledge and their use (or lack) of monitoring and fix-up strategies was part of the assessment repertoire for many past generations of teachers (Pikulski & Shanahan, 1982).

Oral reading plays yet another role in classroom reading instruction in that it is the means whereby students can enjoy literary favorites and classics through teacher read-alouds. Digital texts have increased student access to performances by great story readers such as Jim Dale reading the Harry Potter books. To create community,

offering a setting for expression and presenting a stage for performing and entertaining, oral reading is central (Rasinski & Griffith, 2008). Opportunities for students to select portions of texts or poems to read aloud or participate in Reader's Theater contribute to the development of a classroom literacy community that is vibrant and alive (Wolf, 2004). Teacher read-alouds allow students to experience language and vocabulary they may not yet be able to process on their own. Some research has shown that oral reading fluency correlates well with specific forms of silent reading comprehension (Schatschneider, Torgesen, Buck, & Powell-Smith, 2004). Although oral reading serves several critical functions as described previously, it is erroneous to assume that oral reading proficiency equals silent reading proficiency. There are significant differences between the two processes. These differences are especially pronounced when students read texts presented in digital formats. Sifting through information, deciding what is credible, and choosing how to communicate one's response to information are not typically oral reading processes.

Clarifying the Role of Silent Reading

For almost 40 years prior to the NRP report (NICHD, 2000), teacher educators and staff development specialists routinely recommended independent silent reading practices such as those promoted by Hunt (1970) under the aegis of sustained silent reading (SSR). When using these models of silent reading, teachers were advised to allow students to read silently for extensive periods of time, regardless of their grade level or their proficiency levels. Even though evidence supporting the benefit of spending large chunks of class time on students' self-selected reading was anything but convincing, whole-language proponents in the mid-1980s began advocating for independent silent reading practice to replace core reading instruction programs and oral reading (Hagerty, 1999). In readers' workshops that extended SSR practices to classroom instruction, students—even first graders—were encouraged to choose their own books, often without much teacher guidance or assistance. Reading instruction consisted of brief, randomly sequenced, or incidental whole-class mini lessons and, in rare cases, individual teacher–student conferences.

Recommendations such as these ignored research on silent reading practices that existed at that time. Several projects in the 1980s pointed to the need to adapt silent reading practices to increase greater student accountability and monitoring by teachers in book selection and purpose

setting (Anderson, Wilson, & Fielding, 1988; Manning & Manning, 1984). However, as Stahl (1999) observed, these recommendations were often ignored in whole-language classrooms. Many teachers did not hold conferences with students during independent silent reading, opting instead to read independently and silently themselves, thereby believing themselves to be a model of engaged silent reading.

After nearly 40 years of continuously recommending independent silent reading, the NRP report (NICHD, 2000) cast a shadow of doubt over the practices associated with it, such as SSR. In a meta-analytic review of the research on SSR within classroom settings, NRP members were able to identify 10 SSR studies that met their criteria of rigorous research. In five of those studies, researchers found effects that favored SSR. However, the effect sizes were relatively small. Subsequent analyses of the 10 studies (Lewis, 2002; Reutzel et al., 2008; Wu & Samuels, 2004) noted limitations in their designs and executions. For example, the studies as a group did not report precisely how much time was spent in reading. The die was cast, however, when NRP members (e.g., Shanahan, 2006) strongly suggested, independent of the report, that evidence for time spent on independent silent reading in classrooms—compared with other reading approaches such as guided, repeated, and oral reading with feedback—was not as effective.

Rather than constructively addressing the misinterpretations of the NRP's (NICHD, 2000) concerns, advocates of independent silent reading sharply criticized the conclusions drawn by the NRP (Coles, 2000; Krashen, 2001, 2005). Since that time, several research groups (most represented by chapters in this volume) have reconsidered what it takes to get and keep students' eyes on the page during silent reading. From this sustained research, we describe three features of independent silent reading practice that require attention to improve the silent reading performance of elementary and secondary students in the future.

Instructional Techniques and Practices

Although current research on independent silent reading is not as extensive, or the findings as robust, as those surrounding phonemic awareness and alphabets, there is an emerging research base that indicates that there are specific elements of classroom reading programs that can support the development of proficient silent reading habits.

To understand how these elements can be influenced by teachers and instruction, one needs to understand what distinguishes silent

reading from oral reading. Moving from vocalization or subvocalization to “silent reading” is perhaps one of the hardest aspects of reading there is to “teach” (Wright, Sherman, & Jones, 2010). For most students, this happens gradually with ample opportunities to read. There are also developmental and social factors that likely influence movement toward effective and sustained silent reading. However, if students do not have frequent opportunities to read, there may well be residual subvocalization behaviors that often characterize the silent reading habits of struggling readers. Wright, Sherman, and Jones indicated that teachers need to do more than assign texts for students to read silently and tell them to stop whisper reading.

Samuels, Hiebert, and Rasinski (see Chapter 2) noted that some students may require carefully designed instructional programs to remediate or develop the eye movements that characterize proficient reading. Such programs have yet to be validated by sustained and carefully designed research studies that address how to efficiently train eye movements—and the subsequent effects of doing so on students’ reading automaticity and comprehension. In particular, the success of eye-movement training programs designed to support efficiency in silent reading as discussed in Chapter 2 needs to be disentangled from the eyes-on-the-text phenomenon conflated with current models of eye-movement training.

Proficient silent reading also requires that individuals be able to independently manage their attention. Unlike oral reading where there is a definite task (and a monitor in the form of an adult or a recording device), silent reading requires that readers choose to remain involved in reading, manage their time well, and take steps to correct or fix up failing comprehension when necessary. For example, students might struggle with the decision to keep their eyes on the text instead of skimming or scanning the text or acting like they are reading. As the findings described by Hiebert et al. (see Chapter 3) suggest, perseverance or reading stamina appears to be a considerable challenge for less proficient readers (see also Lee-Daniels & Murray, 2000). When reading silently, students must make internal choices they do not have to make during oral reading.

Of these reader behaviors that are unique to silent reading—managing one’s time, choosing to remain engaged in reading a text, and monitoring and fixing up faulty comprehension—only the topic of monitoring strategies has received much focus (Pearson & Dole, 1987). Monitoring strategies become particularly critical when readers’ purposes are vague or ill-defined and when background knowledge is limited—

circumstances that are often a part of silent reading for less proficient readers (McKeown, Beck, & Blake, 2009). As is often the case, published programs have overdone the number of strategies that are taught and practiced as part of reading comprehension lessons (Block & Duffy, 2008; Dewitz, Jones, & Leahy, 2009). Furthermore, past instruction has not aided students in knowing under what conditions particular strategies are useful or imperative and under what conditions they are not.

Although the work is still limited in scope, we are coming to understand how the stamina of readers can be supported by effective independent silent reading practice conditions. The authors of Chapters 5 and 6 in this volume illustrate several scaffolds needed to increase the perseverance of students' silent reading during allocated independent reading time. What seems to be particularly important is that students not be permitted to flit about selecting texts but be required to complete texts and illustrate their understanding of the content. White and Kim (2010) also argued for involving students in programs that support after-school and summer reading. They also emphasized the need to provide silent reading scaffolds that support engagement with books and the development of reading stamina. What is clear from this group of studies is that Hunt's (1970) suggestion that the same silent reading program be implemented across different developmental and proficiency levels misrepresented the complexity of reading, texts, classrooms, and instruction. Although the simplistic message that all students should read silently in self-selected books may have been a point of departure, there is much that we have learned in the interim about the kinds of scaffolds that can ensure that students increase their capacity and interest in silent reading.

Teacher Support

Change of any kind takes time and information. Fundamental changes in silent reading practices in classrooms can be expected to require substantial amounts of support for the teachers who will be asked to make them. As the teachers' questions—which provided the basis for Hiebert et al.'s (Chapter 3) development of comprehension-based silent reading rate—illustrate, teachers ask many important questions. Often, these are questions for which researchers have few solid answers. Conversations between researchers and teachers are urgently needed on issues associated with independent silent reading so that the questions that teachers ask are addressed by future research and so that the questions that

researchers pursue in relation to independent silent reading are relevant to the real world of classrooms.

The amount of teacher support and scaffolding required to sustain silent reading practice among students in classroom settings can be extensive. Brenner and Hiebert (2010) described a series of modules designed to provide teachers with specific professional development on how to help students keep their eyes on the page during silent reading. Even with access to these modules and on-site coaching from peers, the teachers described in the study nevertheless required a great deal of continuous support to make even small changes in relation to supporting effective independent silent reading practices in their own classrooms.

Digital Contexts

In the digital-global world of the 21st century, proficient silent reading is essential to meeting the challenge of ensuring that more high school graduates are ready for the increasing demands of college and career-related literacy tasks. Literacy proficiencies in traditional print contexts do not necessarily extend seamlessly to those practiced in digital contexts. Effective silent reading in online contexts requires that students adopt a problem-solving stance, where an initial task involves searching for and selecting from available information and a second involves evaluating whether the accessed information is valid and valuable to read. The texts in these tasks are almost always informational in nature, whereas much of past conventional print-based reading instruction has focused heavily on traditional print versions of narrative texts.

Informational and narrative texts differ in structure, conceptual density, and physical features such as diagrams, photo inserts, headings and subheadings, and a table of contents (Duke & Bennett-Armistead, 2003). Readers often skim sections of an informational text, but closely read and reread those sections that provide the precise content they are seeking. In contrast, narrative texts are typically written to be read from beginning to end with a relatively uniform amount of focused attention.

Despite the fact that digital contexts have made the demands for processing informational texts more critical, there is evidence that opportunities for content area learning in elementary schools have decreased rather than increased. In a recent survey, elementary teachers reported devoting around an hour of time weekly to science instruction (Dorph et al., 2007). This amount of time is half of what was reported in a survey conducted in 2000 (Fulp, 2002). If students have not had adequate

experiences with informational text, they are in considerable jeopardy when faced with the additional requirements needed to be successful in negotiating literacy tasks in digital contexts.

The digital age offers considerable opportunities for learners. For educators to ensure that students have the skills that allow them to take advantage of these opportunities, a massive restructuring of the literacy curriculum needs to happen. Support for strong silent reading comprehension is fundamental to this restructuring, but it is not simply a matter of increasing silent reading practice with the texts and processes that have dominated the curriculum. This restructuring requires significant changes to the texts and contexts of instruction as well.

Final Thoughts

If we are to be successful in promoting efficacious silent reading over the next decade, educators need to be more strategic and thoughtful. Unexamined assumptions associated with past independent silent reading practices have led to results that, in the long run, have not supported students in becoming more proficient independent, silent readers. Furthermore, privileging oral reading over silent reading in instruction had not resulted in students transferring oral reading skills to silent reading.

Oral and silent reading both have critical roles in the development of proficient reading. Failing to view oral and silent reading as having complementary rather than competing functions in the development of proficient literacy could jeopardize the futures of our students. Teachers and researchers need to work together to solve the conundrums around how best to support all readers through appropriate uses of both oral and silent reading at different points in students' literacy development.

References

- Anderson, R.C., Wilson, P.T., & Fielding, L.G. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly*, 23(3), 285–303. doi:10.1598/RRQ.23.3.2
- Block, C.C., & Duffy, G.G. (2008). Research on teaching comprehension: Where we've been and where we're going. In C.C. Block & S.R. Parris (Eds.), *Comprehension instruction: Research-based best practices* (2nd ed., pp. 19–37). New York, NY: Guilford.
- Brenner, D., & Hiebert, E.H. (2010). The impact of professional development on students' opportunity to read. In E.H. Hiebert & D. Ray Reutzel (Eds.), *Revisiting Silent Reading: New Directions for Teachers and Researchers* (pp. 198–217). Newark, DE: IRA.
- Clay, M.M. (1985). *The early detection of reading difficulties* (3rd ed.). Portsmouth, NH: Heinemann.
- Coles, G. (2000). *Misreading reading: The bad science that hurts children*. Portsmouth, NH: Heinemann.
- Deno, S.L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52(3), 219–232.
- Deno, S.L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education*, 37(3), 184–192.
- Dewitz, P., Jones, J., & Leahy, S. (2009). Comprehension strategy instruction in core reading programs. *Reading Research Quarterly*, 44(2), 102–126. doi:10.1598/RRQ.44.2.1
- Dillon, S. (2010, March 13). Obama calls for major change in education law. *The New York Times*. Retrieved from www.nytimes.com/2010/03/14/education/l4child.html
- Dorph, R., Goldstein, D., Lee, S., Lepori, K., Schneider, S., & Venkatesan, S. (2007). *The status of science education in the Bay Area*. Berkeley, CA: Lawrence Hall of Science, UC-Berkeley.
- Duke, N.K., & Bennett-Armistead, V.S. (2003). *Reading & writing informational text in the primary grades: Research-based practices*. New York, NY: Scholastic.
- Fulp, S.L. (2002). *Status of elementary school science teaching*. Chapel Hill, NC: Horizon Research.
- Goodman, K.S. (1969). Analysis of oral reading miscues: Applied psycholinguistics. *Reading Research Quarterly*, 5(1), 9–30. doi:10.2307/747158
- Hagerty, P.J. (1999). *Readers' workshop: Real reading*. New York, NY: Scholastic.
- Hunt, L.C. (1970). The effect of self-selection, interest, and motivation upon independent, instructional, and frustration levels. *The Reading Teacher*, 24(2), 146–151.
- Krashen, S. (2001). More smoke and mirrors: A critique of the National Reading Panel report on fluency. *Phi Delta Kappan*, 83(2), 119–123.
- Krashen, S. (2005). A special section on reading research: Is in-school free reading good for children? Why the National Reading Panel Report is (still) wrong. *Phi Delta Kappan*, 86(6), 444–447.
- Lee-Daniels, S.L., & Murray, B.A. (2000). DEAR me: What does it take to get children reading? *The Reading Teacher*, 54(2), 154–155.
- Lewis, M. (2002). *Read more-read better: A meta-analysis of the literature on the relationship between exposure to reading and reading achievement*. (Unpublished doctoral dissertation). Minneapolis/St. Paul, MN: University of Minnesota.
- Manning, G.L., & Manning, M. (1984). What models of recreational reading make a difference? *Reading World*, 23(4), 375–380.
- McKeown, M.G., Beck, I.L., & Blake, R.G.K. (2009). Rethinking reading comprehension instruction: A comparison of instruction for strategies and content approaches. *Reading Research*

Quarterly, 44(3), 218–253. doi:10.1598/RRQ.44.3.1

National Institute of Child Health and Human Development. (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. 00-4769). Washington, DC: U.S. Government Printing Office.

Pearson, P.D., & Dole, J.A. (1987). Explicit comprehension instruction: A review of research and a new conceptualization of instruction. *The Elementary School Journal*, 88(2), 151–165. doi:10.1086/461530

Pikulski, J., & Shanahan, T. (1982). Informal reading inventories: A critical analysis. In J. Pikulski & T. Shanahan (Eds.), *Approaches to the informal evaluation of reading* (pp. 94–116). Newark, DE: International Reading Association.

Rasinski, T., & Griffith, L. (2008). *Building fluency through practice and performance grade 3*. Huntington Beach, CA: Shell Education.

Reutzel, D.R., Fawson, P.C., & Smith, A. (2008). Reconsidering silent sustained reading: An exploratory study of scaffolded silent reading. *The Journal of Educational Research*, 102(1), 37–50. doi:10.3200/JOER.102.1.37-50

Schatschneider, C., Torgesen, J.K., Buck, J., & Powell-Smith, K. (2004). *A multivariate study of factors that contribute to individual differences in performance on the Florida Comprehensive Reading Assessment Test* (Tech. Rep. No. 5). Tallahassee, FL: Florida Center for Reading Research.

Shanahan, T. (2006). Does he really think kids shouldn't read? *Reading Today*, 23(6), 12.

Stahl, S.A. (1999). Why innovations come and go (and mostly go): The case of whole language. *Educational Researcher*, 28(8), 13–22.

White, T.G., & Kim, J.S. (2010). Can silent reading in the summer reduce socioeconomic differences in reading achievement? In E.H. Hiebert & D. Ray Reutzel (Eds.), *Revisiting silent reading: New directions for teachers and researchers* (pp. 67–91). Newark, DE: IRA.

Wolf, S.A. (2004). *Interpreting literature with children*. Mahwah, NJ: Erlbaum.

Wright, G., Sherman, R., & Jones, T.B. (2010). Developmental considerations in transferring oral reading skills to silent reading. In E.H. Hiebert & D. Ray Reutzel (Eds.), *Revisiting silent reading: New directions for teachers and researchers* (pp. 57–66). Newark, DE: IRA.

Wu, Y., & Samuels, S.J. (2004). *How the amount of time spent on independent reading affects reading achievement*. Paper presented at the annual meeting of the International Reading Association, Reno, NV.

ISBN: 978-1-937889-04-3