

THE RELATIONSHIP BETWEEN DELIBERATE PRACTICE AND READING ABILITY

Sean Thomas Hanlon

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Education.

Chapel Hill
2013

Submitted to:

Jeffrey A. Greene

Gregory J. Cizek

Jill W. Fitzgerald

Judith L. Meece

Carl W. Swartz

©2013
Sean Thomas Hanlon
ALL RIGHTS RESERVED

ABSTRACT

SEAN T. HANLON: The Relationship between Deliberate Practice and Reading Ability
(Under the direction of Jeffrey A. Greene)

Many students are not prepared to meet the literacy demands of college and career as defined by the Common Core State Standards (2010). Literacy researchers have struggled to define the frequency and type of reading practice necessary to nurture the development of reading ability. The principles of deliberate practice provide a theoretical framework that could describe the type of practice necessary to develop expertise in reading. The purpose of this study was to explore the relationship between deliberate practice and reading ability. In this study, an educational technology, *Learning Oasis*, was used to deliver deliberate practice and monitor change in student reading ability over time. The hypotheses were that participants that engaged in more deliberate practice, as operationalized in this study, would experience more rapid growth and achieve higher levels of reading ability. Participants in this study ($N = 1,369$) ranged from grades one through twelve and were from a suburban school district in Mississippi. Each participant had at least three measurement occasions separated by at least three months each. The Lexile Framework for Reading was used to estimate participant reading ability during this research. Given the longitudinal nature of the data, a multilevel model was used to explore individual change over time. A negative exponential functional form was determined to best model change in participant reading ability over time. The results showed that, on average, participants that engaged in more deliberate practice (i.e., targeted

practice with immediate feedback completed intensely over a long period of time) grew more rapidly and reached a higher ability level than participants that completed less deliberate practice. Implications for educators, educational technology designers, and researchers are discussed along with potential future areas of research.

ACKNOWLEDGMENTS

Without the contribution and support of many, the tome you now hold in your hands would never have been completed. While it is impossible to acknowledge everybody who offered well-timed advice or words of encouragement, certain individuals require special mention. First and foremost, I extend my gratitude to the students and educators that participated in this research, particularly superintendent “Dr. C.” It has been a long road since that rainy morning in 2006 when two guys from North Carolina showed up in the computer lab at the high school. Thank you for allowing us to be part of the school day; it is a responsibility we continue to take very seriously.

Throughout my time at the University of North Carolina at Chapel Hill, I have been fortunate to work with brilliant scholars. To Lara Jean Costa, I offer my thanks for years of friendship through all manner of shared experiences. To my various committees, I offer my thanks for your scholarship, honesty, and thoughtful feedback. My thanks to Dr. Jeffrey A. Greene, Dr. Gregory J. Cizek, Dr. Jill W. Fitzgerald, Dr. Judith L. Meece, Dr. A. Jackson Stenner, and Dr. Carl W. Swartz.

It is also important to recognize the support I have received from my colleagues at MetaMetrics. In particular, I offer my deepest thanks to Dr. Gary L. Williamson and Juee Tendulkar for their technical expertise, timely feedback, and ongoing support. I hope this research reflects the time, talent, and effort the entire Learning Science and Technology group

(i.e., the group formerly known as New Technologies) devoted to bringing our educational technology to students. To past and present members of the Learning Science and Technology team, I offer my thanks. I also offer my gratitude to other members of the MetaMetrics family that have offered support in a variety of ways through the years. In particular, my thanks to Harold J. Burdick, Dr. Donald S. Burdick, Dr. Malbert Smith III, Timothy J. Klasson, Patricia Carideo, and Chris Whyte.

A special thank you to my advisor and dissertation chair Dr. Jeffrey A. Greene. Jeff, we have come a long way since our collaboration began, and I hope every student receives the high-quality advising and mentoring that you have provided for me over the years. While we still might disagree on certain “stylistic features,” I know your attention to detail and constructive criticism has made me a better thinker and writer. Your guidance on a variety of topics, both academic and otherwise, has been very much appreciated.

The decision to return to graduate school, particularly to leave computer science and embark on a Ph.D. program in educational psychology, was not taken lightly. To that end, I offer a sincere and heartfelt thank you to Dr. A. Jackson Stenner and Dr. Carl W. Swartz for their support and guidance. Jack, I have always appreciated the unique perspective and unflappable demeanor that you bring to a discussion, regardless of the topic or venue. Your honesty and encouragement has always meant a great deal to me. Carl, since our collaboration began we have shared three different offices, released and supported at least eight different classroom-based technologies, listened to Freddy C. and the rest of the ESPN radio cast over incalculable miles in a rental car, racked up enough frequent flyer miles to travel around the world, visited most of the Outback restaurants around the country, and frequented Starbucks

enough times that I think we receive mail there. I continue to learn a great deal from you, and I very much look forward to our continued collaboration.

I offer my appreciation to my family for their love and support throughout the course of this Ph.D. program. I would like to thank the MacDonell family (Cindy, Dave, Scott, and Julie) and Karen and Ryan Hanlon for their kind words of support. To Amanda Smith and Mark Hanlon, I thank them for their true-blue support despite their observation that a 282 page document is really too long to read, let alone write. To my parents, Barbara and Thomas Hanlon, I extend my thanks for a lifetime of encouragement and guidance. It seems like we have come a long way from the EPCOT parking lot and my future career in data entry. Last but certainly not least, I offer my heartfelt thanks to my wife, Kate. J.R.R.Tolkien said “it’s the job that’s never started that takes the longest to finish,” and your limitless encouragement to not only begin a Ph.D. program, but to stay the course and finish was very much appreciated. I think it is safe to say that we can consider this a terminal degree. Thanks for all you do each and every day.

TABLE OF CONTENTS

LIST OF TABLES.....	x
LIST OF FIGURES.....	xii
Chapter	
I. INTRODUCTION.....	1
II. REVIEW OF THE LITERATURE.....	11
The development of expertise.....	11
Theories of reading.....	39
Technology and 21 st century learning and research.....	77
Deliberate practice, theories of reading, and educational technology.....	84
III. METHODS.....	89
Participants.....	89
Data sources.....	91
Procedures.....	92
Instrumentation.....	97
Data analysis plan.....	111
IV. RESULTS.....	131
Exploratory analysis.....	131
Fitting a taxonomy of models.....	138
Hypotheses revisited.....	158

	Deliberate practice as a whole.....	168
	Examination of assumptions.....	168
	Summary of results.....	172
V.	DISCUSSION.....	173
	Deliberate practice and reading ability.....	174
	The role of educational technology.....	183
	Limitations.....	185
	Future research.....	190
	Implications for educators, educational technology designers, and researchers.....	195
	Next steps.....	198
	Conclusion.....	202
	APPENDICES.....	203
	REFERENCES.....	225

LIST OF TABLES

Table

1. Time spent reading daily and words read annually in relation to percentile rank on standardized reading tests.....	3
2. Operational definition of deliberate practice.....	8
3. Qualifying participants by grade.....	90
4. Demographic information for all participants participating in study ($N = 1,369$).....	91
5. Test length of a sample of high-stakes reading assessments.....	101
6. Operational definition of deliberate practice.....	103
7. Grade-based standard-length words per minute (Wpm) ranges.....	109
8. Values in person-period data set.....	118
9. Descriptive statistics for the individual growth parameters obtained by fitting separate within-person OLS regression models ($N = 1,369$).....	133
10. Descriptive statistics for amount of reading and components of deliberate practice..	134
11. Correlations between predictors of interest.....	137
12. Taxonomy of multilevel models for change fitted to participant reading data.....	139
13. Unstandardized fixed effects: Results of fitting a taxonomy of multilevel models for change to participant reading data ($N = 1,360$).....	142
14. Variance components and goodness-of-fit indices: Results of fitting a taxonomy of multilevel models for change to participant reading data ($N = 1,360$).....	144
15. Model comparisons made during fitting taxonomy of models to participant reading data using both Likelihood Ratio Test (LRT) and Wald statistics.....	149
16. VIF statistics for potentially collinear predictors and goodness-of-fit indices for select models.....	152
17. Fixed effects and VIF values for Model G, Model H, and Model I ($N = 1,360$).....	155

18. Variance components and goodness-of-fit indices for Model D, Model G, Model H, and Model I.....	157
19. Observed variation in fitted OLS linear trajectories: Fitted initial status (Lexile), rate of change (Lexile per day), correlation between initial status and rate of change, and R^2 statistics for sample of participants stratified by predictors-of-interest.....	210
20. Observed variation in fitted OLS quadratic trajectories: Fitted initial status (Lexile), rate of change (Lexile per day), correlation between initial status, rate of change, and quadratic term, and R^2 statistics for sample of participants stratified by predictors-of-interest.....	211
21. Means and standard deviations for reading activity completed by participants, by initial grade.....	213
22. Means and standard deviations for reading activity completed by participants, stratified by initial reading ability and lunch status.....	214

LIST OF FIGURES

Figure

1. Cloze items generated by <i>Learning Oasis</i>	94
2. Example item generated by <i>Learning Oasis</i>	95
3. A correctly answered cloze item inside <i>Learning Oasis</i>	96
4. An incorrectly answered cloze item inside <i>Learning Oasis</i>	97
5. Prototypical negative exponential with asymptote depicted.....	123
6. Histogram of measurement occasions.....	135
7. Plot of three prototypical participants. Left panel shows varying the number of items answered. Right panel shows varying the number of words read.....	151
8. Plot of three prototypical participants, varying only number of words read.....	160
9. Plot of three prototypical participants, varying only intense minutes spent reading per day.....	162
10. Plot of three prototypical participants, varying only number of days of reading.....	163
11. Plot of three prototypical participants, varying only standard deviation of elapsed days between reading experiences.....	164
12. Plot of three prototypical participants, varying only initial Lexile reading level.....	166
13. Plot of two prototypical participants, varying lunch status.....	167
14. Plot of three prototypical participants with <i>InitialLexile</i> of 431L and varying in amount <i>Words</i> , <i>IntenseMinPerDay</i> , <i>DaysUsed</i> , and <i>ElapsedDaysSD</i>	169
15. Plot of three prototypical participants with <i>InitialLexile</i> of 780L and varying in amount <i>Words</i> , <i>IntenseMinPerDay</i> , <i>DaysUsed</i> , and <i>ElapsedDaysSD</i>	170
16. Plot of three prototypical participants with <i>InitialLexile</i> of 1129L and varying in amount <i>Words</i> , <i>IntenseMinPerDay</i> , <i>DaysUsed</i> , and <i>ElapsedDaysSD</i>	171
17. Empirical growth plots with superimposed OLS linear trajectories: Random participants from entire sample ($N = 1,369$).....	204

18. Empirical growth plots with superimposed OLS quadratic trajectories: Random participants from entire sample ($N = 1,369$).....	205
19. Empirical growth plots with superimposed Nonlinear Least Squares negative exponential trajectories: Random participants from entire sample ($N = 1,369$).....	206
20. Collection of fitted OLS linear trajectories for all participants in sample ($N = 1,369$). White line represents the average change trajectory for the entire group.....	207
21. Collection of fitted OLS quadratic trajectories for all participants in sample ($N = 1,369$). White line represents the average change trajectory for the entire group.....	208
22. Collection of fitted Nonlinear Least Squares negative exponential trajectories for all participants in sample ($N = 1,369$). White line represents the average change trajectory for the entire group.....	209
23. Examining normality assumptions in the multilevel model for change: Normal probability plots for the raw residuals at level-1 and level-2.....	216
24. Examining normality assumptions in the multilevel model for change: Plots of standardized residuals at level-1 and level-2 versus participant ID numbers.....	217
25. Examining the homoscedasticity assumptions in the multilevel model for change: Raw level-1 residuals vs. level-1 predictors for <i>Initial Lexile</i>	218
26. Examining the homoscedasticity assumptions in the multilevel model for change: Raw level-2 residuals vs. level-2 predictor <i>Words</i>	219
27. Examining the homoscedasticity assumptions in the multilevel model for change: Raw level-2 residuals vs. level-2 predictor <i>DaysUsed</i>	220
28. Examining the homoscedasticity assumptions in the multilevel model for change: Raw level-2 residuals vs. level-2 predictor <i>IntenseMinPerDay</i>	221
29. Examining the homoscedasticity assumptions in the multilevel model for change: Raw level-2 residuals vs. level-2 predictor <i>ElapsedDaysSD</i>	222
30. Examining the homoscedasticity assumptions in the multilevel model for change: Raw level-2 residuals vs. level-2 predictor <i>InitialLexile</i>	223
31. Examining the homoscedasticity assumptions in the multilevel model for change: Raw level-2 residuals vs. level-2 predictor <i>SES</i>	224

CHAPTER 1

INTRODUCTION

In his State of the Union Address, President Obama (2011) provided a sobering vision of the employment demands Americans will encounter in the next decade: “Nearly half of all new jobs will require education that goes beyond a high school education. And yet, as many as a quarter of our students aren’t even finishing high school” (para. 12-34). President Obama’s comments highlighted the growing body of evidence that has shown students are not prepared for life after high school (Bock & Moore, 1986; Bureau of Labor Statistics, 2007; Common Core State Standards, 2010; Dohm & Shniper, 2007; Greene & Forster, 2003; Stillwell, 2010; Sum, 1999; U.S. Department of Labor, 1991). High school graduates are not prepared to meet the literacy demands of college and career (Common Core State Standards, 2010; MetaMetrics, 2008; Williamson, 2008; Wirt et al., 2004). Thirty-two percent of high school graduates are not ready for college-level English composition courses (ACT, 2005) while approximately 40 percent of high school graduates lack the literacy skills sought by employers (Achieve, Inc., 2005). Lacking in reading ability can be costly for high school graduates in terms of college success as well as gaining and maintaining employment.

The gap between students’ reading ability at the conclusion of high school and the complexity of the text encountered after graduation must be closed if students are to succeed in the post-secondary world (Common Core State Standards, 2010; Williamson, 2008).

Reading is one of the most-often studied phenomena investigated by educational researchers, and a multitude of different theoretical models have been proposed (Bernard, 2000; Best & Kahn, 1998; Stanovich, 1992; Thomas, 1996; Tierney, 1994; Tracey & Morrow, 2006). The architects of these models have conceptualized and approached reading in a variety of ways. For example, some have argued that reading is the process of deciphering individual words correctly and accurately from groups of letters (Adams, 1990; Crowder & Wagner, 1992; Ehri, 1998; Perfetti, 1985; Rayner & Pollatsek, 1989; Snow, Burns, & Griffin, 1998). In contrast, another group of researchers have argued that reading is the comprehension, or understanding, that results from the interaction between written words and a reader (Goodman, 1967; Fountas & Pinnell, 2001; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001; Smith, 1971, 1973; Thorndike, 1917).

Regardless of definitions, the adage “practice makes perfect” feels logical when considering the development of reading ability (Gambrell, 2007). Allington (1977) argued that the amount of reading influences student reading performance, and other researchers have supported the importance of reading volume on the development of reading ability (Allington, 1980, 1983, 1984a; Anderson, Wilson, & Fielding, 1988; Cunningham & Stanovich, 1998; Gambrell, 1984; Hiebert, 1983; Knapp, 1995; Krashen, 2004; Meyer & Wardrop, 1994; Stanovich, 2000; Stanovich, West, Cunningham, Cipielewski, & Siddiqui, 1996; Thurlow, Gaden, Ysseldyke, & Algozzine, 1984; Vaughn, Moody, & Schumm, 1998; Wu & Samuels, 2004). The research of Anderson et al. showed correlational evidence that students who read more words annually scored higher on standardized reading tests (see Table 1). Students in the 98th percentile read 4.3 million words, spending approximately 65 minutes reading per day. On the other hand, students in the 50th percentile read 282,000

words annually and spent less than five minutes reading per day. Unfortunately, on average, students are not spending large amounts of time reading (National Endowment for the Arts, 2007).

Table 1. Time spent reading daily and words read annually in relation to percentile rank on standardized reading tests¹.

Percentile Rank	Minutes Reading per Day	Words read per year
98	65	4,358,000
90	21.1	1,823,000
80	14.2	1,146,000
70	9.6	622,000
60	6.5	432,000
50	4.6	282,000
40	3.2	200,000
30	1.8	106,000
20	0.7	21,000
10	0.1	8,000
2	0	0

¹Adapted from Anderson, Wilson, & Fielding (1988).

The National Reading Panel (NRP; National Institute of Child Health and Human Development [NICHD], 2000) challenged the importance of reading quantity, observing a lack of compelling empirical evidence:

Even though encouraging students to read more is intuitively appealing, there is still no sufficient research evidence obtained from studies of high methodological quality to support the idea that such efforts reliably increase how much students read or that such programs result in improved reading skills (pp. 12-13).

After publication, two members of the NRP observed that the report did not oppose encouraging students to read, only that few well-designed studies of the influence of reading volume were available (Samuels, 2002; Shanahan, 2004).

As suggested in the NRP (2000) report, previous research into the influence of reading volume on the development of reading ability has suffered from methodological limitations. In particular, researchers have struggled to quantify the amount of reading completed and estimate changes in student reading ability. A variety of researchers have relied on self-report estimates of reader activity (e.g., Ennis, 1965; Nell, 1988; Stanovich & West, 1989; Wagner & Stanovich, 1996; Walberg & Tsai, 1984). While easy to administer, questionnaires that ask children or their parents about reading behavior suffer from poor reliability and overestimates of actual reading activity as students often inflate their estimate of the amount of reading completed (Cunningham & Stanovich, 1991; Ennis, 1965; Sharon, 1973). Other researchers asked readers to complete diaries describing their reading activity (Allen, Cipielewski, & Stanovich, 1992; Anderson et al., 1988; Greaney, 1980; Greaney & Hegarty, 1987). While this method provides more reliable results compared to a questionnaire, as readers are keeping an on-going record immediately following a reading experience, diaries are susceptible to social pressure and require commitment and effort (Wagner & Stanovich, 1996). Questionnaires and diaries can be useful, but are limited in their ability to quantify the amount of reading practice completed by a student due to the self-report nature of the estimates (Allington, 2009; Wagner & Stanovich, 1996).

Allington (2009) believed that reading volume certainly influenced reading development; however, he observed a lack of evidence related to the amount and type of practice necessary to nurture reading ability. While investigating the development of

expertise, Ericsson, Krampe, and Tesch-Romer (1993) proposed that it is not the amount of practice so much as the amount of deliberate practice that separates experts from novices. Deliberate practice requires activity that is specifically designed to challenge the learner and improve performance. Practice designed to be deliberate can be characterized by five principles: (1) targeted activity that is designed to appropriately challenge the learner, (2) real-time corrective feedback that provides an indicator of performance, (3) distributed practice over a long period of time, (4) intensive practice that does not require the learner to concentrate beyond his or her limits, and (5) self-directed practice when a teacher or coach is unavailable (Ericsson, 1996a, 1996b, 2002, 2004, 2006a, 2006b; Ericsson et. al, 1993). Deliberate practice has been shown to foster expertise in a variety of domains: chess (e.g., Charness, Krampe, & Mayer, 1996; Charness, Tuffiash, Krampe, Reingold, & Vasyukova, 2005; Gobet & Charness, 2006), decision-making in a crisis (e.g., McKinney & Davis, 2004), study habits (e.g., Plant, Ericsson, Hill, & Asberg, 2005), mathematical computation (e.g., Butterworth, 2006), professional writing (e.g., Kellogg, 2006), and sports (e.g., Helsen, Starkes, & Hodges, 1998; Hodges & Starkes, 1996; Starkes, Deakins, Allard, Hodges, & Hayes, 1996). The lack of evidence relating growth in reading ability to reading practice could be an artifact of poorly understanding and capturing the nature and quality of the practice.

While the principles of deliberate practice provide a potential explanatory mechanism for the development of expertise in reading, the methodological limitations of traditional reading research complicate investigations into the relationship between deliberate practice and reading ability. In addition to the challenges associated with using self-report measures (e.g., questionnaires, diaries) to track the amount of reading performed, estimating the

change in student reading ability over time presents unique difficulties. Traditionally, quantifying growth in reading ability was accomplished using a small number of measurement occasions, such as a pretest and a posttest (e.g., Kim, 2006, 2007; Kim & Guryan, 2010; Kim & White, 2008). While this approach provides estimates of student growth in reading ability, it does not provide information about the growth trajectory (e.g., slope, intercept). A growth trajectory represents the shape of student change over time, providing information about the nature (i.e., functional form) of student change as opposed to a change score (e.g., pretest/posttest results) (Rogosa, Brandt, & Zimowski, 1982; Willett, 1988).

Fortunately, technological advances can help researchers overcome the limitations of past research into student growth (i.e., few measurement occasions). *Educational technology*, any software (e.g., online learning systems, games, simulations, digital tutors) or hardware (e.g., desktop computers, laptops, mobile devices) designed to influence student learning, can provide frequent, ongoing estimates of student ability. Educational technology also has the potential to revolutionize educational practice and research by immersing students in personalized learning opportunities while providing researchers with detailed, real-time accounts of student activity during learning (National Education Technology Plan, NETP, 2010).

Deliberate practice offers a set of established theoretical principles that can be incorporated into educational technology; however, educational technology must be specifically designed to foster deliberate practice. The technology must have an underlying scale that allows learners to be matched to appropriately-challenging activities. As students complete activities, they must receive feedback (e.g., visual, auditory) about performance.

Student performance must also be used to provide real-time updates about the learners' performance to ensure appropriate targeting of subsequent activities. The technology must have a systematic way to monitor the intensity of the activity, to ensure that learners are engaged in the activity and not becoming fatigued. Finally, educational technology designed to provide deliberate practice must be accessible year-round so a learner can engage with the technology on-demand. *Learning Oasis* (Hanlon, Swartz, Stenner, Burdick, & Burdick, 2012) is an educational technology designed to foster deliberate practice.

The purpose of the present study was to examine the conflicting findings in the literature on the relationship between reading practice and reading ability. Based on the expertise literature, the quality of reading practice was examined using the principles of deliberate practice. This research was framed using connectionist models of reading that are consistent with the deliberate practice perspective: the Automatic Information Processing Model (LaBerge & Samuels, 1974), Parallel Distributed Processing Model (Rumelhart, 1977, 1994), and Construction-Integration Model (Kintsch, 1988, 1994, 1998). *Learning Oasis* served as the digital learning environment that immersed participants in deliberate practice in reading. *Learning Oasis* provides self-directed literacy practice that is targeted to the individual level of the reader. Learners receive immediate feedback based on performance and are able to access the learning environment from any computer connected to the Internet. The data collected by *Learning Oasis* also provided critical information about the reading experience (e.g., time spent) that could be used to examine how intensely a participant read. *Learning Oasis* provided the data necessary to estimate the amount of deliberate practice completed by a reader, as well as ongoing estimates of reader ability (Table 2). This information was necessary to answer the primary research question: What is the relationship

between the amount of various aspects of deliberate practice and the development of reading ability? The literature suggests that more deliberate practice should be positively related to growth in reading ability.

Table 2: Operational definition of deliberate practice

Component	<i>Learning Oasis</i> Proxy
Amount of conditional reading volume	Number of words read inside <i>Learning Oasis</i>
Amount of real-time feedback	Number of auto-generated cloze items answered
Amount of intensive practice	Number of intense minutes of reading per day
Distributed practice	Number of days a participant read using <i>Learning Oasis</i> ; Standard deviation of the elapsed days between reading experiences
Self-directed	All reading is self-directed (not used in study)

The hypotheses for the present study were:

H1: The *number of words read from targeted text* will positively relate to the rate of change in reading ability over time

H2: The *number of reading comprehension items answered* will positively relate to the rate of change in reading ability over time

H3: The *number of minutes spent reading intensely per day, measured automatically by the educational technology*, will positively relate to the rate of change in reading ability over time

H4: The *number of days a participant read* will positively relate to the rate of change in reading ability over time

H5: The *standard deviation of elapsed days between reading experiences* will negatively relate to the rate of change in reading ability over time

H6: The rate of change in reading ability over time will be negatively related to the *initial reading ability* of the participant

H7: The rate of change in reading ability over time will be negatively related to the *initial grade of the participant*

H8-a: The *socio-economic status (SES) of a participant* will be positively related to initial reading ability. Participants with lower SES, as measured by receiving free lunch, will have a lower initial intercept than participants with higher SES, as measured by paying for lunch.

H8-b: The *SES of a participant* will be positively related to the rate of change in reading ability over time. As predicted by the reading literature, participants with lower SES will have flatter trajectories than participants with higher SES.

Research suggests that SES can influence how well students read, and how quickly their reading ability changes over time (Bradley, Corwyn, Burchinal, Pipes-McAdoo, & Garcia-Coll, 2001; Noble, Farah, & McCandliss, 2006; Fryer & Levitt, 2004; Rathbun & West, 2004; White, 1982; Whitehurst, 1997). Readers with higher SES often spend more time reading at home, and have access to larger numbers of books and other text resources (Raz & Bryant, 1990; Whitehurst, 1997). As reading researchers have explored the relationship between SES and reading ability, the expertise community has not devoted attention to how SES influences the development of expertise. Proponents of deliberate practice (Ericsson, Krampe, Tesch-Romer, 1993) argue that engaging in deliberate practice puts a novice on the path towards expertise, regardless of personal characteristics. For the present study, the perspectives popular in the reading literature guided the hypotheses. That

is, participants with lower SES will have a lower initial reading ability and will experience less rapid change in ability over time compared to participants with higher SES.

The literature suggests that engaging in deliberate practice develops expertise in a domain or discipline. Unfortunately, the influence of practice on reading ability is not well-supported empirically as past research has suffered from methodological limitations. Educational technology, such as *Learning Oasis*, can provide the detailed data necessary to overcome the limitations of past research in this area. Finding a relationship between deliberate practice and reading ability may be the first step toward reconciling researchers' intuition and the lack of empirical evidence identified by the NRP (2000).

CHAPTER 2

REVIEW OF LITERATURE

Empirical evidence related to the influence of reading practice on reading ability is lacking due to methodological limitations and a lack of understanding of the type of practice that leads to the development of reading ability (Allington, 2009). Deliberate practice provides a set of principles that can be used to describe the quality of the practice. This chapter describes: (1) different models of expertise development and concludes with generalizable aspects of expert performance and a discussion of deliberate practice, (2) models of reading that provide theoretical accounts of how the reading process operates, and (3) how educational technology is transforming education and literacy.

The Development of Expertise

What do Alexander the Great, Ludwig van Beethoven, Clara Barton, Susan B. Anthony, J.R.R. Tolkien, and Michael Jordan all have in common? While each is from a different historical era (i.e., 330 BC to the present), these individuals are considered to be experts in their respective fields (i.e., military conquest, musical composition, the treatment of the sick, advocating for civil rights, storytelling using the written word, basketball). The term *expert* describes a person who has special skill or knowledge of some particular field acquired through prolonged and intense experience. *Expertise* refers to the characteristics, knowledge, and skills that differentiate experts from novices and less experienced people (Ericsson, 1996a, 2006a).

According to Chi (2006), there have been two approaches to the study of expertise: absolute and relative. The absolute approach involves the study of truly exceptional people; these experts can be identified using a variety of techniques: retrospection (e.g., how often a composer's music is played), concurrent measures (e.g., tournament results in chess), or independent indices (e.g., moving a knight across the chess board during a timed experiment). While the absolute approach studies only the expert, the relative approach studies experts in relation to novices (i.e., developing experts). This viewpoint is based on the assumption that novices can become experts, and the goal is to understand how to enable the less skilled (i.e., novices) to become more skilled (i.e., experts). This perspective is particularly compelling in the field of education as teachers strive to transform novices in a field (e.g., reading, writing, mathematics) into experts (Chi, 2006).

Over time, a variety of models that describe the differences between a novice and an expert have been proposed in a multitude of domains. Examples include: creative writing (Kellogg, 2006), medical diagnosis (e.g., Dunphy & Williamson, 2004; Schmidt & Boshuizen, 1993), musical performance (e.g., Clarke, 1988; Shaffer, 1981; Sloboda, 1996), nursing (e.g., Benner, Tanner, & Chesla, 1996a, 1996b, 1996c, 1996d), software design (Sonntag, Niessen, & Volmer, 2006), and radiological diagnosis (e.g., Raufaste, Eyrolle, & Marine, 1998). The next sections provide a description of selected models of expertise development. These models were selected for discussion based on their popularity in the literature and their application to multiple domains. They are presented chronologically according to initial publication date. While prominent in the literature, these models are limited in that their authors argue time-based experience alone is enough to achieve expertise in a domain or discipline. That is, according to these models, expertise develops as a function

of time spent inside a domain, regardless of the nature of the activity. In contrast, in expert performance and deliberate practice, the final perspective described, Ericsson argues that the quality of the practice is critical to the development of expertise. This perspective is discussed in detail as it provides the theoretical foundation for the present study.

The Fitts and Posner Three-Stage Theory. Developed by Paul M. Fitts and Michael I. Posner (1967), the Fitts and Posner three-stage model of skill acquisition is a theory that describes the stages involved in the attainment of skilled performance. According to Fitts and Posner, skilled performance is an organized sequence of activities that are goal-directed and accurate (i.e., activities are performed consistently). A sequence of activities can include both sequences of movements (e.g., hitting a baseball) and sequences of symbolic information (e.g., playing chess). Sensory information and response movements provide ongoing feedback about each step during the performance of a skill. According to the theory, performers do not execute long sequences of actions in predefined, unalterable ways (e.g., like a robot); rather, every act depends upon comparisons between (1) input and feedback and (2) progress and goals (Fitts & Posner, 1967).

Fitts and Posner (1967) described two broad classes of learned skills: perceptual-motor and language. Perceptual-motor skills involve responses to actual objects in the physical world and encompass bodily skills (i.e., locomotion), manipulative skills (i.e., interaction with the environment), and perceptual skills (e.g., estimating velocity and distance, recognizing spatial patterns). Language skills involve the manipulation of signs and symbols (e.g., mathematics, representations of problem spaces, language). The model has been applied to a variety of disciplines: surgical training (e.g., Kopta, 1971; Reznick & MacRae, 2006), sports (e.g., McMorris, 2004), rehabilitation of motor control (e.g.,

Timmermans, Seelen, Willmann, & Kingma, 2009), business decisions (e.g., Salas, Rosen, & Diaz-Granados, 2010), and aviation pilot training (e.g., Vidulich, Wickens, Tsang, & Flach, 2010).

The Fitts and Posner theory of skill acquisition consists of three stages (i.e., cognitive, associative, autonomous) that describe complex skill development. The transition from stage-to-stage is gradual and facilitated through experience (Fitts, 1964). This is a key feature of the model: experience alone moves a novice between stages towards expertise.

Stage 1: Cognitive. The first phase of skill acquisition is marked by learners trying to understand the task and what it requires. While instructors often help learners by pointing out perceptual cues and features of the task, performance at this stage is highly variable with a large frequency of errors. Moreover, learners are unable to diagnose why errors are occurring (Fitts & Posner, 1967). For example, beginning golf students learning how to strike the ball might sometimes lift the ball into the air, and other times tap the ball along the fairway. The novice golfers know that something is not right with their swing or grip, but are unable to determine how to correct the error.

Stage 2: Associative. During the associative phase, many of the fundamental components of a task that were introduced during the cognitive phase have been learned. Errors still occur, but less frequently, and learners concentrate on refining the skill. While learners in the associative stage are unable to precisely diagnose the cause of errors, their understanding of the fundamentals of the skill provides a rough explanation of why mistakes occur. Also, the variability of performances begins to decrease (Fitts & Posner, 1967). For example, golfers in the associative stage might not know precisely why they are consistently

slicing the ball, but they might understand it has to do with their grip or weight transfer during their swing.

Stage 3: Autonomous. After significant amounts of practice and experience, learners transition from the associative stage to the autonomous stage. In this phase, skill use has become almost automatic or habitual; Fitts and Posner (1967) compared activity in the autonomous stage to reflexes. Learners are able to perform the skill without thinking, allowing them to focus on particularly complex components of the activity. Additionally, autonomous learners are able to immediately diagnose the cause of errors and choose the necessary corrective action (Fitts & Posner, 1967). For example, autonomous golfers do not focus on the mechanics of their swing; rather, they might focus on trying to achieve the desired spin on a particular shot.

While the Fitts and Posner three-stage model of skill acquisition has been applied to a variety of domains, it is not applicable to the present study. The model admirably describes the stages of expertise acquisition; however, the mechanism that facilitates the shift from stage-to-stage is ill-defined. Experience is proposed as the catalyst that moves a novice to more advanced stages of expertise, but the nature of the experience itself is not described.

The Dreyfus Model of Skill Acquisition. The Dreyfus Model of Skill Acquisition was developed by Hubert L. Dreyfus and Stuart E. Dreyfus (1980) with funding provided by the United States Air Force. While the brothers came from different professional backgrounds, philosophy and operations research respectively, the Air Force asked them to investigate skill acquisition as it related to pilot training (Dreyfus, 1997). The Dreyfus model of skill acquisition has been applied to a variety of disciplines (e.g., decision-making, system management, teaching second language learners, medical diagnosis). Most notably, Benner's

(1984) application of the model of skill acquisition to the field of nursing to establish clinical competence has been highly influential (Wandel, 2003). More recently, Hunt (2008) applied the Dreyfus model to the domain of software development.

In researching how a novice becomes an expert, Dreyfus and Dreyfus focused on unstructured problem domains. That is, a problem area that consisted of large numbers of facts and features that can relate in a variety of ways: social interactions, medical diagnosis, nursing, flying an airplane, chess, and learning a second language. A high level of skill in an unstructured problem domain requires vast amounts of experience with authentic situations, leading to domain-specific expertise. For example, a business-person might exhibit expertise at marketing, while being competent at financial planning, and merely a novice in negotiating a merger (Dreyfus & Dreyfus, 1986).

By placing concrete experience as a prominent factor in their model, Dreyfus and Dreyfus offer an antithetical perspective to Plato and over two centuries of philosophical thought on the nature of expertise. The traditional view (e.g., Plato, Piaget) is that a beginner starts with specific cases and advancing proficiency results in the abstraction and internalization of more complex rules. In contrast, Dreyfus and Dreyfus have argued that skill acquisition moves from abstract rules to particular cases (Dreyfus & Dreyfus, 2005).

Dreyfus and Dreyfus (1986) made a distinction between crude skills and subtle skills. Crude skills (e.g., walking, driving) are generally easy to do as there are large margins for error, time to make corrections, and the results are not final. In contrast, when using subtle skills (e.g., skills used during chess, music, professional sports, or medicine), a miniscule difference in performance can have a significant influence on the result. Additionally, subtle skills often require quick action and experts are not given the opportunity to correct mistakes.

Dreyfus and Dreyfus (2005) argued that crude skills can be performed expertly while thinking about something else, while subtle skills require absolute concentration.

Dreyfus and Dreyfus were able to distinguish between five stages of qualitatively different perceptions (i.e., novice, advanced beginner, competency, proficiency, expertise). Perceptual shifts include changes in both perception of task and perception of method of decision making (Dreyfus & Dreyfus, 1986). Novices advance towards expertise through authentic, domain-specific experience.

Stage 1: Novice. During the first stage of novel skill acquisition, novices learn to recognize various objective facts and features related to a skill and begin to acquire rules for determining action (Dreyfus & Dreyfus, 1986). The features (i.e., elements of the situation), often explicitly provided by an instructor, can be recognized by a novice without specific experience in the domain (i.e., context-free) (Dreyfus & Dreyfus, 1980). Examples of context-free features and rules include: a student driver learning what speed (i.e., a context-free feature indicated by the speedometer) to shift gears (i.e., a rule) when driving a manual transmission car, a novice nurse being taught to read blood pressure monitors, and a beginning chess player being taught how to assign point-values to individual pieces and the strategic rule of trying to exchange your piece for an opponent's if the point value of the piece you capture exceeds the value of the piece you sacrifice (Dreyfus & Dreyfus, 1980, 1986, 2005).

The rules learned by a novice do not produce exemplary performance (Dreyfus & Dreyfus, 2005). For example, a chess player who blindly plays using the “point exchange rule” will not be able to defeat opponents who sacrifice higher-valued pieces to gain a

tactical advantage. These initial rules allow for the accumulation of experience, and lead a novice into the next stage.

Stage 2: Advanced Beginner. As a novice acquires more experience in authentic (i.e., context-specific) situations, the novice begins to distinguish meaningful aspects of the domain (Dreyfus & Dreyfus, 1986). Maxims are context-specific aspects that are recognized based on perceived similarity with prior experience. Unlike rules, these situational maxims require an understanding of the domain (Dreyfus & Dreyfus, 2005). An advanced beginner uses both situational maxims and non-situational (i.e., context-free) features to approach problems. For example, when determining when to shift gears when driving, an advanced beginning driver might use engine sounds (i.e., situational) as well as driving speed (i.e., non-situational) to determine when to shift to a higher gear. A chess player in the advanced beginner stage might be able to recognize situational aspects of a game (e.g., weakened king's side, strong pawn structure) and take this into account when deciding moves (Dreyfus & Dreyfus, 1986). Advanced beginners rely on situation-specific experience to exhibit improved performance when compared to novices. These beginners rely on experience to recognize situations and apply learned rules to advance in a skill; however, eventually the amount of experience leads advanced beginners into the next stage of skill acquisition.

Stage 3: Competence. As an advanced beginner acquires more experience, the number of relevant elements that are recognizable becomes overwhelming (Dreyfus & Dreyfus, 2005). To cope with this overload of information, competent people learn, or are taught, a hierarchical procedure for decision-making. By first selecting a method of organizing the situation (i.e., a perspective), and then examining the most relevant elements based on the chosen perspective, a competent performer is able to simplify and improve

performance (Dreyfus & Dreyfus, 1986). Unlike the procedural rules and maxims used by novices and advanced beginners, competent performers have access to a vast number of situations, often with subtle variations. Therefore, each competent learner must decide what perspective to adopt in each different situation, often while being unsure of the outcome (Dreyfus & Dreyfus, 2005).

During the novice and advanced beginner stages, if the rules or maxims applied to a task result in failure, performers often rationalize the failure as a limitation of the understood rules. In contrast, in the competent stage, the result depends on the learner's choice of perspective. This leads a competent performer to feel responsible for success or failure (Dreyfus & Dreyfus, 1986).

Stage 4: Proficiency. As competent performers experience success and failure based on their chosen perspectives, the performer's representation of the skill (i.e., rules, maxims) is gradually replaced by holistic representations that encompass sets of component features (Dreyfus & Dreyfus, 2005). Put differently, proficient learners are able to intuitively use patterns without breaking a situation into its smaller component rules and maxims (Dreyfus & Dreyfus, 1986). While proficient people are able to spontaneously see the important aspects of a situation, they still must make a conscious decision about how to act; this decision is still based on the rules and maxims employed during the competence stage. A proficient chess player is able to immediately recognize large sets of piece formations, but must deliberate on how best to maneuver (Dreyfus & Dreyfus, 2005).

Stage 5: Expertise. While proficient performers are able to intuitively determine what needs to be accomplished, they still must decide how to proceed. Experts, on the other hand, not only intuitively recognize what must be done, but also immediately know how to achieve

the goal. With enough experience in multiple situations with varying subtleties, experts are able to break situations into subclasses that necessitate unique responses (Dreyfus & Dreyfus, 2005).

Expertise is characterized by immediate, intuitive situational responses. For example, expert chess-players are able to play successfully at a rate of 5 to 10 seconds per move, depending entirely on intuition to recognize the board structure and select moves. Expert drivers are able to perform actions without calculating alternatives (Dreyfus & Dreyfus, 2005).

The Dreyfus Model of Skill Acquisition has had a lasting influence on the expertise literature, particularly in the field of nursing. In arguing that would-be experts require domain-specific experiences to achieve expertise, Dreyfus and Dreyfus stress the central nature of experience. Unfortunately this model is not sufficient for this study as the type of experience necessary to achieve expertise is not defined by the model. As Allington (2009) argued, the quality of the reading practice must be explored if the influence of reading experience on growth in reading ability is to be understood.

The Model of Domain Learning. Developed by Patricia A. Alexander (1997, 2003a), the Model of Domain Learning (MDL) describes the journey of a novice towards competency (i.e., expertise) in an academic domain. An academic domain is defined as the formalized body of conceptual knowledge organized around concepts and principles (Matthews, 1994; West & Pines, 1985). Domains are often socially constructed (Phillips, 1995) and in a state of perpetual formation (Alexander, 1997). The MDL or its components have been successfully applied to a variety of domains: reading (Alexander, 2003b; Fox,

2009), special education (Alexander, Sperl, Buehl, Fives, & Chiu, 2004), music therapy (Langan & Athanasou, 2005), and Internet navigation (Schrader, Lawless, & Hayley, 2008).

In the MDL, Alexander posits that achieving expertise requires subject-matter knowledge, motivation, and strategic processing. These components are theorized to influence one another in different ways within each stage (Alexander, 2003a). Subject-matter knowledge is divided into two opposing types of knowledge: domain knowledge and topic knowledge. Domain knowledge represents the breadth of declarative, procedural, and conditional knowledge a learner possesses (Alexander, Schallert, & Hare, 1991). In contrast, topic knowledge describes the depth of knowledge in a given domain (Alexander, Kulikowich, & Schulze, 1994). For example, learners knowledge of computer science can be described by how much they know about the field in general (i.e., domain knowledge) as well as how much is known about specific topics (e.g., polymorphism, recursive algorithms, back-face culling) in the domain. As learners progress from acclimation towards expertise, these forms of knowledge grow in tandem (Alexander, 1997).

Like subject-matter knowledge, interest plays an important role in the MDL. Both individual interest and situational interest (Hidi, 1990) influence learner progress towards expertise. Individual interest, a learners' association or deep-seated investment in a domain, is enduring (Alexander, 1997; VanSledright & Alexander, 2002). In contrast, situational interest (i.e., temporary arousal triggered by the immediate context) is fleeting (Hidi, 1990). While situational interest remains a force at all stages of the MDL, individual interest often provides the motivating power behind growing beyond acclimation. Learners' individual interests can energize and motivate their thoughts and actions in goal-directed ways. When learners enter a novel domain, their situational interest will be high, helping them to persist

as they grow towards competence and begin to develop individual interest in the domain (Alexander, 1997).

The third component of the MDL is strategic knowledge. Strategic knowledge, a special form of procedural knowledge, is intentionally and purposefully invoked by a learner whose goal is to maximize performance and avoid obstacles to understanding (Alexander et al., 1991; Garner & Alexander, 1989; Paris & Winograd, 1990). Described differently, these strategies are the tools applied by a learner when acquiring, transforming, and transferring information (Pintrich, Marx, & Boyle, 1993). Alexander (1997) included general cognitive (e.g., note-taking), metacognitive (e.g., self-testing), and self-regulatory processes inside strategic knowledge. According to the MDL, the frequency and complexity of strategy use varies by stage (i.e., acclimated learners often use simpler strategies than experts). Also, as these strategies must be purposefully applied, strategy use is directly tied to learner goals and can be reflective of student motivation (Alexander, 1997).

The MDL is a theory whereby learners progress through three stages: acclimation, competence, and proficiency/expertise. These stages are non-recursive and non-regressive (i.e., once a given stage is achieved, the achievement is sustained absent any dramatic changes in the cognitive state of the learner). The movement from early to later stages is characterized by recognizable transformations in perceptions and understanding, as well as the way learners see themselves within a domain. Additionally, the transition between stages occurs gradually (i.e., there is no singular moment of change). The MDL is not tied to chronological age; rather, it is related to experiences and schooling within a domain (Alexander, 1997, 2003a).

Stage 1: Acclimation. A learner in the acclimation stage has minimal knowledge of a specific domain. Knowledge is fragmented and unprincipled in organization and is neither cohesive nor well-integrated. Acclimated learners are dependent on guidance and scaffolding to acquire domain knowledge because (1) they struggle to distinguish between domain and non-domain as well as important and unimportant information, and (2) their limited subject-matter knowledge can complicate strategy usage. Given the limited knowledge-base of an acclimated learner, misconceptions (i.e., serious misunderstandings in concepts) are limited in quantity, but their influence can be severe if not corrected (Alexander, 1997, 2003a). Acclimated learners often rely on situational interest to persist on a task (Mitchell, 1993).

Stage 2: Competence. The competence stage is characterized by a transformation in the knowledge base, level of interest, and strategic abilities. Competent learners' domain-knowledge is better organized and shows a breadth and depth beyond that possessed by learners in the acclimation stage. This improved organization allows a competent learner to acquire and integrate new information more rapidly than an acclimated learner (i.e., one who acquires new knowledge piecemeal) (Alexander, 1997). Enhanced subject-matter knowledge enables competent learners to more readily engage in domain-specific strategies that eventually are used with greater automaticity (Anderson, 1982, 1987; Chi, 1985; Glaser, 1984; Harris & Graham, 1996). Also, increased content-area knowledge often leads to an increase in an individual's interest for a domain (Alexander, 1997; Alexander et al., 1994; Renninger, 1992). According to Alexander (1997), a learner can transition from acclimation to competence if at least one of the three forces (i.e., subject-matter knowledge, interest, and strategic processing) is sufficiently developed. For example, a student could become

competent based on well-organized content-knowledge, immense interest in a domain, or superior strategy usage.

Stage 3: Proficiency/Expertise. While competence could be achieved through the development of content-knowledge, interest, *or* strategy-use, proficiency/expertise can only be achieved when all three forces are well-developed. Experts have a broad, rich, and highly-integrated knowledge base, a long-term commitment to the domain (i.e., interest), and are highly strategic when faced with complex and novel problems (Alexander, 1997, 2003a). Given the extraordinary levels of knowledge, interest, and strategic processing necessary to achieve expertise, Alexander (1997) posited that it is unreasonable to assume that everyone is capable of reaching or is willing to pursue proficiency. Learners who achieve proficiency become less concerned with solving routine problems or accumulating knowledge, and focus on problem formulation and knowledge generation (Chi, 1985; Garner, 1987).

The MDL is unique among models of expertise development as it explicitly includes a motivational component. Moreover, the expectation that not every student achieves expertise in every domain is intuitively accurate as learners specialize later in their academic careers. Given the focus of the present study, to explore the type of reading practice necessary to increase reading ability, this model was not applicable as it does not describe the type of practice necessary to achieve expertise. However, future research could incorporate the MDL as the motivational aspect could certainly be a contributing factor to students' reading development.

Limitations of solely experience-driven models of expertise development. The stage theories of Fitts and Posner (1967), Dreyfus and Dreyfus (1986), and Alexander (1997) offered detailed descriptions of the phases a learner experiences while advancing from novice

to expert in a domain or skill. However, the mechanisms of change (i.e., how a learner actually advances from one stage to another) are less clearly defined beyond the importance of experience. Fitts and Posner argued that extended practice and experience are necessary to achieve automated skill performance. Dreyfus and Dreyfus posited that to achieve competence (i.e., immediate, intuitive situational responses) learners must experience variations of a situation multiple times. Lastly, Alexander listed experience and schooling as the contributing factors that enable learners to develop expert levels of content-knowledge, interest, and strategy-use.

The notion that expertise is the result of experience, and that it takes at least ten years of learning in a domain to become an expert led researchers (e.g., Lawrence, 1988; Posner, 1988) to define expertise as a predictable progression from novice to expert with experience and time driving advancement. Unfortunately, this method of classifying experts (i.e., based solely on amount of experience and time) failed in a variety of domains. Doane, Pellegrino, and Klatzky (1990) showed that the performance of highly-experienced computer programmers was not consistently superior to the programming efforts of students. In the domain of psychology, Dawes (1994) showed that the success of clinical psychologists in treating patients was not linked to their length of training. Studies in the areas of auditor evaluations (Bedard & Chi, 1993) and x-ray diagnosis (Ericsson, 2004) showed that level of performance decreased as a function of the length of professional experience after formal training.

Given the limitations of purely experience-based definitions of expertise, Ericsson and Smith (1991a) proposed that experts are individuals that consistently produce superior performance on tasks in their domain that are representative and authentic. Rather than define

experts as individuals with a predetermined amount of experience in a domain, Ericsson and Smith defined expertise based on performance. For example, in the domains of sports and music, experts are those performers who produce superior performance regardless of the context (e.g., rehearsal, practice, concerts, competition). The scientific study of expertise requires the design, capture, and analysis of superior performances on representative tasks (Ericsson, 2006b). Ericsson and Smith also argued that identifying the theoretical mechanisms of expertise acquisition, and how they can be acquired and enhanced are important components to the study of expertise. Ericsson and Smith believed that experience alone was not enough to develop expertise; rather, particular types of experience could foster expert performance. The next sections discuss these topics.

Expert performance. Expert performance, as defined by K. Anders Ericsson (1996a), is consistent, measurable, and reproducible superior performance on representative domain-specific tasks. These representative tasks are specifically designed to capture the essence of a domain (Ericsson, 1996a; Ericsson & Smith, 1991a). Examples of representative tasks include: asking chess players to generate the best move for specific chess board configurations, presenting typists with preselected text and asking them to reproduce as much as possible in a given time, and asking musicians to play familiar and unfamiliar compositions and then asking them to repeat the performance. Experts are often asked to perform in response to external demands (e.g., during a recital or sporting event), and are able to reproduce expert performance under research conditions (Ericsson & Lehmann, 1996).

A description of expert performance requires three separate conceptual discussions. First, an overview of the generalizable characteristics of expertise provides a foundation

upon which Ericsson's theories are built. Next, the factors that can influence the development of expertise are described.

Generalizable characteristics of experts. The literature on expertise is vast and covers several decades (e.g., Anderson, 1981; Bloom, 1985; Chase, 1973; Chi, Glaser, & Farr, 1988; Clancey & Shortliffe, 1984; Ericsson, 1996b; Ericsson, Charness, Feltovich, & Hoffman, 2006; Ericsson & Smith, 1991b; Feltovich, Ford, & Hoffman, 1997; Hoffman, 1992; Starkes & Allard, 1993; Starkes & Ericsson, 2003). Building on previous work (e.g., Glaser & Chi, 1988; Chi, 2006), Feltovich, Prietula, and Ericsson (2006) presented eight general characteristics of expertise based on the body of literature: (1) expertise is domain-specific, (2) knowledge and content matter are important to expertise, (3) experts possess larger and more integrated cognitive units than novices, (4) expertise involves functional, abstract representations of information, (5) automaticity is important for expertise, (6) expertise involves selectivity, (7) expertise requires metacognition, and (8) expertise is more than basic skill acquisition. Each of these characteristics will be described along with findings based on current theoretical perspectives.

Expertise is domain-specific. Expertise is domain-specific, and people seldom become experts in more than one discipline (Glaser & Chi, 1988). Moreover, there is little transfer of expert performance in one domain to performance in another, regardless of any apparent similarities between the domains (Feltovich et al., 2006). For example, Eisenstadt and Kareev (1979) investigated the memory of board configurations by GO and Gomoku (i.e., board games popular in China) players. While both of these games use the exact same board and employ identical pieces, GO experts performed poorly when asked to recall Gomoku configurations and vice versa. This form of memory experiment is similar to the

methods applied by Chase and Simon (1973) while investigating expertise in chess. The assertion that expertise is domain specific has been confirmed in a variety of domains: medical diagnosis (Johnson et al., 1981), music (Lehmann & Grubber, 2006), problem solving in political science (Voss, Greene, Post, & Penner, 1983; Voss, Tyler, & Yengo, 1983), sports (Hodges, Starkes, & MacMahon, 2006), and taxi-driver route selection (Chase, 1983).

Knowledge and content matter are important to expertise. Domain-specific knowledge has a significant influence on the development of expertise. Newell and Simon (1972) showed that problem solving and skilled performance in a domain were predominantly influenced by domain-specific patterns and actions. Content-area knowledge has also been found to effect cognitive processing. Chi (1978) compared the performance of experienced chess-playing children with the performance of non-chess-playing children on memory and learning tasks involving chess. The students with experience playing chess were better able to apply memory strategies (e.g., grouping, rehearsing) and store more information in short-term memory.

The importance of domain knowledge has also been identified by expert performance researchers. Ackerman and Beier (2006) and Horn and Masunaga (2006) provided evidence that experts' superior performance on representative tasks can be described by mental factors (e.g., expert reasoning and expert working memory) that are different from the factors that characterize novice performance. To achieve mastery in a domain, a learner must have the necessary domain knowledge (Steier & Mitchell, 1996).

Experts possess larger and more integrated cognitive units than novices. Increased experience in a domain leads to improved cognitive organization of task-specific information

(Feltovich et al., 2006). This has been one of the most enduring findings in expertise research (Glaser & Chi, 1988); however different theoretical representations of this phenomenon have emerged. In their pioneering work, Chase and Simon (1973) found that expert chess players were able to recall the positions of significantly more (i.e., four to five times as many) pieces than novice players. This phenomenon was attributed to experts' improved chunking of perception and memory. These chunks (i.e., structures that combine smaller units into larger organizations) develop over-time with experience and allow experts to functionally expand working memory (Baddeley & Hitch, 1974; Miller, Galanter, & Pribram, 1960). For example, an expert chess player might recognize a particular configuration of pawns as a Queen's Gambit (i.e., a single chunk); in contrast, a novice might require a separate chunk for each pawn. The importance of chunking to expert performance has been confirmed by multiple studies (e.g., Egan & Schwartz, 1979; Engle & Bukstel, 1978; Reitman, 1976). According to this perspective, novices and experts are both constrained by the size of working-memory (Cowan, Chen, & Rouder, 2004; Miller, 1956); yet, experts achieve their performance because the size of their chunks is larger than those of novices (Feltovich et al., 2006).

While some researchers believe that experts possess larger chunks, permitting a larger functional working memory, an alternative perspective has emerged. By clearing working memory with an intervening activity, Charness (1976) showed that expert chess players do not rely on working memory for the storage of chess positions. To account for this phenomenon, Ericsson and Kintsch (1995) argued that experts employ long-term working memory (LTWM) to recall chess positions. LTWM is not space-limited like normal working memory (Ericsson & Kintsch, 1995; Miller, 1956). By storing domain-specific

representations in LTWM, experts are able to efficiently retrieve the necessary information (Gobet & Charness, 2006). While each perspective offers a different viewpoint (i.e., where experts store domain-specific information during a task), both approaches agree that experts possess more complex cognitive units than novices.

Expertise involves functional, abstract representations of information. As novices become experts, they acquire complex representations of information that provide immediate and integrated access to the domain-specific knowledge necessary to exhibit expert performance (Feltovich et al., 2006). This finding has been researched in a variety of domains: baseball (Spilich, Vesonder, Chiesi, & Voss, 1979), bridge (Charness, 1979; Engle & Bukstel, 1978), computer programming (Adelson, 1981; McKeithen, Reitman, Reuter, & Hirtle, 1981), physics (Chi, Feltovich, & Glaser, 1981), and medicine (Feltovich, Johnson, Moller, & Swanson, 1984; Johnson et. al, 1981). Based on this research, Zeitz (1997) argued that experts' domain-knowledge representation could be described using complex cognitive schemas known as Moderately Abstracted Conceptual Representations (MACRs). Experts' MACRs enable expert performance by: (1) retrieving appropriate material from memory, (2) integrating information and revealing important information, (3) providing guidance and justification for action, (4) enabling analogical reasoning, and (5) providing abstractions that support reasoning and evaluation of alternatives.

Experts' representations also extend to entire activities (Feltovich et al., 2006). Ericsson and Kintsch's (1995) model of LTWM supports the continual updating of the mental representation based on an experts' monitoring of a task (i.e., situational awareness). The notion that improved representations of knowledge enable event-level planning, reasoning, monitoring, and evaluation (Ericsson, Patel, & Kintsch, 2000) has been supported

in several domains. For example, Klein (1998) showed that expert fire fighters interpret a scenario dynamically, anticipating how a raging fire will evolve. Similarly, Koschmann, LeBaron, Goodwin, and Feltovich (2001) showed that expert surgeons will take actions that have no immediate impact, but facilitate later action during a medical procedure.

Automaticity is important for expertise. Research results suggest that the nature of cognitive operations change with expertise acquisition. Operations that originally are slow, serial, and require conscious attention transform into operations that are fast, less deliberate, and can be processed nearly in parallel with other operations (Schneider & Shiffrin, 1977). The influence of automaticity on expertise has been researched in many domains. For example, Shaffer (1975) showed that expert typists were able to type and recite nursery rhymes in parallel. Similarly, Hatano, Miyake, and Binks (1977) showed that expert abacus operators were able to answer routine questions (e.g., “What is your favorite color?”) without degradation of speed or accuracy in abacus-based operations.

Automaticity is important to expertise as it enables experts to use higher-level skills (e.g., reasoning, comprehension, monitoring, integration) (Feltovich et al., 2006). In complex skills with a variety of cognitive components, the most basic operations (e.g., decoding, encoding) must be automated if higher-level skills are to develop (Endsley, 2006; Logan, 1985). For example, Lesgold and Resnick (1982) showed that when students learn to read, if the basic decoding and encoding of letters and words is not automated, comprehension skills do not substantially develop.

Expertise involves selectivity. Newell (1973) argued that expertise involves the recognition of familiar features of a problem and the application of past-experience based on those features (Ross, Shafer, & Klein, 2006). Given that two situations are often not

identical, experts rely on abstracted features to classify problems (Chi et al., 1981).

Selectivity, the process of separating important aspects of a problem from superfluous information is critical to expertise (Chi et al., 1981; Hinsley, Hayes, & Simon, 1978; Patel & Groen, 1991). In domains with predominantly static situational tasks (e.g., typing), a positive relationship between selectivity and performance has been discovered. Said differently, when performers are able to recognize familiar features of a task, they tend to perform at a higher level (e.g., Rieger, 2004). In contrast, tasks that are often dynamic and vary from experience to experience are harder to master as each encounter offers unique or hard to distinguish features (Chi et al., 1981).

Expertise requires metacognition. Metacognition, knowledge about one's own knowledge and performance (Flavell, 1979), is vital to expert performance. The elegance associated with expert thinking (Bartlett, 1958) derives, in part, from metacognitive processes that allow experts to test their understanding and partial solutions to problems. This ongoing monitoring prevents experts from suffering novice-level mistakes (e.g., blind alleys, errors, extensive retracing of steps, starting anew) during problem solving (Feltovich et al., 2006).

Metacognition allows experts to recognize when a particular situation is outside typical domain-specific experience and requires a shift in approach (Feltovich, Spiro, & Coulson, 1997). Research suggests that metacognition can become automated (Reder & Shunn, 1996) and that metacognitive strategies can be explicitly learned in both general (Kruger & Dunning, 1999) and domain-specific contexts (e.g., Kuiper & Pesut, 2004; Paris & Winograd, 1990).

Expertise is more than basic skill acquisition. Experts have more knowledge that is better organized than novices. Experts continually reorganize and refine their representations of domain-knowledge and procedures to enable efficient problem-solving (Ericsson & Lehmann, 1996). Expertise is viewed not as simple skill or fact acquisition, but as a complicated construct of adaptations (i.e., mind and body) to specific domains. These adaptations, which include self-monitoring and control mechanisms, set experts apart from novices (Feltovich et al., 2006).

Based on the traditional view of expertise proposed by Simon and Chase (1973), researchers have argued that both novices and experts are limited by a fixed-capacity for processing. Experts distinguish themselves from novices by adapting to overcome the limits of attention (Shipp, 2004), working memory (Baddeley, 2000, 2002), and long-term memory access (Brown, 1991; Brown & MacNeill, 1966). Expertise involves the development of integrated representations of knowledge that minimize the impact of information-processing limits (Feltovich et al., 2006). For example, the concept of chunking was proposed as an adaptation that enables experts to overcome the limitation of working memory. Studies (e.g., Meyer & Kieras, 1997; Schumacher et al., 2001) have shown that experts are able to perform simultaneous tasks, overcoming the bottlenecks to processing that hamper novices.

Experts possess characteristics that distinguish them from novices in a domain. Improved cognitive organization, broader content-area knowledge, and developed psychological processes all allow an individual to perform at an expert level. According to Ericsson (1996a, 2002), the development of these expert characteristics is facilitated by engaging in practice that is deliberate.

Factors that can influence the development of expertise. According to Ericsson (2006b), superior performance can be influenced by both teachers and experience. These factors provide novices with the support and experience necessary to progress towards expertise. The remainder of this section describes the importance of these influences.

The role of the teacher. Over time, humans have developed methods for accumulating and preserving knowledge and designed tools to transfer that knowledge to future generations. Thanks to improved methods of training and instruction, individuals no longer must rediscover existing knowledge; aspiring experts can quickly acquire the knowledge and techniques employed by the pioneers of their domain (Ericsson, 2006b). For example, the 13th century philosopher Roger Bacon argued that using self-study (i.e., method of learning contemporary to Bacon), it would require 30 to 40 years to master mathematics (Singer, 1958). Today, the equivalent material (i.e., calculus) is taught in high schools across the country. The development of expertise requires instruction by teachers that help aspiring experts gain access to the body of domain-specific knowledge (i.e., predefined concepts, notation systems, equipment, and measurement devices) (Ericsson, 2006b).

The role of experience. To perform at an expert level requires extended immersion in domain-related activities (Ericsson, 1996b, 2004, 2006b; Ericsson & Lehmann, 1996). Standardized tasks and assessments have allowed researchers to measure performance during development and monitor changes over time (Ericsson, 2006b; Ericsson & Lehmann, 1996). This longitudinal data led to the identification of three trends that link experience in a domain to expertise (Ericsson, 2006b). First, all people improve gradually; there is no evidence of abrupt advances in development. Even child prodigies progress steadily over-time, albeit at a faster speed.

The second trend is that superior performance is not limited by the functional capacity of the body or mind since performance continues to improve beyond the age of physical maturity (Ulijaszek, Johnston, & Preece, 1998). Also, the typical chronological age of reaching expertise varies by domain. For example, athletes reach expertise in the mid-20s whereas scientists and artists typically do not achieve expertise until their 30s and 40s (Schulz & Curnow, 1998; Simonton, 1997).

The final trend supports the findings of Bryan and Harter (1899) and Simon and Chase (1973) that achieving expertise requires at least ten years and 10,000 hours of domain-specific experience. This finding has been confirmed in the domains of music composition (Hayes, 1981), sports, science, and the arts (Ericsson et al., 1993). The ten year requirement serves only as a minimum as some domains require more than ten years of experience; outstanding scientists and authors typically publish their first work around age 25 and their best work a decade later (Raskin, 1936).

These trends are consistent with the models of expertise development proposed by Fitts and Posner (1967), Dreyfus and Dreyfus (1980), and Alexander (1997). Expertise is not a binary state whereby somebody suddenly jumps from novice to expert in a single moment. To achieve expertise in a domain requires the aspiring expert to gradually acquire domain-specific experience over time.

While Ericsson acknowledges the importance of experience, he argues that the quality of the experience itself is important to achieve expertise as only certain types of domain-specific activity lead to improvements in performance. Deliberate practice, the topic of the next section, involves domain-specific activities that have been specifically designed to enhance performance (Ericsson, 2006b). Unlike the models of Fitts and Posner (1967),

Dreyfus and Dreyfus (1980), and Alexander (1997), Ericsson specifically describes the qualities of the practice that are necessary to develop expertise.

Deliberate practice. Deliberate practice (Ericsson, 1996a, 2002, 2004, 2006b; Ericsson et al., 1993), a core tenant of Ericsson's theory of expertise development, assumes that expertise is acquired gradually over-time and that a specific type of experience is necessary to achieve and maintain expertise. Deliberate practice involves the careful design of training tasks, often by a teacher or coach, which a learner can master sequentially. Learners are presented with tasks that are outside their current performance-level but can be mastered within hours of practice. This practice is focused on the important aspects of the task and leads to gradual improvement through repetition with feedback. Learner concentration during practice is critical to successful deliberate practice as mastering a task requires more than mindless repetition or playful practice.

From an expert performance perspective, the largest threat to expertise development is a performance asymptote whereby an aspiring expert has reached a stagnant state and is no longer improving. This phenomenon, predicted by theories of skill acquisition (e.g., Anderson, 1982; Fitts & Posner, 1967), occurs when a learner is not appropriately challenged and automaticity makes additional experience irrelevant. To avoid this stasis, would-be experts must continue to acquire experience that is coupled with deliberate practice. By engaging in demanding tasks that require problem solving and stretch performance, experts are able to acquire and refine cognitive mechanisms that support continued learning and avoid arrested development (Ericsson, 2006b).

Research suggests that there are limits to the daily duration of deliberate practice. Experts from a variety of domains engage in practice for no more than one hour without rest

(Ericsson et al., 1993). Even with periodic rest or recuperative naps, experts in a variety of domains reported that the amount of practice never exceeded five hours per days (Cowley, 1959; Ericsson, 1996a; Ericsson et al., 1993; Plimpton, 1977). If aspiring experts are not able to restore their equilibrium through rest and nightly sleep, they risk suffering from an incapacitating burnout (e.g., grievous injury) (Ericsson, 2006b).

Deliberate practice is not limited to expert-directed practice. Learners often engage in deliberate practice when a coach or teacher is unavailable to guide the practice session. For example, expert violinists report rehearsing specific musical skills during private practice time, independent of their music instructor (Ericsson et al., 1993).

The principles of deliberate practice have been investigated in a variety of fields. Examples include: chess (e.g., Charness, Krampe, & Mayr, 1996; Charness, Tuffiash, Krampe, Reingold, & Vasyukova, 2005; Gobet & Charness, 2006), decision-making in a crisis (e.g., McKinney & Davis, 2004), impact of studying on student achievement (e.g., Plant, Ericsson, Hill, & Asberg, 2005), mathematical computation (e.g., Butterworth, 2006), and sports (e.g., Helsen, Starkes, & Hodges, 1998; Hodges & Starkes, 1996; Starkes, Deakins, Allard, Hodges, & Hayes, 1996). In the domain of musical performance, the tenants of deliberate practice have been discovered in the training regimen of expert musicians (Ericsson et al., 1993; Lehmann & Gruber, 2006; Sloboda, Davidson, Howe, & Moore, 1996). Expert musicians often engage in solitary rehearsal (Green, 2002) that is intensely focused on particular components of a performance (Winner, 1996). Feedback is received via sound in the form of misplayed notes or incorrect pacing, and is often heard and diagnosed iteratively over the course of a rehearsal by expert musicians (Lehmann, 2002). Ericsson et al. (1993) showed that expert musicians engaged in over 10,000 hours of deliberate practice.

In the field of professional writing, deliberate practice has been observed in the habits and techniques of expert writers (Henry, 2000; Kellogg, 2006; MacKinnon, 1993; Paradis, Dobrin, & Miller, 1985). Prolific writers often carefully organize their daily work, allotting limited time to write in order to avoid exhaustion (Boice, 1994, 1997; Cowley, 1958; Kellogg, 1986; Plimpton, 1963). Writing workshops (e.g., The Iowa Writer’s Workshop) and graduate programs often provide appropriately challenging writing tasks and immerse aspiring authors in feedback from instructors (Adams, 1993).

Ericsson’s notion of deliberate practice provides a powerful mechanism that can be used to guide a learner from novice to expertise in a domain. According to the principles of deliberate practice, a task designed to develop expertise will provide: (1) targeted practice designed to “stretch” the learner, (2) real-time corrective feedback, (3) distributed practice over a long period of time, (4) intensive practice that does not require the learner to concentrate beyond their limits (i.e., to avoid burnout), and (5) self-directed practice when a teacher or coach is unavailable. Through practice that is designed according to these principles, would-be experts acquire domain-specific knowledge and experience that facilitates expertise development. Activity is tailored to an appropriately challenging level with ongoing indicators of performance, helping novices expand into more complex aspects of a domain or discipline. This expansion facilitates development of the knowledge representations and mental skills necessary to achieve expertise (Feltovich et al., 2006).

Deliberate practice and experience-based perspectives on expertise. Models of expertise development provide theoretical accounts of how novices become experts in a specific discipline or domain. The solely experience-based models of expertise development (e.g., The Fitts and Posner Three-Stage Theory, The Dreyfus Model of Skill Acquisition, The

Model of Domain Learning) argue that experience alone in a domain is sufficient to achieve expertise. In contrast, Ericsson's theories of expert performance and deliberate practice frame the journey from novice to expertise in terms of the quality of the practice completed.

According to Ericsson, expertise can be fostered by providing learners with feedback-rich practice that is targeted to their individual abilities. If aspiring experts concentrate and self-direct practice activity over an extended period of time, they are on the road towards achieving expertise. Educational technology offers a platform that can provide learners with opportunities to engage in deliberate practice when learners do not have access to feedback from a human teacher or coach. The core concepts of deliberate practice are consistent with theoretical accounts of the reading process. These models are discussed in the next section.

Theories of Reading

As one of the foundational "Rs" of education, the reading construct has attracted research interest from a diverse set of academic disciplines: general education, linguistics, neuroscience, psychology, and sociology (Thomas, 1996; Tracey & Morrow, 2006). The explanatory and predictive properties of the models proposed by researchers in these varying fields have informed research and impacted teaching and learning (Bernard, 2000; Best & Kahn, 1998; Stanovich, 1992; Tierney, 1994). A single theoretical model that explains all phenomena related to reading (e.g., reading process, reading development, reading disability, reading instruction) would be useful in providing a comprehensive, widely-accepted perspective of reading. However, a single unified theory of reading has remained elusive, and literacy researchers have adopted multiple models to explain the complexities of reading (Pressley & McCormick, 1995; Tierney, 1994; Woolfolk, 1998).

One group of reading models has remained at the forefront of reading research for over four decades. Proponents of these connectionist models of reading posit that the reading process relies on cognitive units of knowledge that become interconnected and related through experience (Adams, 1994; Rumelhart & McClelland, 1986a; Seidenberg & McClelland, 1989). Three such connectionist models that frame the present study include: the Automatic Information Processing Model (LaBerge & Samuels, 1974), Parallel Distributed Processing Model (Rumelhart, 1977, 1994), Construction-Integration Model (Kintsch, 1988, 1994, 1998).

Over time, the field of reading research has transformed in both gradual and dramatic ways. Examining these shifts historically allows researchers to reflect with the benefit of hindsight and understand the larger context of contemporary models (Alexander & Fox, 2004; VanSledright, 2002). While it is impossible to exhaustively examine the multitude of reading theories proposed throughout history, a perspective that emphasizes the most influential models is useful in understanding the field (Tracey & Morrow, 2006). The following sections offer such an historical perspective. Models were selected based on their importance historically, their contemporary popularity, because they were precursors to existing connectionist models, or because they had connections to theories of expertise development. The selected perspectives are presented according to the era they were proposed. The connectionist models that frame this study are addressed in more detail throughout the historical review.

Foundational theories (400 B.C. to 1899). The history of reading models is entangled with the history of general education theories. Mental Discipline Theory and Associationism are two early theories that provided a foundation for future models of reading

(Bigge & Shermis, 1992; Brumbaugh & Lawrence, 1985; Gutek, 1972; Schwartz & Robbins, 1995; Sternberg, 1996). These models were selected for review because they influenced the contemporary connectionist models of reading being used in this study.

Mental Discipline Theory. The origins of Mental Discipline Theory can be traced to Greek philosophers Plato (ca. 428-437 B.C.) and Aristotle (384-322 B.C.). Unlike other thinkers of their era (e.g., Homer) who used myths and legends to understand the universe, Plato and Aristotle tried to explain the universe in rational terms (Gutek, 1972). The premise of Mental Discipline Theory is that the mind is like a muscle whose various components (e.g., memory, will, reason, and perseverance) need to be exercised regularly. The mind, like physiological muscles, is only strengthened through exercise and the mind can operate automatically after adequate exercise (Bigge & Shermis, 1992).

Mental Discipline Theory has influenced the educational and psychological literature for over 2,500 years (Brumbaugh & Lawrence, 1963). In the domain of reading, Mental Discipline Theory's influence can be seen in both instructional approaches and theoretical research (Tracey & Morrow, 2006). For example, the instructional technique of repeated reading, where students read a passage aloud multiple times and receive feedback and guidance from a teacher, has been connected to improvements in word recognition, fluency, and accuracy (Armbruster, Lehr, & Osborn, 2001). Researchers interested in the influence of repeated practice have shown that practice can improve reading achievement, oral language development, phonemic awareness, and exposure to print (Snow et al., 1998). These instructional approaches and research findings are consistent with Mental Discipline Theory.

Mental Discipline Theory is consistent with models of expertise development in that repeated practice is important to achieving expertise. Ericsson's notion of targeted practice is

consistent with the concept of the mind as a muscle. Novices who engage in practice that is targeted are strengthening their mind through activity that is neither too easy nor too difficult.

Associationism. Associationism is a theory of education and psychology that attempts to explain how learning occurs. According to this theory, learning happens when events or ideas become connected with one another in the mind. Aristotle, considered the first associationist, proposed three kinds of connections that assist memory and learning: contiguity, similarity, and contrast. Contiguity is the notion that events and ideas that occur together spatially and temporally become associated in the mind. Similarity is the idea that objects with similar features and properties become linked. Finally, contrast represents connections made between things that are different (e.g., light and dark) (Sternberg, 1996).

Both Mental Discipline Theory and Associationism remained the prominent educational theories of learning for centuries. From the time of Plato (ca. 428 B.C.) and Aristotle through the Enlightenment in the 18th century, these two theories remained at the forefront of thinking about education and learning (Bigge & Shermis, 1992; Gutek, 1972; Sternberg, 1996). One of the most influential associationists was John Locke (1632-1704), whose Tabula Rasa Theory provided a turning point in how learning was viewed. In arguing that people are born without internal, innate knowledge, Locke's theories turned attention away from the importance of knowledge that is already internal towards an emphasis on external influences on learning (Brumbaugh & Lawrence, 1985).

While Aristotle emphasized the internal mental connections necessary for learning, and Locke focused on the influence of external information, both theorists are considered Associationists since they were each interested in how knowledge was constructed and acquired (Brumbaugh & Lawrence, 1985; Tracey & Morrow, 2006). According to Sternberg

(1996), Associationism provided the foundation upon which behaviorism and models of cognition based on mental connections were built.

Associationism also had a significant impact on theories of expertise development. Aristotle's ideas of contiguity and similarity of mental objects is consistent with the characteristics of expertise compiled by Feltovich et al. (2006). Contiguity and similarity are captured by the generalization that experts have more integrated cognitive units and better abstract representations of information than novices. Ericsson's (1996a, 2002, 2004, 2006b) deliberate practice supports Locke's perspective that learning can be facilitated via external forces.

Behaviorist theories (1900 to 1960s). While the late 1800s and early 1900s saw increased research into the cognitive processing of reading as measured through perception (e.g., eye-movement research), by 1910 the focus had shifted to instructional approaches. This emphasis on teaching and testing was a result of the popularity of Behaviorism, and would remain predominant well into the 1960s (Thomas, 1996; Venezky, 1984; Woolfolk, 1998). Behavioral theorists believed that the study of observable actions related to learning was the only way to make psychology a true science as experimental manipulation would be empirically visible in these actions (Thomas, 1996). Behaviorist theories are all grounded in two assumptions. First, behavior is the result of a response to a stimulus. Second, external stimuli can be manipulated to modify behavior (Greeno, Collins, & Resnick, 1996).

The behaviorist movement changed the perspective from perceptual processing during reading to reading as a behavior consisting of isolated skills that could be independently reinforced to improve student achievement (Tracey & Morrow, 2006). These component skills included: visual discrimination (i.e., the ability to discriminate shapes and

letters), auditory discrimination (i.e., the ability to discriminate the sounds of the alphabet), vocabulary (i.e., word knowledge), and comprehension (i.e. understanding) (Woolfolk, 1998). Kame'enui, Simmons, Chard, and Dickson (1997) conducted research into the impact of direct instruction of reading as part of Project Follow Through. Founded on behavioral principles, direct instruction is the instructional approach whereby teachers guide students to activities based on the sub-skills of reading (e.g., phonics, vocabulary, comprehension) and provide feedback on performance. The study showed that direct instruction produced gains in basic skills (e.g., word recognition), cognitive skills (e.g., comprehension), and affective feelings (e.g., self-esteem). Two specific behaviorist theories have influenced the field of reading: Connectionism and Operant Conditioning. Each of these theories is based on Associationism and will be discussed in subsequent sections.

Connectionism. While Pavlov and Watson and other classical conditioning researchers were focused on the factors that preceded behavior, Edward L. Thorndike believed that stimuli that occurred after behavior could influence future behaviors. Thorndike's theory, Connectionism, consisted of four laws: the Law of Effect, the Law of Readiness, the Law of Identical Elements, and the Law of Exercise. The Law of Effect, also known as the Principle of Reinforcement, stated that if a behavior was followed by a satisfying change in the environment, the likelihood of the behavior being repeated increases. Conversely, if a behavior was followed by an unsatisfying change in the environment, the probability of the behavior repeating decreases. The Law of Readiness held that learning was facilitated when simpler tasks were completed before related, more complex tasks. The Law of Identical Elements stated that learning a second task was easier if the components of the task are similar to an already learned task. Finally, the Law of Exercise held that the more

stimulus-response connections are practiced, the stronger the links become (Slavin, 1997; Thorndike, 1903).

Since it was first proposed in 1903, Connectionism has influenced reading instruction and reading researchers. According to Hiebert and Raphael (1996), each of Thorndike's laws was integrated into classroom instruction in the early 1900s. The Law of Effect was applied when teachers would praise students for correctly answering reading questions. The Law of Readiness was incorporated by educators in the sequencing of words students were to learn. The Law of Identical Elements and Law of Exercise led to students practicing specific sets of target words to ensure strong connections. More recently, Koskinen (1993) applied the Law of Identical Elements to the reading-writing connection where it was shown that students who received integrated literacy (i.e., both reading and writing) instruction performed better than students in the control group on measures of language development, comprehension, and writing.

Deliberate practice encompasses many of the features of Thorndike's Connectionism. Feedback received during an activity is representative of the Law of Effect. The Law of Readiness, which represents the ordering of tasks from simple to complex, is consistent with the idea of targeting whereby learners receive tasks in a specific order based on their ability. Lastly, the Law of Exercise is realized in deliberate practice whereby it takes intense practice distributed over time to achieve expertise.

Operant Conditioning. Established by B.F. Skinner, Operant Conditioning is a behaviorist theory based on the premise that Classical Conditioning and Connectionism only explain a small amount of learned behaviors. Skinner believed that learning is not automatic and unintentional, as learned responses are produced voluntarily. People deliberately operate

on the environment to produce various consequences (Woolfolk, 1998). By focusing on reinforcement and behavior, Skinner believed it was possible to change behavior (Skinner, 1953).

According to Woolfolk (1998), Skinner's behaviorism completely changed how educators and researchers thought about teaching and learning. Skinner called the application of his theory to the classroom "programmed instruction." That is, instruction is decomposed into small, successive steps that are designed to maximize the probability of student success while minimizing the chances of failure or frustration. As each small step is completed, learners receive reinforcement (Brumbaugh & Lawrence, 1985). Operant Conditioning Theory also influenced the construction of educational software. The behavioral principles of breaking up complicated tasks into manageable pieces and providing immediate feedback to learners based on performance on learning tasks often appear in educational software (Thomas, 1996).

Skinner's Operant Conditioning was also explicitly applied to the reading domain. For example, Burns and Kondrick (1998) conducted research where readers were given tokens for correctly reading target words and paragraphs and answering comprehension questions. These earned tokens could be exchanged for money. The pretest and posttest results, as measured by the Woodcock Reading Mastery Tests-Revised (Woodcock, 1987) and Gray Oral Reading Tests-Revised (Wiederholt & Bryant, 1986), showed a significant improvement in reading.

Like Connectionism, Operant Conditioning theory is consistent with the principles of deliberate practice. Skinner's emphasis on the importance of feedback during a learning experience is one of Ericsson's pillars of expertise development. Operant Conditioning also

involves targeted activity in that tasks are decomposed into manageable pieces that can be mastered with minimal chance of failure.

Constructivist theories (1920s – Present). Constructivism is a theory of learning based on the active construction of knowledge by learners. Learning occurs when new knowledge is integrated with existing knowledge (Woolfolk, 1998). According to Smith (1971), learning is not an event that occurs occasionally based on stimulus or reinforcement; rather, learning is ongoing and occurs naturally.

Constructivism is based on three guiding principles. First, learning takes place through unobservable internal mechanisms. This is in direct opposition to the behaviorist perspective that learning is observable. Second, learning results from the continual hypothesis-testing conducted by a learner whereby hypotheses about a concept (e.g., word, idea) are either proven or disproven and updated during a learning experience. Lastly, learning results from inferences, or filling in gaps in knowledge (Ruddell & Ruddell, 1995).

Constructivism has been applied to the field of reading as an explanatory mechanism for how readers construct messages (e.g., comprehend) when reading (Anderson & Pearson, 1984). Since the 1920s, different constructivist theories have heavily influenced perspectives on reading: Schema Theory, Transactional Response Theory, and Psycholinguistic Theory (Anderson & Pearson, 1984; Gutek, 1972; Rosenblatt, 1938/1983; Tracey & Morrow, 2006). Each of these theories will be discussed in the following sections.

Schema Theory. Adherents of Schema Theory are interested in how knowledge is created, organized, and used by learners (Tracey & Morrow, 2006). According to the theory, each individual has schemas, or knowledge structures, that organize what is known about an entity (e.g., people, places, things, language, processes, and skills) (Bartlett, 1932). The

complexity of a schema greatly influences learning: more elaborate schemas result in easier learning and less developed schemas can make learning new information challenging (Anderson & Pearson, 1984).

Schema Theory holds that knowledge structures are flexible and constantly changing. These changes occur through three processes: accretion, tuning, and restructuring. Accretion occurs when a learner receives information that does not require the modification of the existing schema. On the other hand, tuning occurs when new information is incorporated into an existing schema. Finally, restructuring happens when a schema must be recreated because the antiquated version was insufficient (Anderson & Pearson, 1984; Hiebert & Raphael, 1996; Tracey & Morrow, 2006).

Anderson and Pearson (1984) applied Schema Theory to the field of reading. They argued that readers have different sets of schemas that are used during reading: schemas for content (e.g., people, places, and things), schemas for reading processes (e.g., word recognition, skimming, inference, and summarizing), and schemas for different text structures (e.g., narrative, informative, and persuasive). Differences in reader comprehension can be attributed to differences in reader schemas. To activate prior knowledge and improve comprehension, Anderson and Pearson proposed the use of brainstorming and webbing (i.e., structured ideas).

Other reading researchers have framed their research using Schema Theory. For example, Droop and Verhoeven (1998) researched the impact of background knowledge on reading comprehension in first and second language reading. In the study, third grade students read three texts of similar difficulty with varying contexts: their own cultural background (i.e., Dutch), a foreign culture (i.e., Moroccan and Turkish), and a neutral

culture-free context. After collecting read-aloud, retelling, and comprehension assessment data, it was shown that student comprehension was highest when reading text in their first language that was set in their own culture. This finding was consistent with Schema Theory as the background knowledge of the students provided a benefit. For students reading in their second language, there were no noticeable differences in student performance. Droop and Verhoeven attributed this finding to a lack of language proficiency limiting the influence of prior knowledge.

Schema Theory has had a lasting impression on the field of expertise development. Expertise researchers have argued that experts possess more integrated cognitive units than novices and that expertise involves abstract representations of information (Feltovich, et. al, 2006). This emphasis on the structure of knowledge can be traced back to Schema Theory.

Transactional Response Theory. Proposed by Louise Rosenblatt (2004), Transaction Response Theory extends Schema Theory to the reading construct. This model is based on the premise that no two readers or pieces of text are identical. The transaction between a reader and a passage of text is unique and individualized based on the schema and emotions of the reader (Rosenblatt, 1938/1983). The meaning constructed during the reading process resides neither in the reader nor the text. Rather, meaning and comprehension come into existence in the transaction between reader and text. This transaction can be influenced by a variety of factors such as the context in which the reading occurs, the reader's purpose or the reader's interest level. In defining reading, Rosenblatt referenced William James' concept of a "choosing activity." That is, according to Rosenblatt, reading is a constantly self-revising impulse that directs selection, synthesis, and organization (Rosenblatt, 2004).

Transactional Response Theory is built on the assumption that all readers adopt a particular stance when reading. This stance is conceptualized as a continuum that ranges from “predominantly aesthetic” to “predominantly efferent.” Readers who read from an aesthetic position “live through” the text by focusing on the feelings, ideas, situations, personalities, emotions, tensions, conflicts, and resolutions that the text evokes. In contrast, efferent readers are focused on what is to be extracted and retained after reading (Rosenblatt, 2004).

Rosenblatt’s Transactional Response Theory influenced both classroom instruction and reading research. In the classroom, teachers who employ “text to self,” “text to text,” and “text to world” connections between reader and text are applying Transactional Response Theory (Tracey & Morrow, 2006). Zarrillo and Cox (1992) presented readers with a wide range of post-reading activities designed to stimulate both efferent and aesthetic responses (e.g., literature journals, talking about books, storytelling, puppetry, dioramas, and multimedia production). Bean and Rigoni (2001) showed that students reading multicultural novels used the text as a way of making sense of both their own lives and the lives of their peers and others in their community. This finding is consistent with the Transactional Response Theory perspective that readers can connect with text on a variety of levels.

Rosenblatt’s emphasis on the activity of the reader is related to Ericsson’s deliberate practice. Self-directed activity, a component of deliberate practice, encompasses a learner choosing to engage in a learning experience. In borrowing James’ notion of a “choosing activity”, Rosenblatt stresses that reading is an impulse that directs activity. This is consistent with self-directed practice.

Psycholinguistic Theory. The field of psycholinguistics is positioned at the intersection between psychology and language. Psycholinguistic researchers study how the language system interacts with how information is acquired, interpreted, organized, stored, retrieved, and employed (Shannon, 1990). From the psycholinguistic perspective, reading is primarily a language process where readers rely on language cueing systems to rapidly read text. The three most referenced cueing systems are syntactic, semantic, and graphophonic. Each of these systems helps readers to predict the upcoming words and sentences. Syntactic cues are related to the grammatical structure and syntax of language. Semantic cues are linked to the meaning of the words and sentences. Finally, graphophonic cues are derived from the visual patterns of letters and words and their sounds. According to Psycholinguistic Theory, young readers are able to use these cueing systems unconsciously after they have been internalized through oral language (Goodman, 1965a, 1965b, 1967, 1968; Smith, 1971).

Psycholinguistic Theory holds that readers use their knowledge about both language and the world to drive thinking during a reading experience. Readers use this knowledge while reading to make predictions, or hypotheses, about what the text will say. The hypothesis testing that happens rapidly and unconsciously is a critical component of Psycholinguistic Theory. When the text matches a given hypothesis, reading proceeds smoothly without pause. On the other hand, when text differs from a hypothesis, reading pace slows as the reader attends to the letters and words more closely. Psycholinguistic theorists believe that this hypothesis-testing process allows readers to rapidly move through a text by sampling words and comparing them to their prediction (Goodman, 1965a, 1965b, 1967, 1968; Smith, 1971). Goodman (1967) described reading as a “psycholinguistic

guessing game” where the reader tries to reconstruct meaning based on their knowledge of language and the world.

Psycholinguistic Theory has made a lasting impression on reading instruction and research. The use of running records during guided reading is an instructional practice that can be directly traced to this theory. During guided reading lessons, teachers meet with small groups of students of similar reading ability and provide them with text matched to their ability. During the lesson, each student reads aloud, and the teacher tracks performance using a running record. These records provide diagnostic information about the “miscues” of the reader. From a Psycholinguistic perspective, these miscues can be used by a teacher to determine what types of cueing systems a student is using during reading (Clay, 2000; Fountas & Pinnell, 1996).

Psycholinguistic Theory laid the foundation for Whole Language Theory, a theory of learning that has been influential since the 1980s (Smith, 1971). Whole Language Theory is an instructional approach that stresses the use of authentic literature in the context of meaningful and cooperative experiences designed to enhance student interest in learning. According to this perspective, like oral language, reading is a natural process that is acquired by immersion in high-quality literacy environments and exposure to meaningful literature. This theory is based on the belief that reading, writing, speaking, and listening are interconnected and that improvement in one domain will promote improvements in the others (Bergeron, 1990; Smith, 1971).

While Psycholinguistic Theory has been highly influential in impacting literacy instruction, from a research perspective, the value of Psycholinguistic Theory for explaining the reading process has declined in recent years. According to Stanovich (2000), the

Psycholinguistic hypothesis that the superior word-recognition skills of advanced readers are a result of superior context-use skills (i.e., using the surrounding words to infer meaning) is false. Research has consistently shown that better readers do not rely on context for word recognition (Stanovich, 2000).

Information processing theories (1950s-Present). In the mid-20th century, the cognitive revolution and growing ubiquity of computers shifted researcher interest away from observable behaviors towards understanding and modeling unobservable mental processes. Information processing theories of reading attempt to describe the internal workings of the mind during the reading process (Hiebert & Raphael, 1996; Schwartz & Reisberg, 1991; Sternberg, 1996; Woolfolk, 1998). Dozens of theories that describe reading have been proposed in the last 60 years; however, the number of models and slight variations between theories make a comprehensive review difficult (Ruddell & Unrau, 2004). The remainder of this section describes the most influential information processing models of reading. The components and principles of each model are discussed, with special attention paid to the models that provide the theoretical foundation for the present study.

Rauding Theory. Carver (1977) tried to quantify the reading process by investigating the most important component variables. According to Carver, readers employ five different cognitive skills during a reading task: skimming, scanning, rauding, learning, and memorizing. Rauding, a term proposed by Carver, is the most commonly used skill by a reader and is commonly thought of as “reading.” The term describes the comprehension process that occurs whether a student is reading (i.e., using the printed words on a page) or auditing (i.e., listening to spoken words). Rauding proceeds at a constant rate and involves no studying. When readers accelerate their reading speed, they shift to scanning or skimming.

Readers are forced to slow down to learn or memorize when presented with difficult text or when given extra time to reread (Carver, 1992a).

Rauding Theory uses a variety of factors to determine reader comprehension. Rauding efficiency level, Carver's term to describe general reading ability, consisted of two factors that research suggests are the critical components of reading comprehension: rauding accuracy level (i.e., knowledge of vocabulary) and rauding rate level (typical rauding rate) (Carver & Leibert, 1995). Rauding Theory uses mathematical relationships between rauding efficiency level, rauding accuracy level, rauding rate level, text difficulty, and time provided for reading to estimate comprehension. A critical concept of Rauding Theory is that comprehension is not influenced by prior knowledge or text type. The reader is conceptualized as an information processor and reading depends on the reader's accuracy level and rate level (Carver, 1992a).

According to Carver and Leibert (1995), Rauding Theory has generated two important hypotheses that have been investigated via research. First, reading improvement requires readers to engage with text that is closely matched to their ability. This idea is consistent with the targeting of activity required by Ericsson's deliberate practice. Second, readers should read text that is easy enough to maintain an adequate reading rate. Reading text that is too easy will not expose students to new words, and reading text that is too hard will slow the reading rate and hinder comprehension.

Automatic Information Processing Model. The Automatic Information Processing Model is a bottom-up theory of how readers process text. It is considered bottom-up because the reading process is conceptualized as a series of steps that begin when a reader sees text and ends with comprehension. Higher-level operations (e.g., meaning making) do not effect

lower-level operations (e.g., word recognition). Proposed by LaBerge and Samuels (1974), the Automatic Information Processing Model is one of the most-often referenced reading theories in the research literature (Blanchard, Rottenberg, & Jones, 1989). The foundational concept of this model is that automaticity in decoding ability separates fluent readers from struggling readers (Samuels, 2004).

There are five components in the Automatic Information Processing Model: visual memory, phonological memory, episodic memory, semantic memory, and attention. The reading process begins with the visual processing of text in the visual memory. The visual memory consists of a series of feature detectors that analyze the curves and lines of incoming letters and combine them into spelling patterns and word codes. Through repeated exposure and practice, letters become unitized and perceived as a single unit. As these units accumulate and their creation is increasingly automated, the attention a reader expends in the visual memory decreases, allowing mental resources to be allocated elsewhere (LaBerge & Samuels, 1974; Samuels, 2004; Wolf & Katzir-Cohen, 2001).

After a text stimulus is processed by the visual memory, the phonological memory attaches sounds to the visual units. The acoustic units of the phonological memory are features, phonemes, syllables, and words. Like the units of the visual memory, the units of the phonological memory are arranged in a hierarchy. While the visual memory moves from features to letters to spelling patterns to words, the phonological memory proceeds from features to phonemes to syllables to words (Samuels, 2004).

Episodic memory and semantic memory are the third and fourth components of the model. Episodic memory applies a time, place and context marker to events and knowledge and is responsible for a reader recalling the details of a past event. This structure is organized

around the “when,” “where,” and “who” of past events. Semantic memory is responsible for attaching meaning to words and for enabling comprehension. Essential information (e.g., learning how to sound out words) is stored in semantic memory, while episodic memory stores associated details (e.g., where I was sitting when I learned how to sound out words) (LaBerge & Samuels, 1974; Samuels, 2004).

The central component to the Automatic Information Processing Model is attention. According to Samuels (2004), attention is a prerequisite to learning and comprehension. LaBerge and Samuels (1974) describe two distinct types of attention: external and internal. External attention describes observable behavior that involves the visible direction of sensory organs (e.g., focusing the eyes on the words of the page). Not surprisingly, as external signs (e.g., looking in books) of attention increase, reading scores increase in both males and females (Samuels & Turnure, 1974).

While external attention is important and necessary for comprehension, internal attention is even more crucial. Internal attention has three characteristics: alertness, selectivity, and limited capacity. Alertness describes how vigilant and active a reader is when trying to understand a text. Selectivity refers to the degree to which readers are thoughtful about choosing which aspects of their context they will process (e.g., multiple lines of text). Finally, limited capacity refers to the innate processing limitations of the human mind. Internal attention, while not readily observable like external attention, is critical to the reading process.

According to Samuels (2004), reading is a two-step process of decoding printed material and comprehending the decoded words. In framing attention as requiring cognitive energy, LaBerge and Samuels (1974) argued that beginning readers divide attention (i.e.,

cognitive resources) between decoding and comprehension. As decoding becomes automated, it requires less attention, freeing up mental energy to focus on comprehension. Therefore, the level of decoding automaticity separates fluent readers from struggling readers (LaBerge & Samuels, 1974; Samuels, 2004). This emphasis on automaticity is consistent with the expertise literature in that automaticity is critical for a novice to become an expert (Feltovich et al., 2006; Schneider & Shiffrin, 1977).

Interactive and Interactive-Compensatory Models. The Automatic Information Processing Model is a bottom-up model of reading where text input is processed sequentially and higher-level processors do not influence lower-level functions. In the late 1970s and 1980s, new models of reading were proposed that challenged the bottom-up paradigm (Ruddell & Unrau, 2004). The remainder of this section will describe the limitations of linear processing models that led to the earliest interactive models.

Rumelhart (1977, 1994) identified three major limitations of the linear, non-interacting stages initially proposed by Gough (1972) and expanded by LaBerge and Samuels (1974). First, bottom-up models argue that letters are perceived first and combined into higher level units later; however, research suggests that letter perceptions often depend on surrounding letters. The research of Pillsbury (1897), Huey (1908/1968), and Reicher (1969) showed that letter accuracy and speed perceptions increased when letters appeared as part of a word. These findings suggest that word-level perceptions (i.e. lexical knowledge) can influence letter-level perceptions. Like lexical knowledge, orthographic knowledge has been shown to improve perception of letter strings (McClelland & Johnston, 1977; Miller, Bruner, & Postman, 1954; Reicher, 1969).

The second limitation of linear processing models of reading is that reader perception of words depends on the syntactic environment (Rumelhart, 1977, 1994). Research into oral reading has provided evidence of the influence of syntax on word perception. Kollers (1970) showed that 70% of the substitution errors (i.e., an incorrect word substituted for a correct word) made by adult readers were the same part of speech as the correct word. Similarly, Weber (1970) determined that over 90% of the substitution errors made by first graders were grammatically consistent. These findings support the hypothesis that syntax surrounding a word can influence perception. Supporting results were offered by Stevens and Rumelhart (1975) as well as Miller and Isard (1963).

Like the syntactic environment, the semantic environment of a word influences word perception; however, bottom-up models argue that meaning is constructed long after words are perceived. In measuring reaction time between different pairs of words (i.e., related words such as bread and butter, and unrelated words such as bread and doctor), researchers determined that related words could be read faster, suggesting that the identification of the first word allows the reader to process the second word more quickly (Meyer & Schvaneveldt, 1971; Meyer, Schvaneveldt, & Ruddy, 1974; Schvaneveldt & Meyer, 1973). Other researchers have also provided evidence that the higher-level process of meaning can influence the lower-level process of word identification (e.g., Tulving & Gold, 1963; Tulving, Mandler, & Bauml, 1964).

Based on the limitations of the Automatic Information Processing Model, Rumelhart (1977, 1994) proposed the first non-linear model of reading: the Interactive Model. In this model, reading occurs when text stimulus enters the reading system, where a variety of processors operate simultaneously. After the features of text are extracted, the information is

passed to the pattern synthesizer that acts as a workspace for the knowledge processors. The four knowledge processors (i.e., orthographic, lexical, syntactical, and semantic) work together and complement each other. The orthographic processor is related to visual input while the lexical processor handles word knowledge. The syntactic processor refers to word order within sentences and semantic knowledge is related to meaning construction.

Stanovich (1980) extended Rumelhart's Interactive Model when he proposed the Interactive-Compensatory Model. This model is identical to Rumelhart's, with the addition of a compensatory mechanism between knowledge processors. Stanovich argued that, in addition to operating simultaneously and cooperatively, the four processors (i.e., orthographic, lexical, syntactical, and semantic) act in a compensatory manner. If a given processor is not working efficiently or lacks sufficient data, the other processors will compensate. For example, if parts of a text being read are blurry, the orthographic processor might struggle identifying words so the other processors will try to compensate (e.g., using syntactic and semantic information to guess the illegible word).

The Interactive Model and Interactive-Compensatory Model are historically important as the first non-linear representations of the reading process. Over the last 35 years, multiple iterations of these models have been proposed. The Parallel Distributed Processing Model, described in the next section, is an extension of the Interactive Model that continues to make a lasting impression on the reading research community and was applicable to this study (Ruddell & Unrau, 2004; Tracey & Morrow, 2006).

Parallel Distributed Processing Model. The model originally proposed by Rumelhart (1977, 1994) was the precursor to a new series of interactive models, known as parallel distributed processing models. These models are built on the assumption that learning occurs

as the learner encounters relationships among patterns and events. The utility of these models is that, like the Interactive Model, they are neither bottom-up nor top-down. Rather, all relevant processes act simultaneously and interactively, sharing information. This continuous cooperation and coordination is shaped by the past experiences of the reader (Adams, 1994; Bereiter, 1991; Rumelhart & McClelland, 1986a; Seidenberg & McClelland, 1989). The Parallel Distributed Processing Model is one of the three models used to ground the proposed study because it frames the reading process as a series of interconnected processes (i.e., reciprocal rather than linear processors) that function based on the past experiences of a reader.

As a connectionist model, the Parallel Distributed Processing Model is built on the assumption that all cognitive information is stored as a series of connections between units and that these connections become stronger with repeated stimulation. The cognitive units are interconnected both within and between model components, and the connection strengths will vary based on past exposure (Brown, 1987; Norris, 1994; Zorzi, Houghton, & Butterworth, 1998a). The network of connections between units has led some researchers to call this connectionist approach a *neural network* as units can influence neighboring units through shared connections (Adams, 1994; Rumelhart & McClelland, 1986a).

In a connectionist model, learning occurs when the connections between particular units strengthen over time as a result of learner experience. When a connection is first formed between two units, it is associated with a random, but small connection weight. Over time, as the connection is strengthened, the connection weight increases (Rumelhart, Hinton, & Williams, 1986). Through repeated exposure over time, a connection between units can become strong enough that the learner is able to access the information automatically

(Adams, 1994). This strengthening process occurs gradually over time, and a variety of researchers have attempted to use the gradual unfolding of the neural network to investigate the development of reading (e.g., Ans, Carbonnel, & Valdois, 1998; Berninger et al., 2000; Brown, 1997; Harm & Seidenberg, 1999; Hulme, Snowling, & Quinlan, 1991; Manis, Seidenberg, Doi, McBride-Chang, & Peterson, 1996; Seidenberg & McClelland, 1989; Van Orden, Pennington, & Stone, 1990; Zorzi, Houghton, & Butterworth, 1998b).

The Parallel Distributed Processing Model was originally proposed by Rumelhart and McClelland (1986b); however, the Seidenberg and McClelland (1989) version is most often cited (Berninger et al., 2000). Adams (1990, 1994) framed her literacy research and instructional approach using the Parallel Distributed Processing Model. The model consists of four processors: the orthographic processor, the phonological processor, the meaning processor, and the context processor (Adams, 1990, 1994; Rumelhart & McClelland, 1986b; Seidenberg & McClelland, 1989). These separate processors are compensatory and operate simultaneously and cooperatively during reading. The model presents each processor as a discrete operating unit; however, this is solely for descriptive purposes. The associations (i.e., connections) between units of knowledge depends on the ways they have become interrelated and connected through experience, not on the particular processing unit in which they reside (Adams, 1994). The interrelations between knowledge units, as they are excited or inhibited, are responsible for the reading process (Adams, 1994; Rumelhart & McClelland, 1986b; Seidenberg & McClelland, 1989).

The Orthographic Processor. The reading process begins in the orthographic processor when text is visually perceived. English readers process text in a left-to-right, line-by-line, and word-by-word manner. Each word is processed whether reading connected text

or isolated words, regardless of text difficulty, and independent of semantic, syntactic, or orthographic predictability (Adams, 1990). The idea that a reader processes nearly every word is a finding that has been replicated broadly (Patterson & Coltheart, 1987). The orthographic processor can be considered a database of orthographic knowledge (Byrnes, 2001).

While a reader processes almost every word during reading, how the eyes move has received researcher attention. Just and Carpenter (1987) showed that the eyes of a reader do not move smoothly through a text during reading. Rather, a reader's gaze will jump from word to word, focusing briefly on the center of each string of letters before jumping to the next. Readers seldom skip more than one word in a sequence, and the word skipped is often a short frequent word (e.g., of, the, a).

Adams (1994) provided two useful illustrations of how the neural network of a reader works during the word recognition stage of reading. The first example is what occurs when a reader's gaze arrives at the word "the." Since "the" is a frequently occurring and familiar word, the letters will be strongly interconnected within the mind of the reader. When the reader sees the word, the units corresponding to each letter are stimulated. Due to the strong connections between the letters, they will pass stimulation to the others. This simultaneous stimulation will lead to the letters being recognized quickly and to remain in the reader's mind as a familiar, cohesive spelling pattern. In contrast, when a reader views a nonword such as "tqe," the activation is not nearly as effortless. The "t" and "e" units will stimulate each other as well as the "h" unit; however, the "h" unit does not pass back any stimulation since it was not seen on the page. Meanwhile, the "q" unit will pass stimulation to its strongest connected neighbors (e.g., the letter "u"). Eventually the visual stimulation from the

page will bring about maximum stimulation on the necessary letters and the reader will see “tqe”; however, the perception of each letter will have taken longer (McClelland & Rumelhart, 1981).

The inter-letter associations of a reader facilitate more than just word recognition. These connections assist a reader in processing letter order and breaking words into syllables (Adams, 1994). Both skilled and unskilled readers suffer from a physical challenge when processing letter order; however, skilled readers never make errors when reporting word order and unskilled readers often do (Estes, 1977). The assumption that this limitation is a result of perceptual deficiencies has been proven incorrect (Liberman, Shankweiler, Orlando, Harris, & Berti, 1971). Rather, research suggests that skilled readers overcome this word ordering limitation through well-formed connections between letters. During the perception of text, this associated knowledge helps a skilled reader to parse and support the noisy transmission of letter order. Adams (1979) showed that good readers rarely make mistakes when reporting letter order for both real words and regularly spelled non-words (e.g., mave, jome). Yet, when presented with irregular non-words (e.g., xtsio, gtsi), good readers make as many errors as poor readers.

A well-formed network of inter-letter associations can also assist readers in breaking words into syllables. Research suggests that skilled readers’ ability to read long words is linked to their ability to divide words into syllables (Mewhort & Campbell, 1981). Skilled readers partition words into syllables automatically during the course of reading. The connections between letters enable readers to detect boundaries based on weak connections between letters. For example, in a skilled reader, the connection between “d” and “n” will be much weaker than the connection between “d” and “r” as the former letters appear together

less frequently than the later pair of letters. These connection weights between letters allow skilled readers to divide long words into syllables (Adams, 1981; Seidenberg, 1987).

The connections formed between letters are critical to the reading process as they allow readers to automatically recognize letters and words. Through experience and repeated exposure, the connectionist network of a reader will respond to larger sequences of letters that represent whole words and spelling patterns. The learned associations between individual letters are responsible for the easy manner in which skilled readers process text (Adams, 1994).

The Phonological Processor. A well-defined network of orthographic units allows a fluent reader to recognize words immediately, without needing to “sound out” parts of words (Spoehr, 1981). Research results suggest that skilled readers automatically produce sound-based representations, even though it is not required (Perfetti, Bell, & Delaney, 1988; Tannenhaus, Flanigan, & Seidenberg, 1980; VanOrden, 1991). Like the other processors in the Parallel Distributed Processing Model, the phonological processor consists of a complex network of units. The auditory record of any word, syllable, or phoneme is represented by the activation of a particular, inter-related set of connections between units (McClelland & Elman, 1986).

The phonological processor is connected to both the orthographic processor and the meaning processor. As words are perceived visually and the corresponding orthographic units are stimulated, the appropriate connections to phonological units are stimulated. Meanwhile, the stimulation of phonological units excites orthographic units and helps inform letter and word perception. Phonological activation stimulates units in the meaning processor

which also receives input from the orthographic processor. In this manner, meaning is created both from orthographic and phonological stimulation (Adams, 1990, 1994).

The phonological processor provides a reader with two critical functions. First, it acts as a redundancy system for the orthographic processor in enabling successful word recognition. Phonological translation is not required for recognizing words that have been seen frequently. The connections between the letters of these words are strong enough that, when encountered, word recognition occurs immediately (Adams, 1990; Byrnes, 2001). The challenge for developing readers is that there are relatively few frequently occurring words (Kucera & Francis, 1967). Carroll, Davies, and Richman (1971) predicted that fifty percent of the words students see are accounted for by only 109 unique words. While these 109 words will be seen often enough to be immediately recognizable, the tens of thousands of remaining words a reader will encounter in their lifetime will not be seen frequently enough to become automatically recognizable. The phonological processor acts as a back-up system for recognizing less-familiar words by harnessing existing spelling-sound associations to create phonological representations. This ensures that words encountered less frequently will be recognized with the speed required for comprehension (Adams, 1990, 1994; Byrnes, 2001).

The second critical function of the phonological processor is that it enhances a reader's ability to process and remember text. The language comprehension system works with cohesive grammatical units (e.g., phrases or sentences). Whether listening or reading, the words of a message are perceived one by one and tentatively interpreted at the time of initial perception; however, the message is only fully interpreted at the conclusion of the phrase or sentence (Jarvella, 1971; Kleiman, 1975). Fluent readers will process each word of

a sentence, and pause at the sentence ending while they construct meaning (Just & Carpenter, 1987). For a reader to reflect on their interpretation of a segment of text, the entire clause or sentence must exist in memory. As the human auditory system is designed to recall ordered patterns of information, the phonological representation of a text persists long enough for a reader to interpret meaning (Adams, 1990; Byrnes, 2001). Research results suggest that preventing skilled readers from subvocalizing (i.e., thinking or speaking the words to one's self) severely disrupts their ability to comprehend streams of text (Baddeley, 1979; Levy, 1977, 1978; Waters, Caplan, & Hildebrandt, 1987).

The Meaning Processor. The meaning processor, the third component of the Parallel Distributed Processing Model, attaches meaning to words. Like the orthographic and phonological processors, the meaning processor consists of a connected network of meaning units. These units do not correspond to whole words; rather, word meanings are represented as inter-associated sets of primitive meaning elements. This network of meaning allows a reader to focus on the various meanings of a word based on context. The units of the meaning processor are stimulated by the orthographic and phonological units as well as context processor units (Adams, 1990, 1994).

Like the rest of the Parallel Distributed Processing Model, the meaning processor operates in a connectionist manner where meaning units are strengthened through repeated stimulation. When an unknown word is first encountered, the spelling and pronunciation stimulations arrive at the meaning processor from the orthographic and phonological processors. As the word meaning is not known, the context of the word begins to strengthen certain connections. A single context typically does not provide enough information to

determine word meaning; however, repeated exposures will enhance and modify the existing connections representing word meaning (Adams, 1990).

Learning the meaning of a word is a gradual and imprecise process. The likelihood that a reader will learn a word from a single exposure in a meaningful context varies from 5% to 20% (Nagy, Anderson, & Herman, 1987). The average fifth grader reads 1 million words per year, of which approximately 650,000 are read out of school (Nagy, Herman, & Anderson, 1985). Of these words, 16,000 to 24,000 will be unknown (Anderson & Freebody, 1983) which means that the average fifth grade student learns between 800 and 1,200 words per year just from context during reading. This is a substantial chunk of the 3,000 new words students are expected to learn per year (Miller & Gildea, 1987). Students who read more words yearly will have better developed meaning processor units. Research suggests that students in the 90th percentile read 200 times more text than students in the 10th percentile (Nagy, Anderson, and Herman, 1987).

Besides developing word meaning through exposure to large amounts of reading, direct vocabulary instruction has been shown to increase both word knowledge and reading comprehension (Stahl & Fairbanks, 1986). This instruction should provide examples of the word used in context as well as definitional information. Both the number of times a student encounters a word and the richness of the exposures predict how strongly a word's meaning is understood. Through rich and varied exposures, readers gain advantages in understanding the connotation and sub-meanings of a word (McKeown, Beck, Omanson, & Pople, 1985).

The Context Processor. The context processor manages the construction and monitoring of the ongoing interpretation of the text being read. More specifically, the context processor selects word meanings that are appropriate to a text (Adams, 1990). The context

processor sends stimulation to the meaning processor based on the word meaning it expects; however, it does not prevent the stimulation of inappropriate meanings. Seidenberg, Tannenhaus, Leiman, & Bienkowski (1982) reported that people showed signs of interpreting ambiguous words in sentences multiple ways. For example, when presented with the sentence ‘*Kate saw several spiders, roaches, and bugs*’, people briefly registered the last word to mean both “insects” and “spying devices.” This suggests that the context processor does not eliminate word meanings.

The context processor helps a reader to quickly and efficiently process a text. It is not designed to replace the words of the author. Studies have shown that context will only affect reading accuracy and speed of fluent readers when the researcher disrupts orthographic processing (Stanovich, 1980, 1984). In contrast, context has been shown to influence word-recognition performance in young and disabled readers (Biemiller, 1970; Weber, 1970).

The Parallel Distributed Processing Model can be closely linked to theories of expertise development. Consistent with contemporary perspectives of expertise acquisition, the Parallel Distributed Processing Model positions the experience of the reader as the critical element to the acquisition of reading ability. Through repeated exposures to words and sentence structures over long periods of time, another component of expertise development, the neural network of a reader becomes better connected. Eventually through repeated practice, the network of connections is accessible without the reader spending cognitive resources. These connections within and between processors are enhanced, allowing the reader to move through a text automatically recognizing words and building meaning. This automaticity is also consistent with the expertise literature. The present study

used the Parallel Distributed Processing Model to frame how deliberate practice can lead to improved reading ability.

Construction-Integration Model. Both bottom-up models (e.g., Automatic Information Processing Model) and interactive models (e.g., Parallel Distributed Processing Model) provide theories about how the mind of a reader operates from letter perception through comprehension during reading. In contrast, the Construction-Integration Model (Kintsch, 1988, 1994, 1998) attempts to articulate how text representations are constructed and enable comprehension. Comprehension, according to Kintsch (2004), is automatic meaning construction through constraint satisfaction absent purposeful, conscious effort. This automatic comprehension is supplemented by purposeful problem solving when understanding breaks down.

The Construction-Integration Model is based on the connectionist principle of associated cognitive units and consists of three foundational ideas. First, a variety of mental representations of a text exist during reading. Second, comprehension is made possible through the cognitive processes of construction and integration. Lastly, working memory plays an important role in comprehension (Kintsch, 1988, 1994, 1998). Each of these topics will be addressed in subsequent sections.

Levels of Knowledge Representation. The Construction-Integration Model posits that text is represented in three distinct and connected layers: linguistic, conceptual, and situational. The linguistic level represents the actual words and phrases of a text. When reading instructional text, the exact wording does not remain in the linguistic level for long as the purpose of the reader is to acquire knowledge (Sachs, 1967). However, exact wording can be recalled when necessary (e.g., poem, joke, specific argument) (Kintsch & Bates, 1977).

The conceptual level of a text represents what the words and sentences mean. The cognitive units of the conceptual level are propositions. These propositions represent the meaning of a text and can be atomic (e.g., predicate and one or more arguments) or complex (i.e., network of atomic propositions). Propositional representations of text are intentionally crude as their purpose is not to represent the full meaning of complex text (Kintsch, 2004). Rather, propositions provide an aggregated idea unit that facilitates comprehension (Kintsch, 1974, van Dijk & Kintsch, 1983). According to Kintsch (1974), reading difficulty is a function of sentence length, word familiarity, and the number, coherence, and structure of the ideas presented.

Once propositions are created from the text, two different mental organizations are constructed that represent the conceptual level: microstructure and macrostructure. The microstructure represents the meaning of the words and sentences translated into idea units. In contrast, the macrostructure reflects the global organization of these ideas into higher-level units (e.g., setting, problem, resolution). Taken together, the microstructure and the macrostructure are the textbase, or the semantic meaning of a text (Kintsch, 1988, 1994, 2004; van Dijk & Kintsch, 1983).

When multiple skillful readers encounter the same text, they will construct similar textbases based on the words on the page. The textual representation is differentiated by the situation level. The situation level represents the ideas provided by the text integrated with background information from the reader's prior knowledge. The situation model depends on the purpose and goals of the reader, as well as the amount of prior knowledge (Kintsch, 1988, 1994). While the textbase is composed of propositional structures, the situation model is often represented with images as well as text. According to Kintsch (2004), educators are

interested in the situation model after reading as it represents the synthesis of the printed word with background information.

Cognitive Processes. According to the Construction-Integration Model, the various representations of a text (i.e. linguistic, conceptual, and situational) are formed and maintained through two cognitive processes: construction and integration. The construction phase produces an initial textbase and situation model that is detailed but possibly incoherent and contradictory. The integration phase modifies these representations into a coherent structure and incorporates that structure with the prior knowledge of the reader (Kintsch, 1988, 1994).

During the construction phase, a textbase is built through a series of four steps (Kintsch, 1988, 1994). First, based on the linguistic input, propositions are formed. While a linguistic parser is not part of the Construction-Integration Model, proposition construction is well defined (Kintsch, 1985; van Dijk & Kintsch, 1983). Proposition construction is completed on-line and frequently during reading, and inaccurate propositions are often created (Frederiksen, 1981; Kintsch, 1988). During the second step in the construction process, the formed propositions are used as cues for the retrieval of associated nodes from the neural network (Kintsch, 1988, 1994). This retrieval process is modeled after the work of Raaijmakers and Shiffrin (1981) as associated nodes are selected based on the strength of the connection between the cue and the node. After propositions are created, and associated information is acquired from the neural network of the reader, additional inferences are generated. It is impractical to expect that all the inferences required for comprehension are obtained through the first two steps. In addition to the undirected elaboration that occurs in the second step, more focused activity is necessary to generate the necessary inferences. For

example, bridging inferences are required when the textbase being constructed is incoherent (Haviland & Clark, 1974; Kintsch, 1974). During the final step in the construction process, the weighting factors of the interconnections between the selected elements are updated in the neural network.

According to the Construction-Integration Model, text comprehension is assumed to be organized into cycles that correspond to short phrases and sentences (Kintsch & van Dijk, 1978; Miller & Kintsch, 1980). During each cycle, a network is constructed based on the linguistic information and carry-over from the previous cycle. During the integration phase, this network is pruned, and merged with the existing network used for comprehension. The representation built by the construction phase is not yet useful for informing comprehension as it is often incoherent and inconsistent (Kintsch, 1988, 1994). Activation is spread around the propositional network that was constructed, and eventually the nodes that hang together remain active while outliers and isolated elements are deactivated (Kintsch, 2004; Rumelhart & McClelland, 1986b). This spreading activation occurs until the system stabilizes. If the integration process fails, new constructions are added to the network, and integration is attempted again. When the system stabilizes, the network consists of information at many levels: lexical information, text propositions, knowledge-based elaborations (e.g., inferences), and macro-propositions (Kintsch, 1988, 1994).

Working Memory. The Construction-Integration Model is built on the assumption that readers are able to overcome the limitations of working memory. Research suggests that people can only hold seven plus or minus two items in working memory at a time (Miller, 1956). However, successful text comprehension requires a reader to manage large amounts of information: perceptual features, linguistic features, propositional structure, microstructure,

macrostructure, situation model, control structure, goals, lexical knowledge, general knowledge, and episodic memory for prior text (Kintsch, Patel, & Ericsson, 1999). Each of these components would exceed the limits of normal working memory. Of course, people are able to understand and remember texts while they read, suggesting that they have overcome their mental limitations (Kintsch, 2004).

One theoretical concept that enables readers to overcome the limitations of working memory is long-term working memory (LTWM) (Ericsson & Kintsch, 1995; Kintsch, 1998; Kintsch et al., 1999). LTWM is theorized to be a subset of long-term memory that is directly retrievable via specific cues in short-term working memory. The links between short-term memory cues and corresponding long-term memory information must be stored in stable, fixed memory structures that allow direct and automatic retrieval. Accessing information in LTWM is automatic and does not require cognitive resources (Kintsch et al., 1999).

LTWM is a skill only accessible to experts in specific domains or disciplines. According to Kintsch (2004), reading and listening are practiced over a lifetime, and as long as the text being read is simple and about familiar topics, most adults can be considered expert readers. During reading, new nodes are formed in memory (i.e., propositions derived from the text) that are linked in complex patterns based on the text and the reader. The strength of the links between nodes varies based on the level of construction and integration. If comprehension is successful, a mental structure was synthesized that supports retrieval via LTWM. Readers who possess appropriate comprehension strategies, knowledge of linguistics, words, and topics, and the language skills necessary to construct and integrate during reading are able to harness LTWM to facilitate reading (Kintsch, 2004; Kintsch et al., 1999).

Theoretical perspectives of reading have a long history, dating back to the ancient times of Aristotle and Plato. Throughout history, reading has been viewed through the dominant theoretical lens of a particular time-period. Today, connectionist models remain prominent in the research literature about how a reader processes text. Connectionist models of reading (e.g. Automatic Information Processing Model, Parallel Distributed Processing Model, Construction-Integration Model) posit that the reading process functions due to the interconnections between cognitive units. As a reader experiences novel words and sentence structures, the associations between the nodes of the neural network are enhanced gradually over time. These connectionist models served as a foundation to frame how deliberate practice can improve reading.

Connectionist models of reading and deliberate practice. Connectionist models of reading frame the reading process as dependent on the interconnections between cognitive units. These connections are formed and strengthened gradually over time. This perspective is consistent with the principles of deliberate practice whereby novices engage in practice specifically tailored to foster expertise development. Each of the five components of deliberate practice can be described using features of existing connectionist models of reading.

Targeted practice. Targeted practice is necessary to develop expertise; by engaging in appropriately-challenging tasks, the learner does not reach a performance asymptote (Anderson, 1982; Ericsson, 2006b; Fitts & Posner, 1967). From the connectionist perspective, reading that is appropriately targeted will lead to the strengthening of specific connections between cognitive units. During reading, the ideal connections to be strengthened are those that are neither already sufficiently strong, nor those that are

impossible to stimulate. For example, if the challenge of a reading task is far below the ability of a reader, the connections being stimulated will likely already be strong. On the other hand, a reading task that is too difficult will severely complicate the construction and integration processes, making it unlikely that the connections between units will be strengthened. When reading text of appropriate challenge, a reader is stimulating the connections in the neural network that lead to the acquisition of strengthened associations.

Real-time corrective feedback. Novices engaged in deliberate practice use real-time corrective feedback to monitor progress and adjust performance (Ericsson, 2006b). From the connectionist viewpoint, the ongoing feedback a learner receives during reading provides insight into the quality of the mental representation produced by the integration units (Kintsch, 1988, 1994, 1998). This feedback can either be internal to the reader (e.g., realizing a portion of text was not well-understood) or provided by an external source (e.g., praise from a teacher). Positive feedback signals the reader that the reading process is proceeding properly and the mental representation constructed is accurate. In contrast, negative feedback will lead to a reader questioning their mental representation and possibly focusing more on the orthographic or phonological components of reading.

Intensive. Research suggests that deliberate practice requires intense concentration for controlled periods of time (Cowley, 1959; Ericsson, 1996a; Ericsson et al., 1993; Plimpton, 1977). LaBerge and Samuels' (1974) notion of internal attention reflects a similar concept. Readers that actively and selectively attend to a reading task are exhibiting intense concentration.

Distributed. Ericsson (1996a) argued that expertise takes an extended period of time to develop (e.g., at least 10 years and 10,000 hours). Connectionist theories of reading are

based on the notion that the connections between cognitive units are only strengthened through repeated exposure over time. It is no more possible to become an expert reader with only a dozen hours practice than to become a chess champion after playing a handful of games.

Self-directed. Novices engage in self-directed practice when they choose to practice when a coach or mentor is unavailable to direct learning (Ericsson, 1996a). Attention, a critical component of the Automatic Information Processing Model (LaBerge & Samuels, 1974), requires a reader to choose to engage in a reading task. During reading, the focus and attention of a reader are self-directed as fluent readers are able to select what to focus on, when to employ repair strategies (e.g., re-reading), and when to take a break.

The present study adopted a connectionist perspective on reading in investigating the influence of deliberate practice on reading ability. Connectionist models of reading are predominant in the contemporary reading literature and are consistent with current perspectives on expertise development, particularly deliberate practice. An important attribute of connectionist models of reading is that targeted experience over an extended period of time is critical to the reading process. Through repeated exposure to appropriate text, readers' construct better-connected neural networks, leading to improved reading. From an expertise perspective, connectionist models of reading are inconsistent with solely time-based models of expertise development. Time spent reading does not necessarily lead to expertise.

From both a deliberate practice and connectionist model of reading perspective, the quality of the practice is critical to acquiring expertise. Unfortunately, providing novice readers with opportunities to engage in deliberate practice, in an effort to enhance their

neural network for reading, poses significant challenges. Deliberate practice requires self-directed activity that is targeted to the individual student along with ongoing feedback. This can be challenging for educators who teach dozens of students. Fortunately, technology can provide the digital platform necessary to provide deliberate practice opportunities as well as the research supporting the acquisition of expertise in reading.

Technology and 21st Century Learning and Research

The growing ubiquity of technology in modern society has signaled the emergence of new educational trends. In 2010, the U.S. Department of Education's (DOE) Office of Educational Technology published the National Education Technology Plan (NETP), *Transforming American Education: Learning Powered by Technology*. The NETP called for the revolutionary transformation of the American education system; the core component of the proposed metamorphosis is technology. According to the NETP (2010), the challenge for the education system is to leverage the learning sciences and modern technology to build engaging, relevant, and personalized learning experiences for all learners. The idea of learning experiences tailored to the specific needs, learning preferences, and interests of a student (i.e., personalized learning) is a radical shift from the education of the masses approach described by Amirault and Branson (2006).

The role of the 21st century teacher is also evolving. In the past, teachers were typically viewed as either subject-matter experts or experts in educational techniques (Amirault & Branson, 2006). The NETP (2010) proposed a connected teaching model whereby solo practitioners are replaced by teams of linked educators operating in classrooms that are fully connected to provide 24/7 (i.e., available every hour of every day) access to data and analytic tools. Still a critical component to transforming education, educators take

on the role of facilitator and collaborator, guiding students during learner-directed instructional tasks.

Traditionally, assessment was designed to indicate whether students had learned after receiving instruction (Amirault & Branson, 2006). The model of learning proposed in the NETP (2010) entails improved methods of assessing learned material and diagnosing student deficiencies. In contrast to traditional post-instruction assessment, the NETP recommended that students be assessed during learning tasks (i.e., when there is still time to improve student performance). Research suggests that this blended model whereby assessment is interwoven and embedded with instruction can improve student learning (Black & Wiliam, 1998; Butler, 2010) and potentially reduce the need for test-taking solely for accountability purposes (NETP, 2010).

Today, learning occurs predominantly in classrooms using paper-based textbooks, while technology use varies based on a multitude of factors (e.g., availability of technology, educator willingness and ability to integrate technology in the classroom). Meanwhile, outside school, students are immersed in technology that provides 24/7 access to information and each other (e.g., social networking) (NETP, 2010). According to the Kaiser Family Foundation (2009), 8- to 18-year olds spend an average of seven hours and 38 minutes using entertainment media (e.g., television, music/audio, Internet, electronic games, movies) in a typical day (i.e., more than 53 hours per week). This extraordinary usage is attributed to an explosion in mobile and online media, as well as user multi-tasking (e.g., watching television while interacting with Internet content). Technology can be used to enhance learning in both formal (i.e., inside school) and informal (i.e., outside school) learning settings (Barron, 2006).

Technology can play a critical role in the future of American education (NETP, 2010); however, instructional technology must be built on established educational research and validated through rigorous scientific methodologies (Shavelson & Towne, 2002). Merely placing technology (e.g., computers and software) into classrooms does not guarantee that students will develop the skills necessary to compete in the 21st century. To foster student learning, technology applications must be learner centered, focus on learning outcomes, and be theory-based and relevant. Each of these requirements is critical to the architecture of educational technology designed to foster reading development through deliberate practice.

Requirement 1: Technology applications must be learner-centered. Compared to traditional technology-free approaches (e.g., paper-bound textbooks), technology can provide a more flexible set of learning resources to learners. The NETP (2010) distinguished between three types of instruction that can be delivered via technology: individualized, differentiated, and personalized. Individualization describes instruction that is paced to the learning needs of each student. Content to be learned (i.e., instructional objectives) is identical for all students, but students can advance through material at their own pace (e.g., spend more time on a difficult topic or skip material already understood). Differentiation represents instruction that is customized to the learning needs of each student. Like individualization, the instructional objectives are identical for each learner; however, the instructional methods vary based on student preferences (e.g., learning independently or collaboratively). Personalization, which encompasses both individualization and differentiation, describes instruction that is paced to learning needs, customized to learning preferences, and tailored to the particular interests of learners. Educational technology can support all three approaches to instruction (NETP, 2010). This emphasis on the learner is a departure from traditional paper-based instruction

where a one-size-fits-all approach is often adopted due to limited time and resources; it is a significant challenge for a single educator to provide each student with a personalized learning experience.

Mayer (2001, 2005) distinguished between a learner-centered approach and a technology-centered approach to the design of educational technology. A technology-centered approach focuses on the cutting-edge characteristics of a technology, and how best to expose learners to a particular technology. This approach requires learners to adapt to the technology, and not surprisingly, has a long history of failing to revolutionize education with a variety of technologies (e.g., motion pictures, radio, television, early computers) (Cuban, 1986). In contrast to a technology-centered approach, a learner-centered approach focuses on how students learn and ways that technology can enable learning (Mayer, 2001, 2005). As opposed to forcing a learner to adapt to technology, the technology should adjust to the learner. Said differently, educational technology should provide a student with a personalized experience specifically tailored to the abilities and interests of the learner. For example, Discovery Education's *Science Techbook* provides a digital learning environment where learners are free to explore science concepts through reading passages at a variety of levels, multimedia simulations, text-to-speech support, and interactive glossaries (Discovery, 2012).

An educational system designed to encourage continuous and lifelong learning (Banks et al., 2006) that is available on-demand (Bransford et al., 2006) offers a compelling vision of learning in the 21st century; however, the development of learner-centered technology demands thoughtful design. The differences between individualized, differentiated, and personalized learning are subtle, yet significant as each approach requires technology to provide different types of learning experiences. Given the evolving nature of

technology-enhanced learning, educational technology designed to be learner-centered must be flexible enough to seamlessly provide all three types of instruction depending on various factors (e.g., topic being taught, educator and student preferences). This carefully-considered design also needs to be balanced against the threat of educational technology shifting towards a technology-centered approach whereby student learning takes a secondary role to cutting-edge technology.

Requirement 2: Technology applications must focus on learning outcomes.

Technology designed to promote learning should focus on clear learning outcomes instead of nebulous goals (Anderson et al., 2001); applications that claim to be instructional without offering a clear description of what students learn (i.e., vague goals) are limited in educational utility (Mayer, 2008). In contrast, technology that offers clearly defined learning outcomes (i.e., what a student learns as a result of instruction) provides a way of monitoring what is learned. Technology-based systems can monitor student progress in real-time, providing on-going estimates of student ability and progress. This progress can be accessed and acted upon immediately by educators, as opposed to waiting for the results of formative or summative assessments. The measurement and tracking of these outcomes enables researchers, teachers, and students to determine the effectiveness of a particular educational technology (Pellegrino, Chudowsky, & Glaser, 2001).

Technology promises to play a critical role in successfully measuring the complex skills necessary for success in the 21st century (NETP, 2010). A variety of assessment types have been proposed (e.g., Scalise & Gifford, 2006) and promising results in assessing required skills have emerged. For example, Vendlinski and Stevens (2002) showed that problem-solving and complex reasoning could be measured using technology by examining

the tools and steps used by students while solving complex science problems in a digital environment. Similarly, Dede (2009) and Dieterle (2009) were able to measure students' science inquiry skills, sense of efficacy in science, and science concept knowledge within a virtual environment.

Technology applications that focus on specific learning outcomes are critical to improving student learning. By tracking student development on specific skills and abilities, the effectiveness of different technology solutions can be compared, and the optimal approach to instruction selected. The key component to this technology-based assessment is that the assessments are seamlessly integrated into the technology. For example, Carnegie Learning's (2012) *Adaptive Math Software Solutions* learning system guides students through critical mathematical concepts and adjusts the difficulty and instruction based on student performance. By allowing students to be monitored during learning, technology enables assessment without requiring separate testing and instructional periods.

Requirement 3: Technology applications must be theory-based and relevant.

Technology designed to promote learning should be grounded in theory and based on educationally relevant research (Shavelson & Towne, 2002). A theoretically-grounded approach offers a clear and testable mechanism for how students learn; researchers are able to develop testable predictions that can be compared to actual student learning data. The antithetical perspective to a theory-grounded method is an ideological approach whereby explanations for student learning are too vague to be tested scientifically (Shavelson & Towne, 2002). A theoretically-grounded approach allows researchers to develop more useful theories of how to improve student learning (Bransford, Brown, & Cocking, 1999).

While instructional technology should be designed based on established theory (i.e., empirically-supported), the relevancy of the scientific evidence is also critical. Mayer (2003) showed that, in the history of educational research, periods exist where research was not relevant to education. Examples include research that focused on arbitrary tasks (e.g., learning lists of nonsense words) or non-authentic environments (e.g., rats navigating a maze). The research used to inform the design of an instructional application must be grounded in theory that is based on relevant research (Shavelson & Towne, 2002). Theory-based educational technologies that are designed, developed, deployed and validated in authentic educational contexts provide evidence that instructional technologies are influencing student learning. This evidence is useful to a diverse set of stake-holders (e.g., students, parents, teachers, administrators, researchers, policy-makers) and should be a requirement before a technology is accepted and deployed at a large scale (e.g., nationwide).

The demand for research-based technology requires a theoretical approach that allows researchers to investigate and provide evidence of the mechanisms enhancing learning. One theoretical approach that can guide the design of educational technology while accommodating the needs of both learners and researchers is deliberate practice. Technology designed according to the principles of deliberate practice (i.e., targeted, self-directed practice that is distributed and intensive and provides real-time corrective feedback) can provide students with appropriate, self-selected content (e.g., activities, material to be learned) and provide instantaneous feedback based on student performance. Moreover, technology is able to track time spent learning to ensure that students are learning over an extended period of time (i.e., distributed), and for limited amounts of concentrated time (i.e., intensive). Technology provides a powerful data collection mechanism whereby extremely

detailed records of student action can be stored (e.g., mouse movement and clicks, key presses, time spent on various sections of a learning task). Known as trace data (Winne, 1982), these fine-grained accounts of student activity captured digitally during learning will provide the information necessary to explore how deliberate practice enhances reading ability.

The architects of both deliberate practice and connectionist models of reading have argued that exposing readers to appropriately challenging text is critical to growing novice readers (Adams, 1990, 1994; Cunningham & Stanovich, 1998; Ericsson, 2006b; Stanovich, 2000). Through repeated exposure to words and sentence structures, the neural network of the reader improves and expands. Educational technology can serve as a platform to deliver the deliberate practice necessary to develop expertise in reading. The following section explores the intersection between deliberate practice, theories of reading, and educational technology.

Deliberate Practice, Theories of Reading, and Educational Technology

Despite a lack of convincing empirical research (NICHD, 2000), reading researchers have staunchly believed that the volume of words read influences the development of reading ability (e.g., Allington, 1980, 1983, 1984a, 2009; Anderson et al., 1988; Cunningham & Stanovich, 1998; Gambrell, 1984; Hiebert, 1983; Knapp, 1995; Krashen, 2004; Meyer & Wardrop, 1994; Stanovich, 2000; Stanovich et al., 1996; Thurlow et al., 1984; Vaughn et al., 1998; Wu & Samuels, 2004). Unfortunately, previous attempts to investigate the relationship between reading volume and reading ability have suffered from methodological limitations in quantifying the amount of reading and estimating changes in student reading ability. To estimate the amount of reading practice completed, past research has employed both

questionnaires (e.g., Ennis, 1965; Nell, 1988; Stanovich & West, 1989; Wagner & Stanovich, 1996; Walberg & Tsai, 1984) and reading diaries (e.g., Allen et al., 1992; Anderson et al., 1988; Greaney, 1980; Greaney & Hegarty, 1987) in an attempt to quantify the amount of reading completed by a study participant. These approaches exhibited poor reliability, suffered from societal pressure, and required commitment and effort on the part of both the reader and the researcher (Cunningham & Stanovich, 1991; Ennis, 1965; Sharon, 1973-1974; Wagner & Stanovich, 1996). Estimating change in reader ability has been similarly limited; traditionally, quantifying change was accomplished using a small number of measurement occasions (e.g., pretest and posttest). While this technique resulted in unbiased estimates of growth, it provided limited information about the nature of change (Rogosa, Brandt, & Zimowski, 1982; Willett, 1988).

Regardless of the limitations of past research, Allington (2009) believed that reading volume certainly influenced reading development; however, he observed a lack of evidence related to the type of practice necessary to foster growth in reading ability. Allington's observation is consistent with the deliberate practice perspective that volume of practice alone is insufficient to achieve expertise in a domain such as reading. Consistent with connectionist models of reading that frame the reading process as a series of interrelated cognitive units, deliberate practice requires activity that is designed to improve performance gradually through specifically designed activity. Practice that is deliberate provides: (1) targeted practice that is designed to appropriately challenge the learner, (2) real-time corrective feedback, (3) distributed practice over a long period of time, (4) intensive practice that does not require the learner to concentrate beyond their limits, and (5) self-directed

practice when a teacher or coach is unavailable (Ericsson, 1996a, 1996b, 2002, 2004, 2006a, 2006b; Ericsson et. al, 1993).

Deliberate practice requires practice that is specifically tailored to the individual learner. The activity must be targeted, provide appropriate ongoing feedback, be available on-demand over time, and must operate independent of a teacher. These requirements suggest that, given limited time and resources, an educator will be challenged to provide deliberate practice for every student. Educational technology that is learner-centered, focuses on learning outcomes, and is grounded in established theory offers a powerful platform to deliver deliberate practice. By combining the principles of deliberate practice and trace data of student reading activity and progress, technology can help to overcome the methodological limitations of past research on the study of reading volume.

The present study attempted to overcome these methodological limitations by using an educational technology platform, *Learning Oasis* (Hanlon, Swartz, Stenner, Burdick, & Burdick, 2012). *Learning Oasis* provides a virtual environment where students engage in deliberate reading practice. Text is automatically matched to the ability of the reader, and throughout the reading experience students receive immediate feedback about performance. The data collected by *Learning Oasis* during reading was used to estimate the amount of deliberate reading practice completed and provide ongoing estimates of reader ability. This real-time monitoring of student ability allows *Learning Oasis* to immerse students with appropriately challenging text over time, regardless of changes in reading ability. The burden of providing students with individualized reading practice that is deliberate is assumed by *Learning Oasis*, and not the classroom teacher. The present study applied a multilevel model (Bryk & Raudenbush, 1987; Goldstein, 1995; Snijders & Bosker, 1999) to examine how

deliberate practice influences change in reading ability. The primary research question was: What is the relationship between the amount of various aspects of deliberate practice and the development of reading ability? The hypotheses for the study were:

H1: The *number of words read from targeted text* will positively relate to the rate of change in reading ability over time

H2: The *number of reading comprehension items answered* will positively relate to the rate of change in reading ability over time

H3: The *number of minutes spent reading intensely per day, measured automatically by the educational technology*, will positively relate to the rate of change in reading ability over time

H4: The *number of days a participant read* will positively relate to the rate of change in reading ability over time

H5: The *standard deviation of elapsed days between reading experiences* will negatively relate to the rate of change in reading ability over time

H6: The rate of change in reading ability over time will be negatively related to the *initial reading ability* of the participant

H7: The rate of change in reading ability over time will be negatively related to the *initial grade of the participant*

H8-a: The *socio-economic status (SES) of a participant* will be positively related to initial reading ability. Participants with lower SES, as measured by receiving free lunch, will have a lower initial intercept than participants with higher SES, as measured by paying for lunch.

H8-b: The *SES of a participant* will be positively related to rate of change in reading ability over time. As predicted by the reading literature, participants with lower SES will have flatter trajectories than participants with higher SES.

CHAPTER 3

METHODS

The present study was designed to investigate the relationship between reading practice and the development of reading ability. Deliberate practice offers a set of principles that describe how practice should be structured to advance novices towards expertise. Consistent with perspectives of expertise development, the authors of connectionist models of reading argued that reading ability develops as a result of increased structure and strength of a reader's neural network. Deliberate practice might be a way to strengthen these connections, and improve student reading ability. An educational technology, *Learning Oasis*, was used to deliver deliberate reading practice and collect the necessary data to answer the primary research question: what is the relationship between deliberate practice and the development of reading ability?

Participants

The participants for this study were selected from the pool of students who used an educational technology tool called *Learning Oasis*. Since initial launch in June 2007, *Learning Oasis* has been used by students ranging from grades two through twelve in a suburban school district in Mississippi. Only students ($N = 1,369$) with at least three measurement occasions each separated by at least three months were included in the study. This minimum data requirement was to ensure that study participants had sufficient data to

effectively examine change over time (Singer & Willett, 2003). Participants in lower grades had more opportunities to use *Learning Oasis* (e.g., more years of schooling before graduating), leading to larger sample sizes in the lower grades (Table 3). The school district also provided a suite of demographic factors for each participant. These values were supplemented with estimates of district, state, and national level demographic information (Table 4).

Table 3. Qualifying participants by grade

Initial Grade	Count
1	29 (2.1%)
2	311 (22.7%)
3	196 (14.3%)
4	142 (10.4%)
5	116 (8.5%)
6	152 (11.1%)
7	95 (6.9%)
8	116 (8.5%)
9	103 (7.5%)
10	92 (6.7%)
11	17 (1.2%)
Total	1,369

Table 4. Demographic information for all participants participating in study ($N = 1,369$)

	N (Sample)	Percentage			
		Sample	District ¹	State ¹	Nation ²
Gender					
Female	700	51%	48%	49%	51%
Male	669	49%	52%	51%	49%
Ethnicity					
African American	541	40%	36%	50%	15%
Asian/Other	8	1%	1%	1%	8%
Caucasian	763	56%	57%	46%	54%
Hispanic	57	4%	6%	3%	23%
Lunch Status					
Receives free lunch	670	49%	52%	62%	35%
Pays for lunch	699	51%	48%	38%	65%

¹ Based on 2010/2011 NCLB Report Card (Mississippi Department of Education, 2013)

² Based on 2010/2011 data from the U.S. Census Bureau (2011), Aud, Fox, and KewelRamani (2010), and U.S. Department of Agriculture (2007)

Data Sources

This secondary data analysis was based on data collected over multiple years as participants used *Learning Oasis*. This research did not control access to the educational technology, so the sample of participants reflects a sample of convenience. Educators and students were given the option to use the educational technology by the district superintendent, and 92% of students in grades one through twelve read at least one article inside *Learning Oasis*. All elementary school teachers (i.e., grades one through five) and all middle and high school English-Language Arts educators used *Learning Oasis* to supplement existing literacy instruction. The sample for the present study reflects participants who read enough to meet the minimum inclusion criteria (i.e., three measurement occasions separated

by three months each); of the 92% of students who read at least one article, 54% met this minimum inclusion criterion.

The author was the architect that designed and engineered *Learning Oasis* and was part of the team that helped to implement the educational technology in classrooms. *Learning Oasis* was developed at MetaMetrics, Inc., where the author was employed while conducting this research. This study was approved by the University of North Carolina at Chapel Hill Institutional Review Board (IRB Number 12-2539).

Procedures

Learning Oasis is an educational technology that immerses students in instructional activities designed to foster the development of expertise in reading, writing, and vocabulary. Each activity in *Learning Oasis* is designed according to the tenants of deliberate practice: (1) targeted practice that is designed to “stretch” the learner, (2) real-time corrective feedback, (3) distributed practice over a long period of time, (4) intensive practice that does not require the learner to concentrate beyond their limits, and (5) self-directed practice when a teacher or coach is unavailable. *Learning Oasis* captures detailed trace data (e.g., time spent reading, item response time, answers selected) that provided fine-grained accounts of participant activity digitally recorded during learning. As a web-based technology, learners were able to use *Learning Oasis* twenty-four hours-per-day, 365 days-per-year as long as they had a computer with access to the Internet.

Educators employed *Learning Oasis* as a supplemental reading program differently in their classrooms, depending on their goals and access to technology. In some locations, particularly elementary classrooms, *Learning Oasis* was used as a literacy center. Every day, participants took turns using the classroom computers or laptops to complete *Learning Oasis*

activities during centers. Other educators assigned weekly quiz grades based on *Learning Oasis* activity. In these classrooms, typically in high school, participants were responsible for managing their own learning by finding time to complete activities. Participants without computers at home often used the school computer labs before and after school, or during lunch or study hall. Yet other educators assigned *Learning Oasis* as nightly homework. In these instances, participants without access to technology received an alternate assignment. A critical aspect of *Learning Oasis* is that participants in the system choose what activities to undertake. Teachers typically did not make specific assignments (e.g., “read this specific article about photosynthesis”). Rather, teachers might have made general assignments related to usage (e.g., “read 20,000 words”). Data describing these differences in implementation were not available for this research.

For purposes of the present study, data collected during participant reading were examined. The data captured by *Learning Oasis* provided the information necessary to both monitor changes in participant reading ability and estimate the amount of deliberate practice completed by a participant. *Learning Oasis* used a relational database to store the record of each reading experience completed by a participant (e.g., the content and complexity of the article read, time spent reading, items attempted during the reading, and time spent answering each item). By measuring text and readers on the same developmental scale, The Lexile Framework for Reading provided a mechanism to match readers to text. Inside *Learning Oasis*, participants were able to search for and choose articles at their individual reading level. The pool of searchable articles included over 25,000 unique informational passages from a variety of periodicals (e.g., *Ranger Rick*, *The Economist*, *Scientific American*, or *Sports Illustrated*). After participants chose an article to read, the reading process began. As

participants read, *Learning Oasis* automatically generated semantic cloze items (Figures 1 and 2) wherein words were systematically removed from the text and the participant was asked to select the missing word from four choices (Taylor, 1953). When a correct word was selected, the box was outlined in green (Figure 3). In contrast, when an incorrect word was selected, the correct word was inserted into the blank and the box was colored red (Figure 4). Performance on these cloze items was used to provide an ongoing estimate of participant reading ability measured in Lexiles (Swartz et al., 2011). The updated estimate of reader ability was used to target future reading experiences.

The screenshot shows the Learning Oasis interface. At the top, there are navigation tabs for Reading, Writing, Vocabulary, and Portfolio. The user's name, Sean Hanlon, and options for Dashboard, About, and Logout are in the top right. The main content area displays the following text:

Section: SCIENCE'S COMPASS
 RESEARCH: QUANTUM COMPUTING

FAST SEARCHES WITH NUCLEAR MAGNETIC RESONANCE COMPUTERS

Quantum computers could revolutionize many branches of science through their ability to tackle problems too large for any classical computer. Although the theory is well understood [see the accompanying by Grover on page 228 (1)], actually building a quantum computer has proved extremely difficult, and until recently it has only been possible to demonstrate the very simplest operations. In the last few years, however, the development of computers based on nuclear magnetic (NMR) spectroscopy has been extremely rapid. Researchers at IBM, Massachusetts Institute of Technology, and the University of California at Berkeley (2) and in my laboratory in Oxford (3) have now demonstrated powerful quantum search algorithms with small NMR computers.

All current designs are built out of the same basic components, quantum bits (qubits) and logic gates. Qubits are the quantum analogs of classical bits, but whereas bits can only take two different values, 0 and 1, qubits are not confined to their two basic states, labeled $|0\rangle$ and $|1\rangle$, but can also exist in states such as $|0\rangle + |1\rangle$, called superpositions. A qubit in this state is not simply in state $|0\rangle$ or $|1\rangle$, nor is it in an intermediate state; rather the qubit is in both states simultaneously. Quantum logic gates act on qubits, just as classical logic gates act on classical bits, but quantum gates also work with superpositions and so can perform multiple logic operations at the same time.

A qubit can be implemented with any two-state quantum mechanical system. In NMR computers, the two spin states of a spin-1/2 atomic nucleus in a magnetic field are used. Different atoms in a molecule can be distinguished, and so a molecule can be used as a quantum computer, with each spin-1/2 nucleus providing a single qubit. Simple logic gates that only affect a single qubit are easily implemented with radio frequency fields. These fields interact strongly with nuclear spins, allowing them to be controlled with great precision. To perform interesting computations, however, more complex gates are needed, which allow the state of one qubit to affect other qubits in the computer. This requires some form of interaction between nuclear spins, so that one spin can sense the state of other spins in the molecule. This is easily achieved, because the naturally occurring spin-spin coupling interaction has the desired form.

The NMR signal from a single molecule is far too weak to be detected, and so it is necessary to use a large number of identical copies to the signal. This is not difficult because even a few milligrams of a chemical compound will contain the required number of molecules. It is, however, impossible to ensure that all the copies start the calculation in the same initial state, and so different copies will in effect perform different calculations, making it extremely difficult to extract the desired result. This inability to prepare NMR computers in a well-defined initial state

Figure 1. Cloze items generated by *Learning Oasis*.

learning oasis Reading Writing Vocabulary Portfolio Sean Hanlon Dashboard | About | Logout

Section: SCIENCE'S COMPASS

RESEARCH: QUANTUM COMPUTING

FAST SEARCHES WITH NUCLEAR MAGNETIC RESONANCE COMPUTERS

Quantum computers could revolutionize many branches of science through their ability to tackle problems too large for any classical computer. Although the theory is well understood [see the accompanying by Grover on page 228 (1)], actually building a quantum computer has proved extremely difficult, and until recently it has only been possible to demonstrate the very simplest operations. In the last few years, however, the development of computers based on nuclear magnetic (NMR) spectroscopy has been extremely rapid. Researchers at IBM, Massachusetts Institute of Technology, and the University of California at Berkeley (2) and in my laboratory in Oxford (3) have now demonstrated powerful quantum search algorithms with small NMR computers.

All current designs are built out of the same basic components, quantum bits (qubits) and logic gates. Qubits are the quantum analogs of classical bits, but whereas bits can only take two different values, 0 and 1, qubits are not confined to their two basic states, labeled $|0\rangle$ and $|1\rangle$, but can also exist in states such as $|0\rangle + |1\rangle$, called superpositions. A qubit in this state is not simply in state $|0\rangle$ or $|1\rangle$, nor is it in an intermediate state; rather the qubit is in both states simultaneously. Quantum logic gates act on qubits, just as classical logic gates act on classical bits, but quantum gates also work with superpositions and so can perform multiple logic operations at the same time.

A qubit can be implemented with any two-state quantum mechanical system. In NMR computers, the two spin states of a spin-1/2 atomic nucleus in a magnetic field are used. Different atoms in a molecule can be distinguished, and so a molecule can be used as a quantum computer, with each spin-1/2 nucleus providing a single qubit. Simple logic gates that only affect a single qubit are easily implemented with radio frequency fields. These fields interact strongly with nuclear spins, allowing them to be controlled with great precision. To perform interesting computations, however, more complex gates are needed, which allow the state of one qubit to affect other qubits in the computer. This requires some form of interaction between nuclear spins, so that one spin can sense the state of other spins in the molecule. This is easily achieved, because the naturally occurring spin-spin coupling interaction has the desired form.

The NMR signal from a single molecule is far too weak to be detected, and so it is necessary to use a large number of identical copies to the signal. This is not difficult because even a few milligrams of a chemical compound will contain the required number of molecules. It is, however, impossible to ensure that all the copies start the calculation in the same initial state, and so different copies will in effect perform different calculations, making it

celebrity kaleidoscope commentary authorship

Figure 2. Example item generated by *Learning Oasis*.

learning oasis Reading Writing Vocabulary Portfolio Sean Hanlon Dashboard | About | Logout

Section: SCIENCE'S COMPASS

RESEARCH: QUANTUM COMPUTING

FAST SEARCHES WITH NUCLEAR MAGNETIC RESONANCE COMPUTERS

Quantum computers could revolutionize many branches of science through their ability to tackle problems too large for any classical computer. Although the theory is well understood [see the accompanying by Grover on page 228 (1)], actually building a quantum computer has proved extremely difficult, and until recently it has only been possible to demonstrate the very simplest operations. In the last few years, however, the development of computers based on nuclear magnetic (NMR) spectroscopy has been extremely rapid. Researchers at IBM, Massachusetts Institute of Technology, and the University of California at Berkeley (2) and in my laboratory in Oxford (3) have now demonstrated powerful quantum search algorithms with small NMR computers.

All current designs are built out of the same basic components, quantum bits (qubits) and logic gates. Qubits are the quantum analogs of classical bits, but whereas bits can only take two different values, 0 and 1, qubits are not confined to their two basic states, labeled $|0\rangle$ and $|1\rangle$, but can also exist in states such as $|0\rangle + |1\rangle$, called superpositions. A qubit in this state is not simply in state $|0\rangle$ or $|1\rangle$, nor is it in an intermediate state; rather the qubit is in both states simultaneously. Quantum logic gates act on qubits, just as classical logic gates act on classical bits, but quantum gates also work with superpositions and so can perform multiple logic operations at the same time.

A qubit can be implemented with any two-state quantum mechanical system. In NMR computers, the two spin states of a spin-1/2 atomic nucleus in a magnetic field are used. Different atoms in a molecule can be distinguished, and so a molecule can be used as a quantum computer, with each spin-1/2 nucleus providing a single qubit. Simple logic gates that only affect a single qubit are easily implemented with radio frequency fields. These fields interact strongly with nuclear spins, allowing them to be controlled with great precision. To perform interesting computations, however, more complex gates are needed, which allow the state of one qubit to affect other qubits in the computer. This requires some form of interaction between nuclear spins, so that one spin can sense the state of other spins in the molecule. This is easily achieved, because the naturally occurring spin-spin coupling interaction has the desired form.

The NMR signal from a single molecule is far too weak to be detected, and so it is necessary to use a large number of identical copies to the signal. This is not difficult because even a few milligrams of a chemical compound will contain the required number of molecules. It is, however, impossible to ensure that all the copies start the calculation in the same initial state and so different copies will in effect perform different calculations making it is the correct answer!

Figure 3. A correctly answered cloze item inside *Learning Oasis*.

learning oasis Reading Writing Vocabulary Portfolio Sean Hanlon Dashboard | About | Logout

Section: SCIENCE'S COMPASS

RESEARCH: QUANTUM COMPUTING

FAST SEARCHES WITH NUCLEAR MAGNETIC RESONANCE COMPUTERS

Quantum computers could revolutionize many branches of science through their ability to tackle problems too large for any classical computer. Although the theory is well understood [see the accompanying commentary by Grover on page 228 (1)], actually building a quantum computer has proved extremely difficult, and until recently it has only been possible to demonstrate the very simplest operations. In the last few years, however, the development of computers based on nuclear magnetic resonance (NMR) spectroscopy has been extremely rapid. Researchers at IBM, Massachusetts Institute of Technology, and the University of California at Berkeley (2) and in my laboratory in Oxford (3) have now demonstrated powerful quantum search algorithms with small NMR computers.

All current designs are built out of the same basic components, quantum bits (qubits) and quantum logic gates. Qubits are the quantum analogs of classical bits, but whereas bits can only take two different values, 0 and 1, qubits are not confined to their two basic states, labeled $|0\rangle$ and $|1\rangle$, but can also exist in states such as $|0\rangle + |1\rangle$, called superpositions. A qubit in this state is not simply in state $|0\rangle$ or $|1\rangle$, nor is it in an intermediate state; rather the qubit is in both states simultaneously. Quantum logic gates act on qubits, just as classical logic gates act on classical bits, but quantum gates also work with superpositions and so can perform multiple logic operations at the same time.

A qubit can be implemented with any two-state quantum mechanical system. In NMR computers, the two spin states of a spin-1/2 atomic nucleus in a magnetic field are used. Different atoms in a molecule can be distinguished, and so a molecule can be used as a quantum computer, with each spin-1/2 nucleus providing a single qubit. Simple logic gates that only affect a single qubit are easily implemented with radio frequency fields. These fields interact strongly with nuclear spins, allowing them to be controlled with great precision. To perform interesting computations, however, more complex gates are needed, which allow the state of one qubit to affect other qubits in the computer. This requires some form of interaction between nuclear spins, so that one spin can sense the state of other spins in the molecule. This is easily achieved, because the naturally occurring spin-spin coupling interaction has the desired form.

The NMR signal from a single molecule is far too weak to be detected, and so it is necessary to use a large number of identical copies to 4 the signal. This is not difficult because even a few milligrams of a chemical compound will contain the required number of molecules. It is, however, impossible to ensure that all the copies start the calculation in the same initial state and so different copies will in effect perform different calculations making it **observance** is the wrong choice!

Figure 4. An incorrectly answered cloze item inside *Learning Oasis*.

Instrumentation

The present study required two quantities that were measured using two different instruments: estimates of participant reading ability, and amount of deliberate practice completed by a participant. First, The Lexile Framework for Reading provided a metric that could represent readers and text on a common developmental scale. This conceptual framework was used to measure change in participant reading ability. Second, *Learning Oasis* provided the digital learning environment and data collection mechanism necessary to determine the amount of deliberate practice completed by a learner.

The Lexile Framework for Reading. The Lexile Framework for Reading is a scientific approach to measuring both reading ability and text complexity (Nelson, Perfetti, Liben, & Liben, 2011; Stenner, 2003; Stenner & Stone, 2004; Williamson, 2008). According to The Lexile Framework for Reading, reading ability is a latent trait that influences a reader's chance of success in comprehending professionally edited text (Williamson, 2008) and text complexity is a quantitative factor that represents the difficulty of a text (Common Core State Standards, 2010). A Lexile measure is a numeric value with a trailing "L" (e.g., 750L, 1000L, 1300L) and can be used to represent both a person's reading ability and an estimate of a text's complexity. These values can range from below 0L (i.e., Beginning Reader) to upwards of 2000L (Stenner, Burdick, Sanford, & Burdick, 2006, 2007). Research suggests that students graduating high school must be able to read 1300L text to be considered college and career ready (Common Core State Standards, 2010; Stenner, Sanford-Moore, & Williamson, 2012; Williamson, 2008). A 2001 National Center for Education Statistics report concluded that the Lexile Framework has solid psychometric properties and has been validated across a range of populations (White & Clement, 2001).

The Lexile Framework for Reading relies on two separate, but parallel technologies. First, the Lexile Reading Analyzer is a piece of software that analyzes text and produces an estimate of text complexity. Second, reading tests that measure student reading ability on the Lexile scale provide a mechanism for assessing the reading ability of a student. The output from each of these technologies is a Lexile measure that can be used to match readers to text (e.g., a 1000L reader should be matched to text at 1000L) (Stenner et al., 2007).

Estimating text complexity. Carver (1974) argued that all symbol systems contain both a semantic and syntactic component. For written prose, words are organized according

to specific rules of syntax that are structured into fragments and sentences. Both the syntactic and semantic components of text can vary, influencing its complexity (i.e., readability or comprehensibility). The Lexile Framework for Reading accounts for both of these factors when estimating text complexity (Stenner et al., 2007). The Lexile Framework for Reading is highly correlated with the estimates of text complexity produced by other text complexity formulas (e.g., Advantage/TASA Open Standard, Degrees of Reading Power, SourceRater, Pearson Reading Maturity Metric) that examine a variety of text features (e.g., word length, word frequency, word grade level, word difficulty, word meaning, sentence length, within-sentence punctuation, sentence and paragraph cohesion, clause count, paragraph length, book length) (Nelson et al., 2011; Wright & Stenner, 1998; Wright & Stone, 2004).

Semantic component. When investigating semantic complexity, researchers often employ proxies for the probability that a reader will encounter a word in a familiar context and be able to infer its meaning (Bormuth, 1966). Klare (1963) theorized that the semantic component varies along a continuum ranging from familiar to rare. Carroll, Davies, and Richman (1971) utilized a five-million word corpus and determined that more frequently occurring words had a higher likelihood of being known by a reader. The Lexile Framework for Reading employs a 600-million-word corpus when computing a proxy for semantic complexity (Stenner et al., 2007).

Syntactic component. Klare (1963) argued that the syntactic component of text complexity varied with the demand placed on short-term memory. Sentence length has been shown to be an effective proxy for the syntactic complexity of a piece of text (Crain & Shankweiler, 1988; Liberman, Mann, Shankweiler, & Westelman, 1982; Shankweiler &

Crain, 1986). The Lexile Framework for Reading operationalizes syntactic complexity using a transform of sentence length (Stenner et al., 2007).

Estimating reader ability. A variety of instruments report Lexile reader measures. Each instrument has been linked to the Lexile scale, providing students, parents, and educators with estimates of reader ability denoted in Lexiles. Example instruments include high-stakes end-of-year state assessments (e.g., California, Florida, Georgia, North Carolina, South Carolina, Virginia) as well as norm-referenced assessments, formative assessments, and reading interventions published by a variety of companies (e.g., CTB/McGraw-Hill, Measured Progress, NWEA, Pearson, Prentice Hall, Riverside Publishing, Scantron, Scholastic). Each of these tests or programs produces an estimate of a student’s reading ability denominated in the Lexile metric. *Learning Oasis*, developed by the creators of the Lexile Framework for Reading, is another instructional program that estimates student reading ability. *Learning Oasis* collects the data necessary to calculate and monitor the participant measures used in the study.

Learning Oasis. The cloze items routinely generated by *Learning Oasis* provide useful research data about individual reading experiences (e.g., reading fluency, response latency, bad-faith checks). Given that the number of items is often a function of the length of the reading passage (i.e., shorter passages often have fewer items), it is useful to aggregate reading items into “envelopes” of sufficient length. *Learning Oasis* applies an envelope of 56 items because that is consistent with the mean test length of existing high-stakes reading assessments (Table 5). An envelope does not exceed ninety-days between the first item in an envelope and the last. As students read at their own pace, some will create a 56 item envelope from only a few days-worth of reading. Less active users will take longer in

accumulating 56-items. This enveloping process is important to ensure that estimates of reading ability are not influenced by the number of items in a particular article. Each envelope provides an estimate of a student’s ability, the date of the estimation, and a record of the encounters (i.e., student-text interactions) included in the envelope.

Table 5. Test length of a sample of high-stakes reading assessments.

Test	State	Items	Source
ACT	Any	40	ACT, 2012
Graduate Record Exam (GRE)	Any	50	ETS, 2012
Scholastic Aptitude Test (SAT)	Any	67	Princeton Review, 2012
TerraNova, Third Edition	Any	39-47	CTB, 2012
High School Graduation Qualifying Exam	Alaska	56	Alaska Department of Education and Early Development, 2012
Arizona’s Instrument to Measure Standards (AIMS)	Arizona	70-114	Arizona Department of Education, 2012
California Standards Test (CST)	California	43	California Department of Education, 2012
Transitional Colorado Assessment Program (TCAP)	Colorado	56	Colorado Department of Education, 2012
Connecticut Mastery Test (CMT)	Connecticut	52-57	Connecticut State Department of Education, 2012
Delaware Student Testing Program (DSTP)	Delaware	55	Delaware Department of Education, 2012

Florida Comprehensive Assessment Test (FCAT)	Florida	50	Florida Department of Education, 2012
Criterion-Referenced Competency Tests (CRCT)	Georgia	50	Georgia Department of Education, 2012
Idaho Standards Achievement Test (ISAT)	Idaho	40-45	Idaho State Department of Education, 2012
Iowa Assessments	Iowa	34-45	Iowa Department of Education, 2012
Illinois Standards Achievement Test (ISAT)	Illinois	70	Illinois State Board of Education, 2012
Kansas Reading General Education Test	Kansas	58-84	Kansas State Department of Education, 2012
New England Common Assessment Program (NECAP)	Maine, New Hampshire, Rhode Island, Vermont	52	Maine Department of Education, 2012
Massachusetts Comprehensive Assessment System (MCAS)	Massachusetts	46	Massachusetts Department of Elementary and Secondary Education, 2012
End-of-Grade Reading Comprehension Test	North Carolina	58-62	North Carolina Department of Public Instruction, 2012
Pennsylvania System of School Assessment (PSSA)	Pennsylvania	58	Pennsylvania Department of Education, 2012
Palmetto Assessment of State Standards (PASS)	South Carolina	36-50	South Carolina State Department of Education, 2012
Texas Assessment of Knowledge and Skills (TAKS)	Texas	36-48	Texas Education Agency, 2012
Virginia Standards of Learning (SOL)	Virginia	35-45	Virginia Department of Education, 2012
Mean		36-56	

Estimating the amount of deliberate practice. The data collected by *Learning Oasis* during student reading (e.g., items completed, words read, time spent reading) provided insight into the quality of practice. The present study required a systematic way to identify the deliberate practice completed by a reader (Table 6). Although there is no accepted procedure for quantifying deliberate practice in reading, this study applied concepts from the existing literature to quantify each component of deliberate practice.

Table 6. Operational definition of deliberate practice

Component	<i>Learning Oasis</i> Proxy
Amount of conditional reading volume	Number of words read inside <i>Learning Oasis</i>
Amount of real-time feedback	Number of auto-generated cloze items answered
Amount of intensive practice	Number of intense minutes of reading per day
Distributed practice	Number of days a participant read using <i>Learning Oasis</i> ; Standard deviation of the elapsed days between reading experiences
Self-directed	All reading is self-directed (not used in study)

Targeted practice. The match between reader and text can mean the difference between a skilled reader who is confident and a declining reader who is frustrated (Mesmer, 2008). The International Reading Association (2004) emphasized the importance of students reading appropriately challenging text. Unfortunately, many learners struggle to select appropriately challenging text for themselves (Donovan, Smolkin, & Lomax, 2000; Fresch, 1995). The Lexile Framework for Reading offers a powerful mechanism to match readers to

text by representing readers and text on the same developmental scale (Stenner et al., 2007). *Learning Oasis* provided automatic matching of readers to text.

The Lexile Framework was used by researchers investigating the summer fade phenomenon (i.e., decline in student reading ability over the summer) (e.g., Kim, 2006, 2007; Kim & White, 2008; Kim & Guryan, 2010). Kim (2006, 2007) sent books targeted to the Lexile of each student and asked readers to self-report completing the book. Summer fade was measured using pretest/posttest results. Kim (2007) showed a statistically significant difference between treatment (i.e., received targeted books) and control (i.e., did not receive books) groups. These results suggested that reading targeted books over the summer reduced the negative impact of summer fade.

Unfortunately, the level of match between reader and text that leads to optimal learning is unknown (Shanahan, 1983, 2011). Kim (2006, 2007) sent books to readers that were $\pm 50L$ around the reading ability of the student. This 100L span results in an expected comprehension rate between 71% and 78% (Stenner, 2003). This range is consistent with other perspectives on expected comprehension during reading. Research suggests that 75-100% comprehension represents instructional level reading while reading below 50% comprehension leads to reader frustration (Betts, 1946; Ekwall, 1976; Johnson & Kress, 1965; Harvey & Goudvis, 2007). Carver (2000) argued that 64% or better is the ideal comprehension rate for reading.

Learning Oasis allows readers to choose articles that are $\pm 100L$ around their reading ability. According to the Lexile Framework for Reading, this 200L range leads to an expected comprehension rate between 66% and 82% (Stenner, 2003). This expected comprehension serves as a guideline to ensure that readers are not presented with

extraordinarily easy or insurmountably difficult text. Still, it is important to realize that a theoretical match between reader and text alone is insufficient to predict comprehension. Although theoretical methods of making informed targeting decisions are useful, the ideal match between reader and text ultimately depends on the specific reader and the particular text (Schirmer & Lockman, 2001). A reader can be perfectly matched to a text and perform poorly or exceed expectations based on a variety of factors (e.g., interest, motivation, background knowledge) (Mesmer, 2008).

For the present study, number of words read inside *Learning Oasis* represented the amount of targeted practice. While *Learning Oasis* enforced this targeting automatically, the requirement was relaxed at the very low (less than 600L) and very high end (greater than 1300L) of the Lexile scale due to limited availability of targeted text for students to read. Targeted words read were used to quantify the amount of targeted practice. Using targeted words read as an approximation of targeting is consistent with connectionist theorists of reading whose creators argued that exposure to appropriately challenging words strengthens areas of the neural network that are less-well developed. In contrast, students reading words that are not challenging and too easy will not further strengthen the neural network as the connections between mastered words will already be sufficiently strong, and connections between not-yet-mastered words will remain weak.

Real-time corrective feedback. The importance of feedback during learning has been an important component to theories of learning throughout history. It was the critical component of Thorndike's (1932) Law of Effect, and Skinner (1969) positioned feedback as a type of reinforcement that informs a learner about the accurateness of performance. Contemporary researchers have argued that feedback can motivate learning, provide

information necessary to improve learning, and lead to altered approaches to learning (Borich & Tombari, 1997; Eggen & Kauchak, 2004; McClenaghan & Ward, 1987; Zahorik, 1987).

Research has shown that feedback provided immediately along with information about the accuracy of the response improved learning and retention (Epstein, Epstein, & Brosvic, 2001; Epstein et al., 2002; Epstein & Brosvic, 2002).

Learning Oasis presents readers with automatically-generated cloze items chosen based on the difficulty of the text. A cloze item is created when a word is removed from the passage, and the reader is presented with four choices to complete the sentence (Figure 2). These items are scored in real-time immediately following item completion (Figures 3 and 4). Research suggests that cloze items require readers to manipulate linguistic elements at both the sentence level (e.g., Alderson, 1979; Markham, 1985; Porter, 1983) and between sentences (e.g., Bachman, 1985; Brown, 1983; Chavez-Oller, Chihara, Weaver, & Oller, 1985; Chihara, Oller, Weaver, & Chavez-Oller, 1977; Jonz, 1987). During reading, a reader receives immediate feedback based on the accuracy of a response. While Bormuth (1967) argued that removing every fifth word was ideal, *Learning Oasis* chooses random words of similar difficulty to the text. This modification to the cloze procedure ensures that connective words such as *the* and *then* are not chosen as items. For the present study, the number of items taken by each individual reader represented immediate feedback as each time an answer was chosen, both textual and a color-based feedback was provided about the reader's performance. This feedback provided an indication of performance without requiring additional time that could detract from a reading experience. A participant could not read inside *Learning Oasis* without answering items.

Intensive. Reading fluency, described by the National Reading Panel (NICHD, 2000) as a critical component of reading, provides a useful framework to interpret reading intensity. Fluency is reading text with speed, accuracy, expression, and comprehension (Johns & Berglund, 2006; NICHD, 2000). Reading rate is traditionally denoted in words read per minute (wpm) and refers to how quickly and accurately a reader can move through a text. Carver (1976) proposed clocking reading fluency using standard-length words per minutes. Abbreviated Wpm (i.e., a capital “W”), standard-length words per minute is a measure of how many standard words (i.e., 6 character-spaces) a reader reads per minute. Standard words are not actual words denoted by spaces between them; rather, a standard-length word count for a text is computed by examining all the non-space characters and dividing by six to normalize the word count. This distinction is important as easier texts have shorter words while more difficult texts have longer words. According to Carver (1983, 1990), when measured in standard-length words per minute, reading rate is constant regardless of text difficulty. The present study used reading rate, measured using Wpm, to quantify the intensity of reading practice.

Carver (1972, 1992b, 2000) outlined five distinct gears that a reader can shift between during reading. Lower numbered gears require more intensity and cognitive attention than higher level gears: memorizing (Gear 1), learning (Gear 2), rauding (Gear 3), skimming (Gear 4), and scanning (Gear 5). Memorizing occurs when a reader wants to commit text to memory. During learning, a reader extracts and remembers particular ideas from text. Rauding, the middle gear, involves a reader internally articulating words and sentences in an effort to comprehend. Skimming and scanning occur when a reader wants to get an overview of a text or search for a specific word or phrase. Carver (2000) argued that readers can shift

gears for a variety of reasons: the instructions request faster or slower reading, the task requires more or less than simple comprehension, the reading material is too difficult or too easy, or the time available is insufficient to complete the task via rauding.

While several sets of reading rate norms exist for oral reading (e.g., Good & Kaminski, 2003; Hasbrouck & Tindal, 2006), silent reading norms are limited (Hiebert, Samuels, & Rasinski, 2012). Based on a national sample of over 1,000 students collected by Taylor (1965), Carver (1989, 1990, 1994, 2000) estimated grade-level mean Wpm for the rauding process (Table 7): the typical college student reads between 251 and 277 Wpm (Carver, 1983, 1992b; Rayner, 1975; Zuber & Wetzel, 1981). These reading rates, measured in Wpm, were used to estimate the reading gear used by college students (Carver, 1990). While these data are informative, it is important to note the limitations of these national norms for the present study. First, Carver only provided grade-based silent reading rate norms for the 50th percentile. The range of typical reading rates within a grade is unknown (Hiebert, Samuels, & Rasinski, 2012). Second, Taylor's data were collected nearly 50 years ago and might not be applicable to students reading targeted text in a digital learning environment. This *conditional silent fluency*, or how quickly students silently read targeted text, is not well defined nor well understood.

Table 7. Grade-based standard-length words per minute (Wpm) ranges.

Grade	Mean Rauding Rate (Wpm) ¹	Deliberate Practice Range (Wpm)
1	112	12-212
2	125	25-225
3	137	37-237
4	150	50-250
5	162	62-262
6	175	75-275
7	188	88-288
8	201	101-301
9	213	113-313
10	227	127-327
11	238	138-338
12	251	151-351

¹ Reading rates based on rates provided in Carver (1989, 2000)

As deliberate practice requires a learner to concentrate on a learning task (Ericsson et al., 1993), the present study used reading rate to estimate the intensity of reading practice. Framed using Carver’s reading gears, a reader was considered engaged in deliberate practice when reading in either the learning or rauding gears. Using the grade-level rates provided by Carver (1989) and estimates of the reading rate during the learning gear (Carver, 1992b, 2000), the reading rate during deliberate practice was estimated (Table 7). The range of expected reading rates is based on Carver’s (1992b) observation that great variability exists between individuals at the same grade and suggested quantification of this inconsistency. The Wpm for each reading experience was computed at the conclusion of each reading

experience, and only experiences that were within the appropriate Wpm bands, as identified by Carver (1992b, 2000), were considered intense practice. Intensity was quantified by the number of minutes of intense reading completed per day.

Distributed. To achieve expertise in a domain requires a novice to engage in at least ten years and 10,000 hours of deliberate practice (Bryan & Harter, 1899; Hayes, 1981; Ericsson, 1996b, 2004, 2006b; Ericsson et al., 1993; Simon & Chase, 1973). This time-based distribution of practice is consistent with the literature on spacing effects: exposures to concepts with repetitions separated by time are better recalled than concepts with repetitions that are massed and occur in immediate succession (Carpenter, Cepeda, Rohrer, Kang, & Pashler, 2012; Dempster, 1988; Toppino & Schneider, 1999). The spacing of practice enables the development of expert knowledge structures; this development is a time consuming process that is cognitively demanding (Chi, Bassok, Lewis, Reinmann, & Glaser, 1989; Lesgold et al., 1988; Schooler, 1990). Practice sessions distributed temporally ease the burden placed on a learner developing expertise. For the present study, the distributed nature of the practice was represented using two estimates: 1) the number of days reading inside *Learning Oasis*, and 2) the standard deviation of the time elapsed between days spent reading. Using these metrics, participants with more days read engaged in more distributed practice than participants who read on fewer days. In contrast, participants with smaller standard deviations of time elapsed between readings were engaged in more distributed practice than participants with larger standard deviations.

Self-Directed. To acquire expertise in a domain requires aspiring experts to work independently while managing and guiding their learning (Ericsson et al., 1993). In the domain of reading, the amount of independent, silent reading completed both in and out of

school was found to be significantly related to increases in reading achievement (Allington, 1984b; Anderson, Hiebert, Scott, & Wilkinson, 1989; Anderson et al., 1988). While reading a variety of text (e.g., newspapers, magazine articles, books) is important to the development of reading ability, the majority of readers spend approximately 1% of their free time reading (Anderson et al., 1988).

Learning Oasis is an entirely learner-driven personalized learning platform: students choose the activities to complete and the content to learn. This includes reading articles of high-interest to the student as well as flexibility in the time of day to learn. *Learning Oasis* is available on-demand and students are able to access the learning system from any computer connected to the Internet. Typical usage includes use during school hours (e.g., as a literacy center or using a computer lab or laptop cart), as well as access outside school hours (e.g., from home, library, YMCA). By design, *Learning Oasis* is entirely self-directed. For this reason, no operational contrast exists between self-directed and directed learning. While self-directed activity did not appear in the model, all reading experiences included in this study were self-directed, and were therefore deliberate. Future research should explore the influence of varying the level of self-direction.

Data Analysis Plan

The primary research question for the present study was: What is the relationship between the amount of various aspects of deliberate practice and the development of reading ability? From a deliberate practice perspective, expertise takes an extended period of time to achieve. Therefore, the study examined how readers changed over time.

Change has been a topic of interest to empirical researchers for decades; however, the research community's perspective has evolved over time. In the 1960s and 1970s, most

researchers insisted that change could not be measured effectively and was better ignored (e.g., Bereiter, 1963; Cronbach & Furby, 1970; Linn & Slinde, 1977). Starting in the 1980s, techniques emerged that allowed the effective investigation of individual change (Graham, Singer, & Willett, 2009; Singer & Willett, 2003). These methods grew out of different disciplines, and go by a variety of names: individual growth modeling (Rogosa, Brandt, & Zimowski, 1982; Willett, 1988), multilevel modeling (Goldstein, 1995), hierarchical linear modeling (Raudenbush & Bryk, 2002), random coefficient regression (Hedeker, Gibbons, & Flay, 1994), and mixed modeling (Pinheiro & Bates, 1995). A critical requirement for measuring individual change is longitudinal data (i.e., data collected on the same individuals over time). Cross-sectional data (i.e., data collected without matching people across collection occasions), while easier to collect and more widely accessible, are insufficient for growth modeling as different people are measured at each measurement occasion (Rogosa et al., 1982; Singer & Willett, 2003; Willett, 1988).

The multilevel model for change provides a statistical model for the analysis of individual change over time using longitudinal data (Graham et al., 2009; Singer & Willett, 2003). A multilevel model for change includes two levels: within-person and between-person. The level-1 submodel describes how individuals change over time (i.e., within-person, intraindividual). The purpose of the level-1 submodel is to describe the shape of each person's individual growth trajectory. The level-2 submodel describes how the parameters of individual growth trajectories differ across individuals (i.e., between-person, interindividual). An additional goal of the level-2 submodel is to determine the relationship between predictors (e.g., grade, SES, amount of practice completed) and the parameters of each individual growth trajectory. For the present study, participant change in reading ability over

time was investigated in the context of the amount of deliberate practice completed by a participant. Taken together, these two submodels form a multilevel statistical model that identifies the influence of deliberate practice on growth in reading ability (Bryk & Raudenbush, 1987; Rogosa & Willett, 1985).

The remainder of this section describes the data analysis plan for the study. The first section discusses the data requirements to apply the multilevel model for change. The second section presents the analysis for the study.

Data requirements. Longitudinal data alone are not enough to allow researchers to investigate change over time. To properly apply a multilevel model for change, at least three requirements must be met: (1) multiple waves of data, (2) a reasonable time metric, and (3) an outcome that changes systematically over time (Singer & Willett, 2003).

Multiple waves of data. To effectively model change, researchers must have longitudinal data that describe how each person changes over time. This requirement bears repeating as some empirical researchers have at times made the leap from cross-sectional differences between people (e.g., differences between varying sets of children at multiple ages at a single occasion in time) to generalizations regarding change over time. In such instances, while change might be the actual reason for observed differences, researchers may not be able to definitively identify the effect using cross-sectional data. Studies based on cross-sectional data confound age and cohort effects (Hedeker & Gibbons, 2006; Singer & Willett, 2003).

The number of waves of data is also an important factor. Several researchers (e.g., Bock, 1983; Bryk & Raudenbush, 1987; Rogosa et al., 1982; Rogosa & Willett, 1985; Ware, 1985; Willett, 1988, 1989) have argued for using multiple waves of data collection to

represent change. Two waves of data collection are insufficient as they cannot describe the shape of individual trajectories and researchers are unable to distinguish between true change and measurement error (Rogosa et al., 1982). While researchers can examine individual change using only three waves of data, they are limited to an assumption of linear change. It is not possible to examine curvilinear change without additional measurement occasions. Additional waves provide increased flexibility in selecting models (i.e., linear, quadratic, negative exponential) to represent change (Singer & Willett, 2003).

Multiple waves of data are important, but the collection schedule can vary across individuals (Graham et al., 2009). Data can be collected on a fixed schedule, where each participant is measured the same number of times (e.g., student writing data collected in the spring of grades 4, 7, and 10). The multilevel model for change can also function with data collected according to a flexible schedule (i.e., each individual has a unique schedule for data collection). This flexibility is possible because the multilevel model can accommodate missing or imbalanced data at either level more effectively than traditional repeated measures or random-effects ANOVA approaches (Singer & Willett, 2003). For the present study, the estimates of reading ability based on the ongoing participant activity were used to fit a multilevel model. Each participant had different numbers of measurement occasions measured at distinct points in time. The number of measurement occasions varied from three (the minimum) to 119 ($M = 22.31$). Participants who utilized *Learning Oasis* for a longer period of time had more measurement occasions.

A reasonable time metric. Time is the level-1 predictor in the study of change; therefore, it must be measured reliably in a reasonable metric. A multilevel model for change supports a range of time metrics (e.g., seconds, days, weeks, months, years, sessions,

semesters) and the choice of a metric for time needs to reflect the expected outcome (Graham et al., 2009; Singer & Willett, 2003). In most studies, there are a variety of reasonable time metrics. For example state departments interested in reading often choose to measure time yearly (i.e., North Carolina's current approach) or monthly (e.g., if a formative assessment system is employed that can provide monthly estimates of reading ability).

The time metric informs the data collection schedule; the goal is to collect enough data to create reasonable estimates of each individual's growth trajectory (Singer & Willett, 2003). In a time-structured collection schedule, each individual is assessed on an identical schedule regardless of the temporal spacing (i.e., data collection waves can be equally or unequally spaced in time). A time-unstructured data collection schedule occurs when data collection schedules vary across individual (i.e., each person has a unique schedule). The present study employed a time-unstructured data collection schedule. Time was measured in days elapsed since the participant first started using *Learning Oasis*. This centering facilitated interpretation as the time metric represented how long each learner used *Learning Oasis*. To ensure that each participant had sufficient data to make reasonable estimates of change over time, to be included in the study, a participant needed at least three measurement occasions separated by at least three months.

An outcome that changes systematically over time. Multilevel models for change require instruments that produce outcomes that systematically change over time (Singer & Willett, 2003). Examples of instruments that fulfill these requirements include the Child Assessment Schedule (Hodges, Kline, Stern, Cytryn, & McNew, 1982), the Woodcock Johnson Psychoeducational Test Battery (Bracken & McCallum, 1993), the SCL-90-R (Derogatis & Savitz, 2000), and the Lexile Framework for Reading (Stenner, Burdick,

Sanford, & Burdick, 2007). The measurement properties of a particular instrument are also important. Measurement invariance describes stability in the psychometric properties of a metric across populations and occasions (Mellenbergh, 1989; Meredith & Millsap, 1992; Millsap, 2010). This is critical to longitudinal growth modeling as changes in measures over time should reflect real change in the outcome variable of interest and not be artifacts of the individual, instrument or measurement occasion (Millsap, 2010). Measures must sustain consistent interpretations over time (e.g., identical scores obtained on multiple occasions should imply equal magnitude of the outcome). The psychometric properties (e.g., measurement invariance) of The Lexile Framework for Reading have been established (White & Clement, 2001), and they are sufficient for the measurement of change.

Analysis plan. Analysis of longitudinal growth data using a multilevel model for change was performed in three general steps: (1) creation of a longitudinal data set and exploratory analysis, (2) construction of a multilevel model for change, (3) fitting the model to data and interpreting the results. The following sections describe the steps that were taken to apply the multilevel model for change to participant reading data collected by *Learning Oasis*. The analysis plan followed these steps.

Creation of a longitudinal data set and exploratory analysis. The first step to examining change over time was to organize the longitudinal data into a format that facilitates analysis. Singer and Willett (2003) recommended a person-period format for multilevel analysis. In a person-period format, each row in the dataset represents a single measurement occasion for an individual; therefore, students with multiple measures over time have multiple rows. This format is preferred as it allows the seamless addition of newly acquired data without major modification to the format of the dataset.

The data set for the present study consisted of a variety of values, ranging from person identifiers to quantities representing deliberate practice (Table 8). Each of the quantities representing deliberate practice was measured by *Learning Oasis*. After the dataset was collected and organized, exploratory analyses were conducted to identify important features of the data and to prepare for the application of a statistical model. This exploration was performed in two ways: (1) exploring individual change over time and (2) exploring differences in change across people.

Exploring individual change over time. Insight into how each person changes over time is acquired by examining both graphical and statistical evidence. A random selection of participants, stratified by a variety of predictors was used for exploratory purposes. Empirical growth plots (i.e., a scatter plot with time on the x-axis and participant reading measure on the y-axis) provided information about how a participant's estimated reading ability changed over time (e.g., increasing, decreasing, or constant) (Singer & Willett, 2003).

After exploring participant growth data graphically, an ordinary least squares (OLS) regression line was employed to better understand where participants began (i.e., intercept), how fast they changed (i.e., slope), and how well the data fit both a linear and curvilinear model (i.e., R^2 statistic). To facilitate interpretation, time was measured in days elapsed since a participant entered *Learning Oasis*. This anchoring aided interpretation as the time variable represented how long a participant had used the learning system (Graham et al., 2009; Singer & Willet, 2003).

Table 8. Values in person-period data set.

Variable Name	Description	Component of Deliberate Practice
user_id	Unique identifier of a study participant	
Lexile	Estimated reading ability (measured using <i>Learning Oasis</i>)	
Ctime	Centered time (days elapsed since entering <i>Learning Oasis</i>)	
words	Total targeted words read	Targeted practice
items	Total items	Immediate feedback
minutes_per_day	Number of minutes per day spent reading intensely	Intensity of practice
days_used	Number of days <i>Learning Oasis</i> was used	Distributed practice
elapsed_daysSD	Standard deviation of elapsed days	Distributed practice
initial grade	Initial grade of the participant when they started <i>Learning Oasis</i>	
initial Lexile	Initial Lexile of the participant when they started <i>Learning Oasis</i>	

Exploring differences in change across people. After investigating how individuals changed over time, similarities and differences in change across people (i.e., interindividual differences in change) were explored. By plotting the OLS regression lines for a random set of prototypical individuals simultaneously, along with the average change trajectory, the contrast between individual trajectories and group change was examined. In adopting a parametric model (e.g., linear) for individual change, the specific parameters (e.g., intercept and slope) were used to express interindividual differences in change. Descriptive analyses of the estimated intercepts and slopes provided: estimates of initial status and rate of change, observed variability in initial status and rate of change (i.e., standard deviations of intercepts and slopes), and the relationship between initial status and rate of change (i.e., correlation between estimated intercepts and slopes) (Singer & Willett, 2003).

Constructing a multilevel model for change. At the core of a multilevel model for change are two submodels. The level-1 submodel describes how individuals change over time (i.e., trajectory of estimated reading ability), and the level-2 submodel describes how these trajectories vary across individuals (Graham et al., 2009; Singer & Willett, 2003). This section describes these two levels at a conceptual level.

Level-1: Individual change. Also known as the individual growth model, this model represents the expected growth in reading ability for each participant over time. The level-1 submodel is:

$$Y_{ij} = [\pi_{0i} + \pi_{1i}(Time_{ij})] + \varepsilon_{ij} \quad (1)$$

This model assumes a straight line represents each person's true change over time and random measurement error (ε) causes deviation from the true status on each occasion.

Subscripts i and j represent persons and measurement occasions respectively. The first part of

the equation, shown in brackets, represents the hypothesis that change in reading ability is a linear function of the time predictor. The individual growth parameters (i.e., π_{0i} and π_{1i}) characterize the trajectory (i.e., intercept and slope) of the i -th participant. While all participant change trajectories are assumed to have the same algebraic form (i.e., linear), each participant can have a unique trajectory (i.e., intercept and slope are free to vary by person) (Graham et al., 2009; Singer & Willett, 2003).

Level-2: Systematic interindividual differences in change. The level-2 submodel provides insight into the relationship between interindividual differences in change trajectories and time-invariant aspects of the individual (Graham et al., 2009; Singer & Willett, 2003). There are four specific features of this model: (1) outcomes must be individual growth parameters from the level-1 model, (2) each level-1 growth parameter must appear in a separate equation of the level-2 submodel, (3) each equation must specify a relationship between an individual growth parameter and the predictor variable, and (4) the model must allow individuals who share common predictor variables to vary in individual change trajectories (Singer & Willett, 2003). The level-2 submodel can be represented as:

$$\pi_{0i} = \gamma_{00} + \gamma_{01}PREDICTOR_i + \zeta_{0i} \quad (2)$$

$$\pi_{1i} = \gamma_{10} + \gamma_{11}PREDICTOR_i + \zeta_{1i} \quad (3)$$

These formulas treat the components of an individual's growth trajectory (intercept, π_{0i} and slope, π_{1i}) as outcomes associated with the selected predictor, PREDICTOR. Four level-2 parameters (i.e., γ_{00} , γ_{01} , γ_{10} , and γ_{11}), known as fixed effects, capture systematic differences between change trajectories that are attributable to the level-2 predictor (i.e., PREDICTOR). The residuals (ζ_{0i} and ζ_{1i}) represent the portions of the level-2 outcomes

that remain unexplained by the level-2 predictor (Graham et al., 2009; Singer & Willett, 2003).

Fitting the model to data and interpreting the results. The present study used the statistical package R 2.15 for fitting the multilevel model for change to data. The fixed effects computed by the software package quantified the relationship between time-invariant predictors and individual change trajectories. Each participant had a single set of time-invariant predictors that represented the amount of deliberate practice completed by the reader. In the study, these fixed effects described the relationship between individual growth parameters and the amount of deliberate practice completed. These values were interpreted in the same manner as ordinary regression coefficients with the understanding that the outcomes π_{0i} and π_{1i} , represented the level-1 individual growth parameters (i.e., intercept and slope). Substituting the numeric values yielded the fitted level-2 model; Equation (2) describes the initial status and Equation (3) describes the slope. The results were interpreted graphically by plotting the trajectories of prototypical individuals (Graham et al, 2009; Singer & Willett, 2003). While multilevel software packages provide estimates of the fixed effects, they also provide statistics that describe the variance components associated with each effect. These quantities help researchers determine the remaining variability in each estimate that was not captured by the predictors.

The taxonomy of models that were systematically fit to the data is described in the remainder of this section. The present study explored three functional forms during model fitting: linear, quadratic, and negative exponential. The first form represented time as a linear predictor. The quadratic model included time², providing an estimate of the curvature of student growth. Williamson, Appelbaum, and Epanchin (1991) posited that a linear model

worked well for most students and was useful heuristically in representing student change in reading ability. In contrast, Francis, Shaywitz, Stuebing, Shaywitz, and Fletcher (1996) argued that the quadratic form better represents change in student reading ability over time. However, the mathematical properties of the quadratic include an important downside. The quadratic is parabolic and that means that at some point the prediction of student reading ability could cease to increase and begin to decline. Over the course of time covered by this study (i.e., students in grades 2 through 11), this decline was not expected (Francis et. al, 1996; Williamson et. al, 1991). An alternative functional form, a negative exponential, was tested because it did not have this limitation (Singer & Willett, 2003).

The negative exponential growth model offered a nonlinear change trajectory where exponential growth was dampened by the presence of an asymptote (Figure 5). This asymptote represented a ceiling beyond which the growth trajectory could not rise. A critical aspect of this functional form was that student change would not at any point be predicted to decline, due to the underlying mathematical properties of the functional form (Pinheiro & Bates, 2000; Singer & Willett, 2003). The negative exponential model contained three level-1 individual growth parameters: asymptote (α_i), initial status (π_{0i}), and rate (π_{1i}) (Pinheiro & Bates, 2000). The functional form for the negative exponential model was represented by:

$$Y_{ij} = \alpha_i - (\alpha_i - \pi_{0i})e^{-(\pi_{1i})Time_{ij}} + \varepsilon_{ij} \quad (4)$$

Prototypical negative exponential

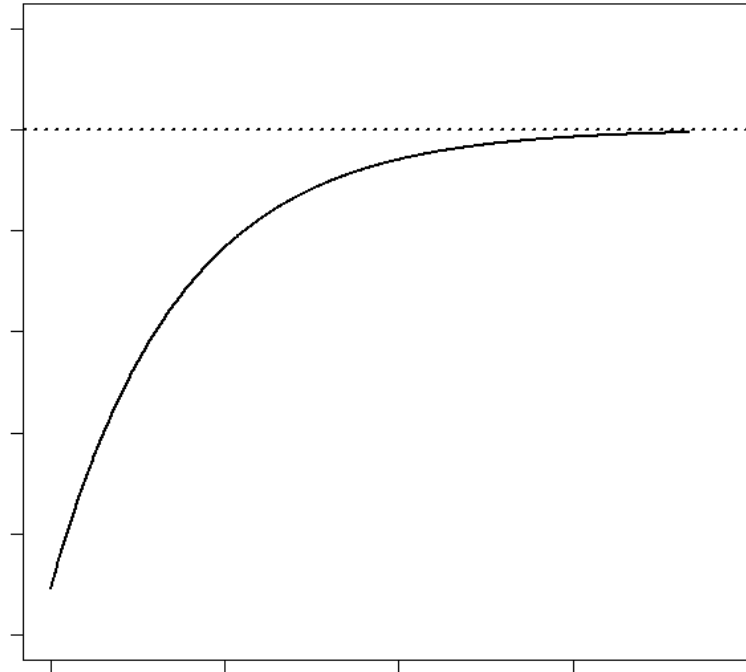


Figure 5. Prototypical negative exponential with asymptote depicted.

Given that theory would suggest that reading ability does not decline in the ways that a quadratic model would allow, the negative exponential model was deemed a plausible additional model to test.

A taxonomy of statistical models. According to Singer and Willett (2003), a taxonomy of statistical models is a systematic sequence of models that help to address research questions. Each model in the taxonomy extends a previous model, allowing inspection and comparison between models. This comparison provides information about the individual and joint effects of predictor variables. This section describes the taxonomy of models for the present study.

The first model fit to data was the *unconditional means model* (Model A). The model can be represented as:

$$Y_{ij} = \pi_{0i} + \varepsilon_{ij} \quad (\text{A.1})$$

$$\pi_{0i} = \gamma_{00} + \zeta_{0i} \quad (\text{A.2})$$

This model did not include change over time or predictors. Rather, it described the outcome variation in reading ability. Fitting the data to this model first was essential as Model A partitioned the variation in the outcome variable, reading ability, into within-person and between-person variance. If either variance component was not statistically significantly different than zero, too little variation existed at that level to be explained. The intraclass correlation coefficient, computed from these variance components, described the proportion of outcome variation that existed between people (Schnabel, Little, & Baumert, 2008; Singer & Willett, 2003).

The second model fit to data was the *unconditional linear growth model* (Model B).

The model can be represented as:

$$Y_{ij} = \pi_{0i} + \pi_{1i}(Time_{ij}) + \varepsilon_{ij} \quad (\text{B.1})$$

$$\pi_{0i} = \gamma_{00} + \zeta_{0i} \quad (\text{B.2.1})$$

$$\pi_{1i} = \gamma_{10} + \zeta_{1i} \quad (\text{B.2.2})$$

This model included time as the only level-1 predictor variable. Adding a predictor variable to level-1 altered the interpretation of the variance components. The level-1 residual variance summarized the scatter of each participant's data around his or her individual change trajectory. The level-2 variances represented the between-person variability in initial status and rate of change respectively. This partitioning helped determine if differences between people predicted differences in true initial reading ability and true rate of change. The covariance structure of the level-2 variance terms was expressed as a correlation coefficient

to determine how initial reading ability and rate of change were related (Singer & Willett, 2003). It was expected that participants who began at a lower reading ability would grow faster than participants who began at a higher level (i.e., Hypothesis 6) (Francis, et. al, 1996). Two additional *unconditional growth models* were fit to the data: an *unconditional quadratic growth model* (Model C), and an *unconditional negative exponential model* (Model D). These models provided information about the amount of variance explained by models of different functional forms.

Assuming the results of fitting Models A, B, C and D to the data suggested that variance in level-1 parameters might be explainable by predictor variables, the time-invariant predictors of deliberate practice were included in the next series of models. There were two types of variables included in the model: question predictors and control predictors (Singer and Willett, 2003). Question predictors represented variables of substantive interest. For the present study, the question predictors were the values describing deliberate practice (i.e., words read, items taken, intense minutes per day, days used, and standard deviation of elapsed days between reading encounters). In contrast, control predictors represented effects that were included in the model to account for variability not related to deliberate practice. For this study, initial grade, initial reading ability, and lunch status served as control predictors. Research suggested that both grade and initial reading ability of a participant would predict growth (e.g., Francis, et. al., 1996).

The model describing deliberate practice and controlling for initial reading ability and initial grade (Model E) can be represented as:

$$Y_{ij} = \pi_{0i} + \pi_{1i}(Time_{ij}) + \varepsilon_{ij} \quad (C.1)$$

$$\pi_{0i} = \gamma_{00} + \gamma_{01}InitialGrade_i + \gamma_{02}InitialLexile_i + \gamma_{03}SES_i + \zeta_{0i} \quad (C.2)$$

$$\begin{aligned} \pi_{1i} = & \gamma_{10} + \gamma_{11}InitialGrade_i + \gamma_{12}InitialLexile_i + \gamma_{13}SES_i + \\ & \gamma_{14}Words_i + \gamma_{15}Items_i + \gamma_{16}IntenseMinutesPerDay_i + \\ & \gamma_{17}DaysUsed_i + \gamma_{18}DaysElapsedSD_i + \zeta_{1i} \end{aligned} \quad (C.3)$$

The level-1 submodel (C.1) contains time, measured in days since beginning to use *Learning Oasis*, as the only level-1 predictor. The first part of the level-2 submodel (C.2) represents initial status and includes the control predictors representing the initial grade, initial reading ability, and SES of the participant. Each of these predictors was hypothesized to significantly predict the intercept. The second part of the level-2 submodel (C.3) represents the rate of change and included the control variables (i.e., initial reading ability, grade, SES) as well as the predictors describing deliberate practice. These predictors only appear in equation C.3 because they can influence the rate of change, but not the initial status. Similar models were constructed for the quadratic (Model E) and negative exponential (Model F) functional forms.

A composite model provides an alternative representation of equations C.1, C.2 and C.3. While mathematically equivalent, it offers a useful way to represent the model as it highlights the cross-level interactions between each predictor and time. The composite model is presented in equation C.4.

$$\begin{aligned}
Y_{ij} = & \gamma_{00} + \gamma_{01}InitialGrade_i + \gamma_{02}InitialLexile_i + \gamma_{03}SES_i + \gamma_{10}Time_{ij} + \\
& \gamma_{11}(InitialGrade_i \times Time_{ij}) + \gamma_{12}(InitialLexile_i \times Time_{ij}) + \gamma_{13}(SES_i \times Time_{ij}) + \\
& \gamma_{14}(Words_i \times Time_{ij}) + \gamma_{15}(Items_i \times Time_{ij}) + \gamma_{16}(IntenseMinutesPerDay_i \times Time_{ij}) + \\
& \gamma_{17}(DaysUsed_i \times Time_{ij}) + \gamma_{18}(DaysElapsedSD_i \times Time_{ij}) + \zeta_{0i} + \zeta_{1i}Time_{ij} + \varepsilon_{ij}
\end{aligned}
\tag{C.4}$$

The hypotheses for this study were:

- H1: The *number of words read from targeted text* will positively relate to the rate of change in reading ability over time
- H2: The *number of reading comprehension items answered* will positively relate to the rate of change in reading ability over time
- H3: The *number of minutes spent reading intensely per day, measured automatically by the educational technology*, will positively relate to the rate of change in reading ability over time
- H4: The *number of days a participant read* will positively relate to the rate of change in reading ability over time
- H5: The *standard deviation of elapsed days between reading experiences* will negatively relate to the rate of change in reading ability over time
- H6: The rate of change in reading ability over time will be negatively related to the *initial reading ability* of the participant
- H7: The rate of change in reading ability over time will be negatively related to the *initial grade of the participant*
- H8-a: The *socio-economic status (SES) of a participant* will be positively related to initial reading ability. Participants with lower SES, as measured by receiving free lunch,

will have a lower initial intercept than participants with higher SES, as measured by paying for lunch.

H8-b: The *SES of a participant* will be positively related to rate of change in reading ability over time. As predicted by the reading literature, participants with lower SES will have less steep trajectories of change than participants with higher SES.

Model interpretation, goodness-of-fit, and examining assumptions. The final steps in the analysis plan were to interpret the fitted model, examine goodness-of-fit indices, and investigate the tenability of the model assumptions. The model was interpreted by examining the values and significance of each predictor variable. The sign and magnitude of the fixed-effect parameter estimates provided insight into how each component of deliberate practice influenced change in reading ability.

A variety of approaches have been proposed to examine model fit. Each approach requires different criteria and offers different interpretive power. The present study applied two different techniques: deviance statistics and Wald statistics. According to Singer and Willett (2003), deviance statistics are preferable to single parameter tests (e.g., *t*-tests) because they offer superior statistical properties and allow composite tests on several parameters simultaneously. The deviance statistic compares the log-likelihood statistics for two models as long as the models were fit to identical data and are nested (Singer & Willett, 2003). Other deviance-based statistics, such as the Akaike Information Criterion (AIC; Akaike, 1974) and Bayesian Information Criterion (BIC; Schwarz, 1978), provide ad hoc estimates of model fit and relax the nesting requirement; however, there is a lack of interpretative framework for these statistics beyond “smaller is better” (Gelman & Rubin,

1995; Raftery, 1995). Both approaches to deviance-based examination of fit were employed given that both nested and non-nested models were explored.

In addition to deviance statistics, the present study examined both single-parameter and composite Wald statistics. Wald statistics, a generalization of dividing a parameter estimate by its standard error, provide a method to test hypotheses about single parameters and composite effects between multiple fixed effects regardless of the method of estimation (Harrell, 2001). Wald statistics are useful for examining sets of predictors as a group (Singer & Willett, 2003). As the study met the requirements for using deviance statistics, Wald statistics provided a second means to examine model fit.

The final step of the data analysis was to investigate the assumptions made in fitting the model. Violations of assumptions can result in biased parameter estimates, inaccurate standard errors, and misleading inferences (Hedeker & Gibbons, 2006; Singer & Willett, 2003). By employing a maximum likelihood estimator and fitting a multilevel model, it was assumed that the level-1 and level-2 errors were independent and normally distributed with constant variance. These assumptions were examined by checking the functional form, and exploring for deviations from normality and homoscedasticity.

The functional form of the model was examined by inspecting plots of outcomes versus predictor variables. At level-1, a random sample was selected and participant growth plots along with an OLS-estimated change trajectory were examined. At level-2, OLS estimates of individual growth were plotted against each individual level-2 predictor. Inspection of the plots confirmed the accuracy of the hypothesized shape.

Visual inspection of plots provided insight into the tenability of the assumptions of normality and homoscedasticity. Normal probability plots, plots of residual values against

normal scores, for the single level-1 and two level-2 residuals provided information as to the normality assumption. On these graphs, departures from normality were indicated by departures from a line. Plots of standardized residuals also provided information about the tenability of the normality assumption. The final assumption, equal variances of level-1 and level-2 residuals for each predictor was examined by plotting raw residuals against predictors. If the residual variability was approximately equal for each predictor, the assumption was determined reasonable.

CHAPTER 4

RESULTS

The results of data analysis are presented in four sections. First, exploratory analysis for all qualifying participants ($N = 1,369$) is presented. Next, the taxonomy of models and results of model fitting are described. Third, the hypotheses for the present study are revisited based on the results of model fitting. Finally, the assumptions of shape, normality, and homoscedasticity are examined.

Exploratory Analysis

Exploratory analysis for all participants was conducted to identify important features of the data. Each participant in the sample met the minimum data requirements (i.e., three measurement occasions separated by at least three months each) and no identifying data were missing. Information about initial grade (Table 3) and key demographic predictors (i.e., gender, ethnicity, lunch status) (Table 4) was available for each participant in the sample so there were no missing data.

Individual change over time and change across participants was explored by both graphical and statistical means using three parametric forms: linear, quadratic, and negative exponential. Empirical growth plots for sixteen random participants were generated; for each selected participant, measures of reading ability were graphed along with linear, quadratic, and negative exponential curves respectively. Plots using these three parametric forms were

also modeled and graphed using all participants in the sample (Appendix A). The white line on each graph represented the average change trajectory of the entire sample. The shape of the average change trajectory using linear, quadratic, and negative exponential fits suggested that participant reading ability, on average, positively increased over time. Further analysis designed to explore change within and between participants was conducted by stratifying participants into groups based on time-invariant predictors-of-interest (i.e., initial grade, initial reading ability, lunch status) (Appendix A).

Sample-wide statistical exploration was conducted by examining the mean, standard deviation, and correlation between each growth parameter (i.e., initial status, rate parameter, quadratic, asymptote) (Table 9). For the linear fit, the correlation coefficient of -0.69 between initial status and rate parameter suggested that participants that began using *Learning Oasis* at a higher Lexile reader measure grew at a slower rate than participants that began at a lower Lexile reader measure. The correlation coefficient between initial status and the rate parameter for the quadratic fit ($r = -0.84$) mirrored this finding. Also of note is that participants who experienced the highest rate of change exhibited the smallest deceleration (i.e., quadratic). The negative exponential fit statistics suggested that the mean asymptote was 1303L and the correlation coefficient between the intercept and rate parameter was -0.02. Based on the computed R^2 values for the linear ($R^2 = 0.45$), quadratic ($R^2 = 0.58$) and negative exponential ($R^2 = 0.71$) fits, it appeared that the negative exponential model fit the data best.

Table 9. Descriptive statistics for the individual growth parameters obtained by fitting separate within-person OLS regression models ($N = 1,369$)

	Linear <i>Mean</i> (<i>SD</i>)	Quadratic <i>Mean</i> (<i>SD</i>)	Negative Exponential <i>Mean</i> (<i>SD</i>)
Initial Status (IS)	838.0L (354.1)	746.6L (884.6)	576.4L (407.9)
Rate Parameter (RP)	0.31 (0.31)	0.71 (1.81)	-5.88 (1.16)
Quadratic/Curvature (QC)		-0.00026 (0.002)	
Asymptote (AS)			1303.2L (775.8)
R^2	0.45 (0.28)	0.58 (0.27)	0.71 (0.19)
Correlations			
IS-RP	-0.69	-0.84	-0.02
IS-QC		0.38	
RP-QC		-0.75	
IS-AS			0.12
RP-AS			-0.56

The data collected by *Learning Oasis* provided detailed accounts of participant reading activity (Table 10). On average, participants had access to *Learning Oasis* for 1,422 calendar days (i.e., includes weekends, holidays, and summer break) and read at least one article on 92 of those days (6.5%). The average number of measurement occasions was 22.3 (Figure 6), and the average number of articles read was 212.2. On average, the estimates of the components of deliberate practice included: 151,574 targeted words read, 2,324 items, 9.03 intense minutes per day, 92.01 days of reading, and a standard deviation of days

between reading encounters of 34.1. Reading data organized by predictors-of-interest was also examined (Appendix B).

Table 10. Descriptive statistics for amount of reading and components of deliberate practice

Component	Mean (SD)	Min	Max
Number of days able to access <i>Learning Oasis</i> (measured from first login to last login)	1,422.1 (388.5)	197	1,961
Number of measurement occasions	22.3 (15.4)	3	119
Number of articles read	212.2 (102.3)	35	905
Amount of conditional reading volume (words read)	151,574 (113,117)	9,173	625,279
Amount of real-time feedback (items)	2,324 (1,213)	319	9,459
Amount of intensive practice (minutes per day)	9.03 (5.09)	2.04	33.64
Distributed practice: Number of days of reading	92.01 (42.33)	7	284
Distributed practice: Standard deviation of days between readings	34.1 (14.6)	8.4	149.0

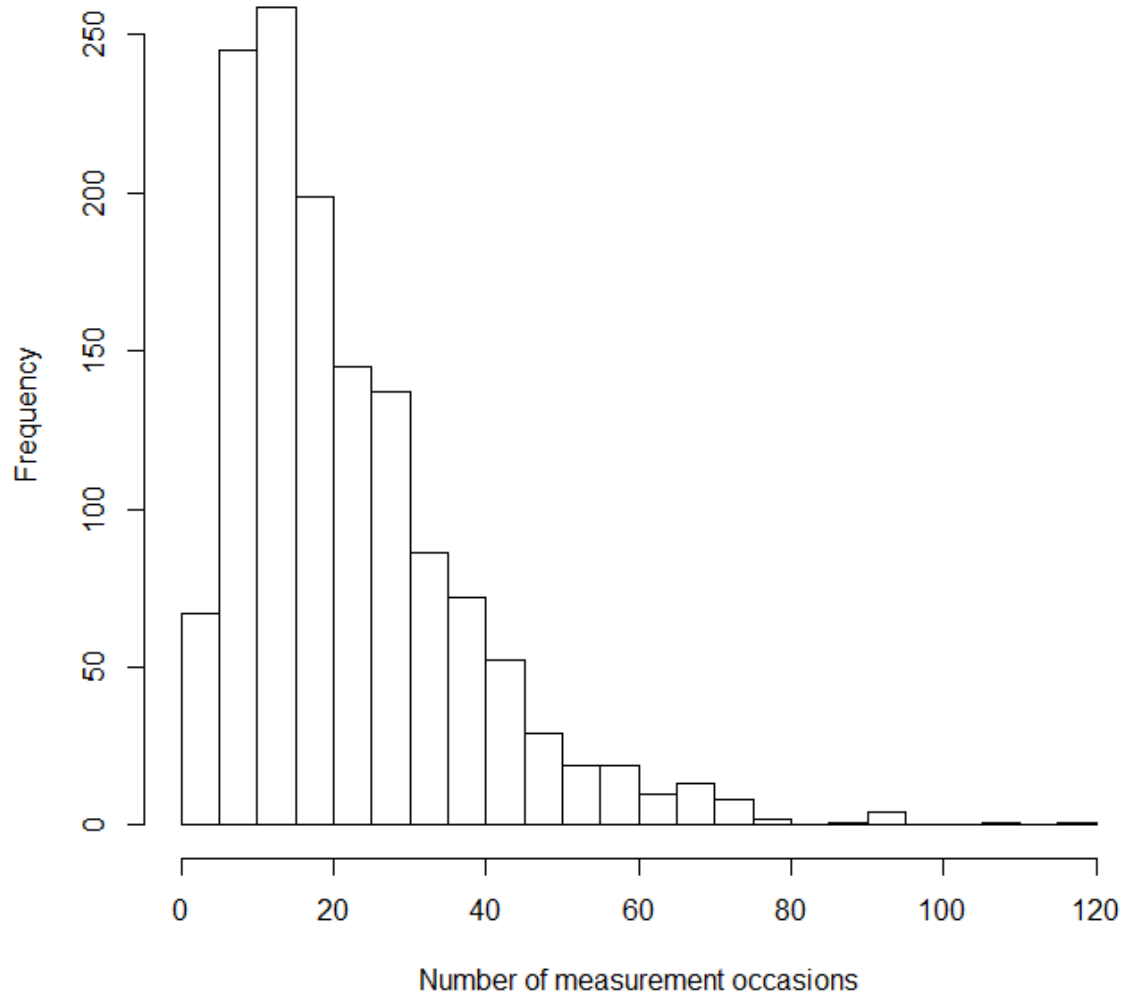


Figure 6. Histogram of measurement occasions.

Correlations between variables of interest were computed (Table 11). Initial grade was highly correlated to Initial Lexile, suggesting that participants who started using *Learning Oasis* in a later grade had higher starting Lexile reader measures than participants starting in an earlier grade. Initial grade was also strongly related to words read and minutes spent reading per day, suggesting that participants in higher grades read more words and spent more time reading per day. Lunch status was statistically significantly related only to

words read ($r = 0.14$) and minutes per day ($r = 0.16$); however the correlations were relatively small. Words read was highly correlated ($r = 0.82$) with items answered. This relationship reflects that the number of items that a participant saw was a function of the length of a passage of text. Given the magnitude of select correlations, variance inflation factors (VIF) were computed to investigate potential collinearity in predictors. Three variables (i.e., initial grade, words, items) had VIF statistics above 4.0, the suggested “rule of thumb” for testing for potential collinearity (Fox, 2008); however, none of them were large enough to merit a priori removal from the study. Subsequent analysis, with the fully fitted mixed model, was conducted to examine possible multicollinearity.

Table 11. Correlations between predictors of interest

	A	B	C	D	E	F	G	H
A: Initial Grade	-	0.78***	0.11***	0.73***	0.41***	0.74***	-0.34***	-0.12***
B: Initial Lexile		-	0.25***	0.59***	0.19***	0.62***	-0.29***	-0.07**
C: Lunch Status			-	0.14***	0.00	0.16***	0.00	-0.04
D: Amount of conditional reading volume (words read)				-	0.82***	0.72***	0.10***	-0.26***
E: Amount of real-time feedback (items)					-	0.41***	0.43***	-0.31***
F: Amount of intensive practice (minutes per day)						-	-0.39***	-0.04
G: Distributed practice: Number of days of reading							-	-0.26***
H: Distributed practice: SD of days between readings								-

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Fitting a Taxonomy of Models

A series of models were fit to participant reading data (Table 12). Model A represented an unconditional means model. Models B, C, and D represented unconditional growth models using linear, quadratic, and negative exponential parameters respectively. Models E, F, and G represented the full model, with all predictors, using linear, quadratic, and negative exponential terms respectively. Additional models were tested once the best fitting functional form was determined. Those models will be discussed later. The fixed effects are presented as unstandardized values that reflect model parameters on the original scale of each fixed effect (Table 13). The *p*-values reflect single-parameter Wald tests for significance.

Variance components, pseudo- R^2 estimates, and goodness-of-fit as measured by deviance, AIC, and BIC, were computed for each model (Table 14). Pseudo- R^2 estimates were computed from the variance components and provided an index of how much variation was explained at a particular level (Singer & Willett, 2003). Subsequent models were compared to the previous accepted model (Table 15) based on the results of the deviance-based Likelihood Ratio Test (LTR), model-level Wald tests, and comparison of AIC and BIC values. As the AIC and BIC values lacked an interpretive framework, the model with a smaller value was judged a better fit (Singer & Willett, 2003). Some models were compared using both approaches (i.e., the nested models), and other models (i.e., the non-nested

Table 12. Taxonomy of multilevel models for change fitted to participant reading data

Model	Level-1 model	Level-2 model
A	$Y_{ij} = \pi_{0i} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \zeta_{0i}$
B	$Y_{ij} = \pi_{0i} + \pi_{1i}(Time_{ij}) + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \zeta_{1i}$
C	$Y_{ij} = \pi_{0i} + \pi_{1i}(Time_{ij})$ $+ \pi_{2i}(Time_{ij})^2 + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \zeta_{1i}$ $\pi_{2i} = \gamma_{20} + \zeta_{2i}$
D	$Y_{ij} = \alpha_i - (\alpha_i - \pi_{0i}) \times$ $e^{-(e^{\pi_{1i}})Time_{ij}} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \zeta_{1i}$ $\alpha_i = \gamma_{20} + \zeta_{2i}$
E	$Y_{ij} = \pi_{0i} + \pi_{1i}(Time_{ij}) + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \gamma_{01}InitialGrade_i + \gamma_{02}InitialLexile_i + \gamma_{03}SES_i + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \gamma_{11}InitialGrade_i + \gamma_{12}InitialLexile_i + \gamma_{13}SES_i + \gamma_{14}Words_i + \gamma_{15}Items_i$ $+ \gamma_{16}IntenseMinPerDay_i + \gamma_{17}DaysUsed_i + \gamma_{18}DaysElapsedSD_i + \zeta_{1i}$
F	$Y_{ij} = \pi_{0i} + \pi_{1i}(Time_{ij})$ $+ \pi_{2i}(Time_{ij})^2 + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \gamma_{01}InitialGrade_i + \gamma_{02}InitialLexile_i + \gamma_{03}SES_i + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \gamma_{11}InitialGrade_i + \gamma_{12}InitialLexile_i + \gamma_{13}SES_i + \gamma_{14}Words_i + \gamma_{15}Items_i$ $+ \gamma_{16}IntenseMinPerDay_i + \gamma_{17}DaysUsed_i + \gamma_{18}DaysElapsedSD_i + \zeta_{1i}$ $\pi_{2i} = \gamma_{20} + \gamma_{21}InitialGrade_i + \gamma_{22}InitialLexile_i + \gamma_{23}SES_i + \gamma_{24}Words_i + \gamma_{25}Items_i$ $+ \gamma_{26}IntenseMinPerDay_i + \gamma_{27}DaysUsed_i + \gamma_{28}DaysElapsedSD_i + \zeta_{2i}$

G	$Y_{ij} = \alpha_i - (\alpha_i - \pi_{0i}) \times e^{-(e^{\pi_{1i}})Time_{ij}} + \varepsilon_{ij}$	$\begin{aligned} \pi_{0i} &= \gamma_{00} + \gamma_{01}InitialGrade_i + \gamma_{02}InitialLexile_i + \gamma_{03}SES_i + \zeta_{0i} \\ \pi_{1i} &= \gamma_{10} + \gamma_{11}InitialGrade_i + \gamma_{12}InitialLexile_i + \gamma_{13}SES_i + \gamma_{14}Words_i + \gamma_{15}Items_i \\ &\quad + \gamma_{16}IntenseMinPerDay_i + \gamma_{17}DaysUsed_i + \gamma_{18}DaysElapsedSD_i + \zeta_{1i} \\ \alpha_i &= \gamma_{20} + \gamma_{21}InitialGrade_i + \gamma_{22}InitialLexile_i + \gamma_{23}SES_i + \gamma_{24}Words_i + \gamma_{25}Items_i \\ &\quad + \gamma_{26}IntenseMinPerDay_i + \gamma_{27}DaysUsed_i + \gamma_{28}DaysElapsedSD_i + \zeta_{2i} \end{aligned}$
H	$Y_{ij} = \alpha_i - (\alpha_i - \pi_{0i}) \times e^{-(e^{\pi_{1i}})Time_{ij}} + \varepsilon_{ij}$	$\begin{aligned} \pi_{0i} &= \gamma_{00} + \gamma_{02}InitialLexile_i + \gamma_{03}SES_i + \zeta_{0i} \\ \pi_{1i} &= \gamma_{10} + \gamma_{12}InitialLexile_i + \gamma_{13}SES_i + \gamma_{14}Words_i + \gamma_{16}IntenseMinPerDay_i \\ &\quad + \gamma_{17}DaysUsed_i + \gamma_{18}DaysElapsedSD_i + \zeta_{1i} \\ \alpha_i &= \gamma_{20} + \gamma_{22}InitialLexile_i + \gamma_{23}SES_i + \gamma_{24}Words_i \\ &\quad + \gamma_{26}IntenseMinPerDay_i + \gamma_{27}DaysUsed_i + \gamma_{28}DaysElapsedSD_i + \zeta_{2i} \end{aligned}$
I	$Y_{ij} = \alpha_i - (\alpha_i - \pi_{0i}) \times e^{-(e^{\pi_{1i}})Time_{ij}} + \varepsilon_{ij}$	$\begin{aligned} \pi_{0i} &= \gamma_{00} + \gamma_{02}InitialLexile_i + \zeta_{0i} \\ \pi_{1i} &= \gamma_{10} + \gamma_{13}SES_i + \gamma_{14}Words_i + \gamma_{16}IntenseMinPerDay_i + \zeta_{1i} \\ \alpha_i &= \gamma_{20} + \gamma_{22}InitialLexile_i + \gamma_{23}SES_i + \gamma_{27}DaysUsed_i + \gamma_{28}DaysElapsedSD_i + \zeta_{2i} \end{aligned}$

models) could only be compared using Wald tests and AIC and BIC comparisons. During initial model building, emphasis was placed on examination of the variance components and goodness-of-fit statistics in order to find the best fitting model. A discussion of the fixed effects and the implications for the present study is presented in the next section.

After fitting Model A, the unconditional means model, the significance of the fixed and random effects ($p < .001$) suggested that there was within and between-participant variance unaccounted for by the model. Each variance component was statistically significantly different from zero and subsequent models included predictors to model within-person and between-person variation. The intraclass correlation coefficient of 0.75 suggested that 75% of change in reading ability was attributable to differences among participants.

The first unconditional growth model explored, Model B, included only linear time as a predictor. The fixed effects (i.e., intercept and slope) were both statistically significantly different from zero ($p < .001$) suggesting that the average true change trajectory had a non-zero initial status and growth rate. Comparing the within-person variance components of Model A to Model B showed that 83% of within-person variation in reading ability was systematically associated with linear time (Table 14). At level-2, the variance components for both initial status and rate of change were statistically significantly different from zero ($p < .001$). This finding suggested that additional level-2 predictors could be added to the model. All goodness-of-fit indices (e.g., LRT [$\chi^2(3) = 18,251, p < 0.001$], Wald, AIC, BIC) supported the conclusion that Model B was a better fit than Model A.

Table 13. Unstandardized fixed effects: Results of fitting a taxonomy of multilevel models for change to participant reading data ($N = 1,360$)

			Model A	Model B	Model C	Model D	Model E	Model F	Model G
			<i>Mean</i>	<i>Mean</i>	<i>Mean</i>	<i>Mean</i>	<i>Mean</i>	<i>Mean</i>	<i>Mean</i>
			<i>(SE)</i>	<i>(SE)</i>	<i>(SE)</i>	<i>(SE)</i>	<i>(SE)</i>	<i>(SE)</i>	<i>(SE)</i>
Fixed Effects									
Initial Status π_{0i}	Intercept	γ_{00}	980.32*** (7.41)	844.33*** (9.28)	780.82*** (10.07)	720.45*** (10.74)	132.84*** (8.37)	28.11*** (8.44)	-60.07*** (8.53)
	Initial Grade	γ_{01}					26.82*** (1.94)	20.90*** (1.93)	14.82*** (1.94)
	Initial Lexile	γ_{02}					0.724*** (0.016)	0.835*** (0.016)	0.929*** (0.016)
	SES	γ_{03}					33.11*** (6.87)	25.38*** (6.86)	9.80 (6.97)
Rate Parameter π_{1i}	Intercept	γ_{10}		0.271*** (0.006)	0.574*** (0.013)	-5.910*** (0.031)	0.668*** (0.028)	1.085*** (0.067)	-7.180*** (0.197)
	Initial Grade	γ_{11}					-0.023*** (0.003)	0.008 (0.008)	0.165*** (0.024)
	Initial Lexile	γ_{12}					-0.0004*** (< 0.0001)	-0.001*** (< 0.0001)	-0.0008*** (0.0002)
	SES	γ_{13}					0.005 (0.009)	0.040 (0.022)	0.151* (0.059)
	Words Read	γ_{14}					6.60e-07*** (< 0.0001)	1.89e-06*** (< 0.0001)	4.07e-06*** (< 0.0001)
	Items	γ_{15}					-0.00008*** (< 0.0001)	-0.0002*** (< 0.0001)	-0.0005*** (< 0.0001)
	Min. per Day	γ_{16}					0.00665*** (0.0015)	0.0064 (0.0037)	0.0794*** (0.0131)
	Days Used	γ_{17}					0.00095*** (0.00015)	0.0035*** (0.00037)	0.0092*** (0.0012)

	SD Days Elapsed	γ_{18}			-0.00074* (0.00030)	-0.0005 (0.0007)	-0.0027 (0.0021)
Quadratic/ Asymptote ¹	Intercept	γ_{20}	-0.0002*** (< 0.0001)	1202.875*** (5.663)		-0.0003*** (0.00005)	1013.59*** (35.89)
	Initial Grade	γ_{21}				-0.00002*** (< 0.0001)	-8.540* (3.771)
	Initial Lexile	γ_{22}				4.60e-07*** (< 0.0001)	0.178*** (0.024)
	SES	γ_{23}				-0.00003* (< 0.0001)	13.44 (8.92)
	Words Read	γ_{24}				-5.39e-10 (< 0.0001)	0.0012*** (0.0001)
	Items	γ_{25}				7.07e-08*** (< 0.0001)	-0.1022*** (0.0089)
	Min. per Day	γ_{26}				-3.50e-06 (< 0.0001)	1.532 (1.521)
	Days Used	γ_{27}				-1.77e-06*** (< 0.0001)	0.915*** (0.189)
	SD Days Elapsed	γ_{28}				8.09e-07 (< 0.0001)	1.453*** (0.386)

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ [p -values computed via Wald statistic]

¹ Parameter estimates of quadratic term (π_{2i}) for Model C and Model F and asymptote term (α_i) for Model D and Model G

Table 14. Variance components and goodness-of-fit indices: Results of fitting a taxonomy of multilevel models for change to participant reading data ($N = 1,360$)

Variance Components			Model A (SD)	Model B (SD)	Model C (SD)	Model D (SD)	Model E (SD)	Model F (SD)	Model G (SD)
Level-1	Within-person	σ_{ϵ}^2	7.33e04*** (270.80)	1.24e04*** (338.31)	1.01e04*** (100.41)	9.78e03*** (98.90)	1.24e04*** (111.24)	1.01e04*** (100.36)	9.79e03*** (98.95)
	In initial status	σ_0^2	2.50e04*** (158.24)	1.14e05*** (338.31)	1.33e04*** (364,143.6)	1.45e05*** (381.38)	1.21e04*** (109.97))	1.04e04*** (102.19)	8.76e03*** (93.61)
	In rate parameter	σ_1^2		0.043*** (0.207)	0.182*** (0.427)	0.856*** (0.925)	0.0156*** (0.1250)	0.0908*** (0.3013)	0.689*** (0.830)
	In quadratic	σ_2^2			4.7e-08*** (0.000217)			3.6e-08*** (0.00019)	
	Asymptote	σ_{α}^2				3.05e04*** (174.74)			1.39e04*** (117.69)
Pseudo- R^2 and Goodness of Fit									
	R^2 (within-person)	R_{ϵ}^2		0.83	0.86	0.87	0.83	0.86	0.87
	R^2 (initial status)	R_0^2					0.89	0.22	0.94
	R^2 (rate parameter)	R_1^2					0.64	0.50	0.20
	R^2 (quadratic)	R_2^2						0.23	
	R^2 (asymptote)	R_2^2							0.54
	AIC		401,452	383,207	378,394	377,536	380,239	375,047	373,917
	BIC		401,477	383,257	378,477	377,620	380,380	375,288	374,159
	Deviance		401,446	383,194	378,374	377,516	380,205	374,989	373,859

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Model C, the second unconditional growth model, included both time and time² (i.e., quadratic) as predictors. The fixed effects (i.e., intercept, slope, quadratic) were all statistically significant ($p < .001$) (Table 13). This suggested that the average true change curve had a non-zero initial status, rate parameter, and quadratic term. The variance components of Model C suggested that 86% of the within-person variation in reading ability was systematically associated with both linear and quadratic time. The variance components for initial status, rate parameter, and quadratic at level-2 were statistically significantly different from zero ($p < .001$) suggesting that additional level-2 predictors could be added to the model. All goodness-of-fit indices (e.g., LRT [$X^2(8) = 4,822, p < 0.001$], Wald, AIC, BIC) supported the conclusion that Model C was a better fit than Model B.

The third unconditional growth model, Model D, included time within a negative exponential functional form. This parameterization slowed the growth of the typical exponential by including an asymptote to the curve. This asymptote was modeled like the other parameters (i.e., intercept, rate parameter), and was represented in both the fixed and random effects, allowing the parameter to vary between participants and be modeled (Pinheiro & Bates, 2000; Singer & Willett, 2003). It is important to note that the model parameters did not have the same interpretation as in the earlier models, even though they are labeled similarly. The initial status parameter represents the intercept as depicted in equation (4); however, the rate parameter was not technically a representation of slope. But, it determined how rapidly the trajectory approached the asymptote. The parameter estimates for initial status, rate of change, and asymptote were all statistically significant ($p < .001$) suggesting that the average true change trajectory had a non-zero initial status, rate of change, and asymptote. The within-person variance component of Model D suggested that

87% of the within-person variation in reading ability was associated with the exponential parameterization. The variance components for initial status, the rate parameter, and asymptote were all statistically significantly different from zero ($p < .001$) suggesting that additional level-2 predictors could be added to the model. Given that Model D was not nested within Model C, it was not possible to conduct a Likelihood Ratio Test; however, the remaining goodness-of-fit indices suggested that Model D was a better fit than Model C.

Model E included all substantive predictors for the present study and a linear representation of time. *InitialGrade*, *InitialLexile*, and *SES* were included as predictors for both initial status and the rate parameter. The rate parameter submodel was supplemented with the predictors representing the components of deliberate practice (i.e., *Words*, *Items*, *IntenseMinPerDay*, *DaysUsed*, and *DaysElapsedSD*). All predictors of initial status were statistically significant ($p < .001$). All predictors of the rate parameter were statistically significant except *SES* (*ns*). Comparing the variance components of Model E to those of Model D (i.e., the best fitting model thus far) showed that the within-person variance increased slightly, likely due to the removal of the exponential term. The variance components representing initial status and rate of change were statistically significantly different from zero ($p < .001$) suggesting additional explainable variance in both initial status and rate of change remained. Model E explained 89% of the variation in initial status and 64% of the variability in the rate parameter (Table 14). It was not possible to conduct a LRT due to the non-nested nature of Models D and E, but the other indices of goodness-of-fit (e.g., Wald statistics at the model level, deviance, AIC, and BIC) suggested that Model D was a better fit than Model E. This suggested that the exponential model without any

predictors except time was a better fit than a linear model with all of the predictors-of-interest.

Model F included all the substantive predictors for the present study and both a linear and quadratic representation of time. All three predictors of initial status (i.e. *InitialGrade*, *InitialLexile*, *SES*) were statistically significant ($p < .001$). Four variables were statistically significant ($p < .001$) predictors of rate of change (i.e., *InitialLexile*, *Words*, *Items*, *DaysUsed*) and four were not statistically significant (i.e., *InitialGrade*, *SES*, *IntenseMinPerDay*, *DaysElapsedSD*). Five predictors were statistically significant predictors of the quadratic term: *InitialGrade*, *InitialLexile*, *Items*, *DaysUsed*, and *SES*. The remaining predictors (i.e., *Words*, *IntenseMinPerDay*, *DaysElapsedSD*) were not statistically significant in predicting the quadratic term (Table 13). Examining the within-person variance component of Model F in relation to Model D suggested that within-person variation declined by 1%. The statistically significantly different from zero ($p < .001$) variance components representing initial status, rate of change, and quadratic suggested that additional variance remained in each parameter. Model F explained 22% of the variance in initial status, 50% of the variance in rate of change, and 23% of the variation in the quadratic term (Table 14). Examination of Wald statistics for composite predictors and the AIC and BIC showed that Model F was a better fit than Model D. This suggested that the quadratic representation with all predictors was a better fit than an exponential model without any predictors except time at level-1.

Model G used a negative exponential functional form and contained the various deliberate practice and control variables predicting initial status, the rate parameter, and asymptote (Table 12). Comparing the variance components of Models F and G showed that

Model G explained 1% more of the within-person variation than Model F. The statistically significantly different from zero ($p < .001$) variance components for initial status, rate parameter, and asymptote suggested that additional level-2 predictors could explain additional variability. Two predictors of initial status (i.e., *InitialGrade*, *InitialLexile*) were statistically significant ($p < .001$). *SES*, the remaining predictor of initial status, was not statistically significant. Seven predictors were statistically significant in predicting the rate parameter (i.e., *InitialGrade*, *InitialLexile*, *SES*, *Words*, *Items*, *IntenseMinPerDay*, *DaysUsed*). *ElapsedDaysSD* was not statistically significant in predicting the rate parameter. Six predictors of the asymptotic term were statistically significant in predicting the asymptotic parameter (i.e., *InitialLexile*, *InitialGrade*, *Words*, *Items*, *DaysUsed*, *ElapsedDaysSD*). The remaining two predictors, *SES* and *MinutesPerDay*, were statistically non-significant (Table 13). Model G explained 94% of the variation in initial status, 20% of the variability in the rate parameter, and 54% of the variation in the asymptote (Table 14). The goodness-of-fit indices suggested that Model G was a better fit than Model F. This suggested that the negative exponential parameterization with all predictors was a better fit than a quadratic model with all predictors.

Table 15. Model comparisons made during fitting taxonomy of models to participant reading data using both Likelihood Ratio Test (LRT) and Wald statistics

	Model 1	Model 2	Better fitting model according to...			
			LRT	Wald	AIC	BIC
1	A	B	B	B	B	B
2	B	C	C	C	C	C
3	C	D		D	D	D
4	D	E		D	D	D
5	D	F		F	F	F
6	F	G		G	G	G

Potential multicollinearity was examined by computing VIF statistics for Model G. Consistent with the exploratory analysis, three variables showed VIF values greater than 4.0 (i.e., *Words*, *Items*, *InitialGrade*). To minimize the influence of multicollinearity, a series of models were created based on Model G (Table 16). The VIF statistics for each of the potentially collinear predictors were computed along with goodness-of-fit statistics. Models G.1, G.2, and G.3 contained all predictors except *Words*, *Items*, and *InitialGrade* respectively. Each model reduced the VIF values of the other two predictors; however, their magnitudes suggested potential multicollinearity remained. Given that none of the three models eliminated potential multicollinearity satisfactorily, two additional models were explored. Model G.4 removed both *Words* and *InitialGrade*, and Model H removed both *Items* and *InitialGrade*. Each model reduced the remaining potentially collinear term to the smallest magnitude observed. While the VIF value for the remaining predictor was similar between the two models, Model H was considered a better fit as the goodness-of-fit indices (i.e., AIC, BIC) were smaller.

While Model H reduced the potential influence of multicollinearity, the goodness-of-fit indices suggested that it was not as good a fit as the full Model G; however, an argument could be made that removing *Items* and *InitialGrade* was appropriate. Both visual and statistical evidence supported the removal of *Items*. Possible multicollinearity between *Items* and *Words* was explored by visually inspecting plots of prototypical participants using the full Model G. Two plots were created, each depicting three prototypical participants who had the same values for all predictors (i.e., the mean) except items taken and words read respectively (Figure 6). The three selected values displayed on each plot reflect the mean (i.e., solid line), one standard deviation above the mean (dotted line), and one standard deviation below the mean (dashed line) for items answered (Figure 7, left panel) and words read (Figure 7, right panel) respectively. Surprisingly, the number of items answered was negatively related to the rate parameter in reading ability. This result was unexpected, and highly unlikely given theory. Comparing the shapes of the two plots (i.e., items and words), they appear to be mirror images of one another with the prototypical participant reading the most words also completing the least number of items. The potential collinearity between *Words* and *Items* was also supported by the magnitude and sign of the fixed effect estimates (i.e., almost mirror images of each other). As argued by Menard (2002), a stark contrast in the fixed effects, as evidenced by the parameter estimates and visual inspection of model plots, can be an indicator of collinearity between predictors. As the items a participant encountered in *Learning Oasis* was largely a function of the number of words read, collinearity between words read and items was not unexpected. Given the evidence, *Items* appeared to be a reasonable candidate for removal from the model due to multicollinearity.

Statistical evidence supported the removal of *InitialGrade* in an effort to reduce multicollinearity. *InitialGrade* was correlated above 0.7 with three predictors (i.e., *InitialLexile*, *Words*, *IntenseMinPerDay*). This potential collinearity was expected given the fact that participants in higher grades typically possess greater reading ability and are able to read faster than participants in lower grades. The implementation model suggested that participants in higher grades were also given more time per day to read inside *Learning Oasis*. This information, along with the results of fitting models with and without *InitialGrade* suggested that the information provided by *InitialGrade* was represented by other variables. This finding was consistent with the notion that participants in higher grades often had more prior ability than participants in lower grades.

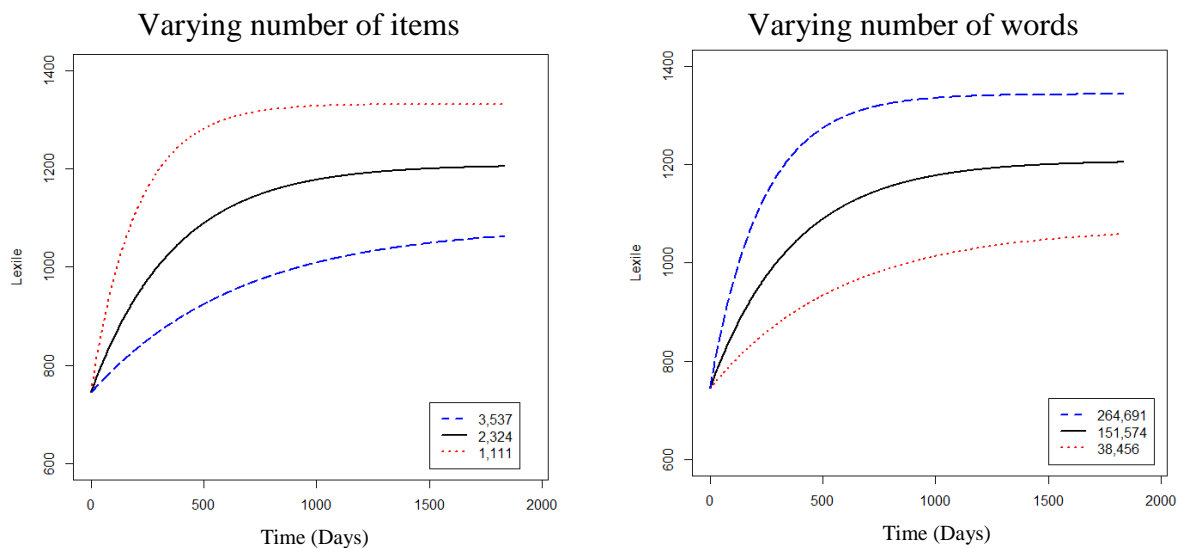


Figure 7. Plot of three prototypical participants. Left panel shows varying the number of items answered. Right panel shows varying the number of words read. The left panel indicates that participants who answered fewer items grew faster and performed at a higher level than participants who answered more items. This is counterintuitive and not consistent with theory.

Table 16. VIF statistics for potentially collinear predictors and goodness-of-fit indices for select models

		Model G	Model G.1	Model G.2	Model G.3	Model G.4	Model H ¹
		<i>Full Model</i>	<i>No Words</i>	<i>No Items</i>	<i>No Init. Grade</i>	<i>No Words No Init. Grade</i>	<i>No Items No Init. Grade</i>
Rate Parameter							
Initial Grade	γ_{11}	5.76	6.92	5.28	-	-	-
Words Read	γ_{14}	15.97	-	7.46	14.57	-	5.38
Items	γ_{15}	11.09	5.63	-	11.04	5.62	-
Asymptote							
Initial Grade	γ_{21}	7.36	6.85	7.52	-	-	-
Words Read	γ_{24}	13.46	-	5.09	13.67	-	4.38
Items	γ_{25}	9.80	5.68	-	9.52	6.18	-
Goodness-of-Fit							
AIC		373,917	374,193	374,074	373,964	374,206	374,182
BIC		374,158	374,391	374,298	374,181	374,406	374,382
Deviance		373,859	374,215	374,029	372,912	374,158	374,134

¹ Model G showed the best data model fit according to goodness-of-fit indices, but suffered from multicollinearity. Of the models that reduced collinearity, Model H was the best fitting.

A top-down simplification approach was applied to Model H in an effort to remove statistically non-significant predictors from the model. Terms were tested sequentially starting with higher-order terms based on increasing absolute t-value magnitude. The simplified model was constructed by comparing the model fit without the predictor-in-question with the previously accepted model. A decision to remove or retain a particular predictor was based on examining goodness-of-fit indices (i.e., LTR, Wald test on composite predictors, AIC, BIC). Six predictors were removed from the model because they did not contribute to model fit. *SES* was removed from the submodel for initial status. *InitialLexile*, *DaysUsed*, and *ElapsedDaysSD* were removed from the rate parameter term as they did not contribute to model fit. Finally, *Words* and *IntenseMinPerDay* were removed from the submodel representing the asymptote (Table 17). These findings were consistent with the single-parameter Wald test conducted during initial fitting of Model H (i.e., these predictors were statistically non-significantly related to their respective level-1 parameter). Each decision about a predictor was unanimous for all indices of fit.

The final model, Model I (Table 17), contained one statistically significant predictor of initial status (i.e., *InitialLexile*), three statistically significant predictors for the rate parameter (i.e., *SES*, *Words Read*, *IntenseMinPerDay*), and four statistically significant predictors of asymptote (i.e., *InitialLexile*, *SES*, *DaysUsed*, *ElapsedDaysSD*). Given that *Words* was a statistically non-significant predictor of the asymptote, the terms removed from the model due to potential collinearity (i.e., *Items*, *InitialGrade*) were added to Model I individually and together to explore any potential influence on the asymptotic parameter. When added to the model individually, both *Items* and *InitialGrade* were statistically non-significant in predicting the asymptote. When both *Items* and *InitialGrade* were added to the

model together, they were both statistically significant predictors of the asymptote; however, the VIF statistics suggested that potential multicollinearity remained. Therefore, Model I, without *Items* and *InitialGrade* was judged the best fitting model that reduced potential collinearity.

Examination of the variance components showed that Model I explained the same within-person variance explained by both Model G and Model H (i.e., 87%). The statistically significantly different from zero ($p < .001$) variance components for initial status, rate parameter, and asymptote suggested that additional predictors could explain additional variability (Table 18). Model I explained 95% of the variability in initial status, 10% of the variation in the rate parameter, and 43% of the variability in the asymptote. While Model I explained less of the variance in the rate parameter and asymptote than Model G, the threat of multicollinearity was minimized according to VIF statistics; therefore, choosing Model I was a more conservative choice. Compared to Model H, Model I was determined to be the better fit based on composite Wald test at the model level, AIC, BIC, and LTR.

Table 17. Fixed effects and VIF values for Model G, Model H, and Model I ($N = 1,360$)

Fixed Effects			Model G		Model H		Model I	
			<i>Full Model</i>		<i>No items or initial grade</i>		<i>Simplified Model H</i>	
			Unstand. (SE)	VIF	Unstand. (SE)	VIF	Unstand. (SE)	VIF
Initial Status	Intercept	γ_{00}	-60.07*** (8.53)	-	-78.73*** (8.39)	-	-72.89*** (8.22)	-
	π_{0i} Initial Grade	γ_{01}	14.82*** (1.94)	2.56	-	-	-	-
	Initial Lexile	γ_{02}	0.929*** (0.016)	2.66	1.053*** (0.010)	1.06	1.049*** (0.010)	1.00
	SES	γ_{03}	9.80 (6.97)	1.09	4.89 (6.70)	1.06	-	-
Rate Parameter	Intercept	γ_{10}	-7.18*** (0.197)	-	-6.80*** (0.19)	-	-6.84*** (0.067)	-
	π_{1i} Initial Grade	γ_{11}	0.165*** (0.024)	5.76	-	-	-	-
	Initial Lexile	γ_{12}	-0.0008 (0.0002)	3.89	0.0002 (0.0001)	2.23	-	-
	SES	γ_{13}	0.151* (0.059)	1.61	0.154* (0.063)	1.47	0.193** (0.061)	1.35
	Words Read	γ_{14}	4.07e-06*** (< 0.0001)	15.97	1.57e-06** (< 0.0001)	5.38	2.38e-06*** (< 0.0001)	2.44
	Items	γ_{15}	-0.0005*** (< 0.0001)	11.09	-	-	-	-
	Min. per Day	γ_{16}	0.0794*** (0.0131)	3.90	0.053*** (0.014)	3.61	0.052*** (0.010)	2.37
	Days Used	γ_{17}	0.0092*** (0.0012)	4.01	0.001 (0.001)	2.82	-	-
	SD Days Elapsed	γ_{18}	-0.0027 (0.0021)	1.75	-0.003 (0.002)	1.67	-	-

Asympt. α_i	Intercept	γ_{20}	1013.59*** (35.89)	-	789.97*** (33.70)	-	832.80*** (26.75)	-
	Initial Grade	γ_{21}	-8.540* (3.771)	7.36	-	-	-	-
	Initial Lexile	γ_{22}	0.178*** (0.024)	3.92	0.300*** (0.020)	2.56	0.312*** (0.015)	1.48
	SES	γ_{23}	13.44 (8.92)	1.64	34.36*** (10.08)	1.49	31.99*** (9.66)	1.42
	Words Read	γ_{24}	0.0012*** (0.0001)	13.46	-2.24e-05 (< 0.0001)	4.38	-	-
	Items	γ_{25}	-0.1022*** (0.0089)	9.80	-	-	-	-
	Min. per Day	γ_{26}	1.532 (1.521)	4.08	3.161 (1.752)	3.83	-	-
	Days Used	γ_{27}	0.915*** (0.189)	4.02	0.766*** (0.175)	3.14	0.673*** (0.122)	1.61
	SD Days Elapsed	γ_{28}	1.453*** (0.386)	1.97	1.629*** (0.427)	1.80	1.341*** (0.333)	1.14

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ [p -values computed via Wald statistic]

Table 18. Variance components and goodness-of-fit indices for Model D, Model G, Model H, and Model I

			Model D <i>Unconditional neg. exponential growth model (SE)</i>	Model G <i>Full model (SE)</i>	Model H <i>Full model without items or initial grade (SE)</i>	Model I <i>Simplified Model H (SE)</i>
Variance Components						
Level-1	Within-person	σ_{ε}^2	9.78e03*** (98.90)	9.79e03*** (98.95)	9.74e03*** (98.68)	9.74e03*** (98.71)
	In initial status	σ_0^2	1.45e05*** (381.38)	8.76e03*** (93.61)	7.73e03*** (87.90)	8.00e03*** (89.43)
	In rate parameter	σ_1^2	0.856*** (0.925)	0.689*** (0.830)	0.740*** (0.894)	0.769*** (0.899)
	Asymptote	σ_{α}^2	3.05e04*** (174.74)	1.39e04*** (117.69)	1.96e04*** (139.86)	1.75e04*** (132.20)
Pseudo-R^2 and Goodness of Fit						
	R^2 (within-person)	R_{ε}^2	0.87	0.87	0.87	0.87
	R^2 (initial status)	R_0^2		0.94	0.95	0.95
	R^2 (rate parameter)	R_1^2		0.20	0.14	0.10
	R^2 (asymptote)	R_{α}^2		0.54	0.36	0.43
	AIC		377,536	373,917	374,182	374,149
	BIC		377,620	374,159	374,382	374,299
	Deviance		377,516	373,859	374,134	374,113

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Hypotheses Revisited

The hypotheses for the present study were examined based on the results of fitting Model I. Given that the negative exponential functional form included an additional estimated parameter (i.e., asymptote), the rate of change was represented using both the rate parameter and asymptote. Additional hypotheses related to each predictor's influence on the asymptote were created to represent these two parameters. In this section, each hypothesis is presented separately along with a discussion of the results of model fitting and a visual representation of the influence of the predictor in question. Each graph contains curves that represent prototypical participants. Only the predictor of interest is varied within a particular graph. This variation, depicted in the legend of each graph, reflects the mean (solid line), one standard deviation above the mean (dashed line) and one standard deviation below the mean (dotted line). All other predictors were held constant at their mean. This allowed the role of individual predictors to be examined visually.

Given that estimates of the rate parameter are related to a non-linear exponential term, and can be difficult to interpret on their own, graphs facilitate interpretation about the influence of a predictor. With this challenge in interpretability, emphasis was placed on the sign and significance of the fixed effects, using the graphical depiction as a way of representing the influence of a particular predictor. These findings reflect Model I, which contains level-2 predictors' cross-level interaction with time as depicted by the multilevel model (Table 12). After the presentation of each hypothesis and associated information, the influence of deliberate practice as a whole is explored.

H1-a: The *number of words read from targeted text* will positively relate to the rate of change in reading ability over time

H1-b: The number of words read from targeted text will positively relate to the magnitude of the asymptote

The *Words* variable represented the amount of conditional reading volume completed by a participant inside *Learning Oasis*. The fixed effect on the rate parameter was statistically significant ($p < 0.001$), and the positive sign suggested that participants who read more words grew more rapidly. The scale for *Words* was actual words read, but the functional form of the model makes interpretation challenging. Visual inspection showed that the prototypical participants who read more words grew more rapidly than the participants who read fewer words (Figure 8). Hypothesis H1-a was retained as number of words read was positively related to the rate of change in reading ability.

In the full model, Model G, the predictor *Words* was a statistically significant contributor to the asymptote term; however, once *Items* was removed from the model (Model H), *Words* was statistically non-significant. *Words* was removed from the asymptotic component of the model during simplification. This suggested that *Words* did not predict the asymptote and that the number of words read by a participant did not predict the upper limit of their reading ability. Hypothesis H1-b was not supported.

H2-a: The number of reading comprehension items answered will positively relate to the rate of change in reading ability over time

H2-b: The number of reading comprehension items answered will positively relate to the asymptote

The predictor *Items* represented how many reading comprehension items a participant answered while using *Learning Oasis*. *Items* was a statistically significant predictor of both the rate parameter and asymptote in Model G; however, it was removed from the model to

reduce the potential influence of multicollinearity. Hypotheses H2-a and H2-b were not examined further using the data and model for this study.

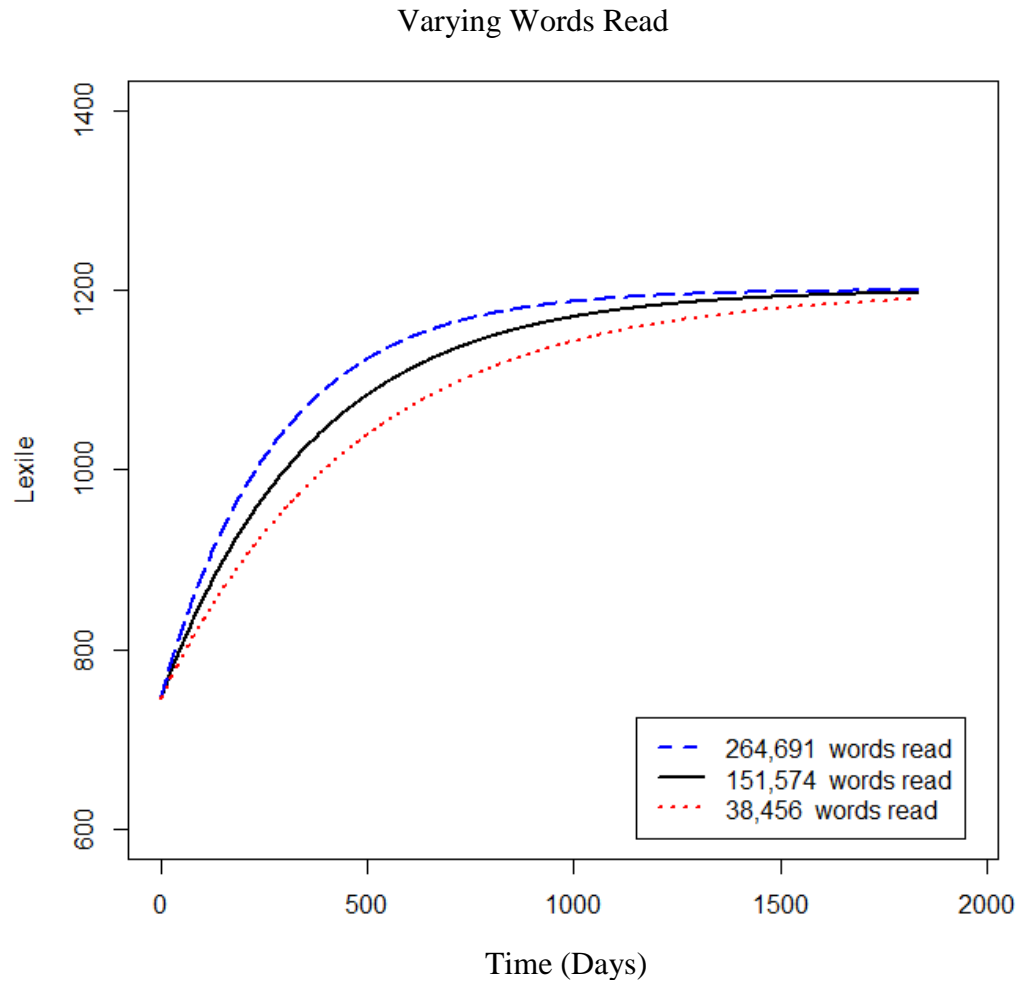


Figure 8. Plot of three prototypical participants, varying only number of words read.

H3-a: The number of minutes spent reading intensely per day, measured automatically by the educational technology, will positively relate to the rate of change in reading ability over time

H3-b: The number of minutes spent reading intensely per day, measured automatically by the educational technology, will positively relate to the asymptote

IntenseMinPerDay represented how many minutes, measured by *Learning Oasis*, a participant read intensely per day. *IntenseMinPerDay* was a statistically significant predictor of the rate parameter. The positive sign of the fixed effect parameter predicting the rate parameter suggested that participants who spent more time reading intensely each day grew at a faster rate than participants who read less intensely per day. Visual inspection of the predictive power of *IntenseMinPerDay* supported this finding and hypothesis H3-a was retained (Figure 9).

IntenseMinPerDay was not statistically significant in predicting the asymptote parameter. It was removed from the level-2 model for the asymptotic term during simplification to reflect that *IntenseMinPerDay* did not predict the asymptote. This means that the number of intense minutes spent reading per day did not predict the upper limit of reading ability. Hypothesis H3-b was not supported.

H4-a: The number of days a participant read will positively relate to the rate of change in reading ability over time

H4-b: The number of days a participant read will positively relate to the asymptote

DaysUsed represented how many calendar days a participant read inside *Learning Oasis*. *DaysUsed* was a statistically significant predictor of the rate parameter in the full model; however, when potentially collinear terms were removed *DaysUsed* was statistically non-significant and it was removed from the model. Therefore, hypothesis H4-a was not supported.

DaysUsed was a statistically significant predictor of asymptote. The positive sign on the fixed effect suggested that participants who read on more days experienced a higher

asymptote than participants who read on fewer days. The graphical representation supported this finding, and hypothesis H4-b was retained (Figure 10).

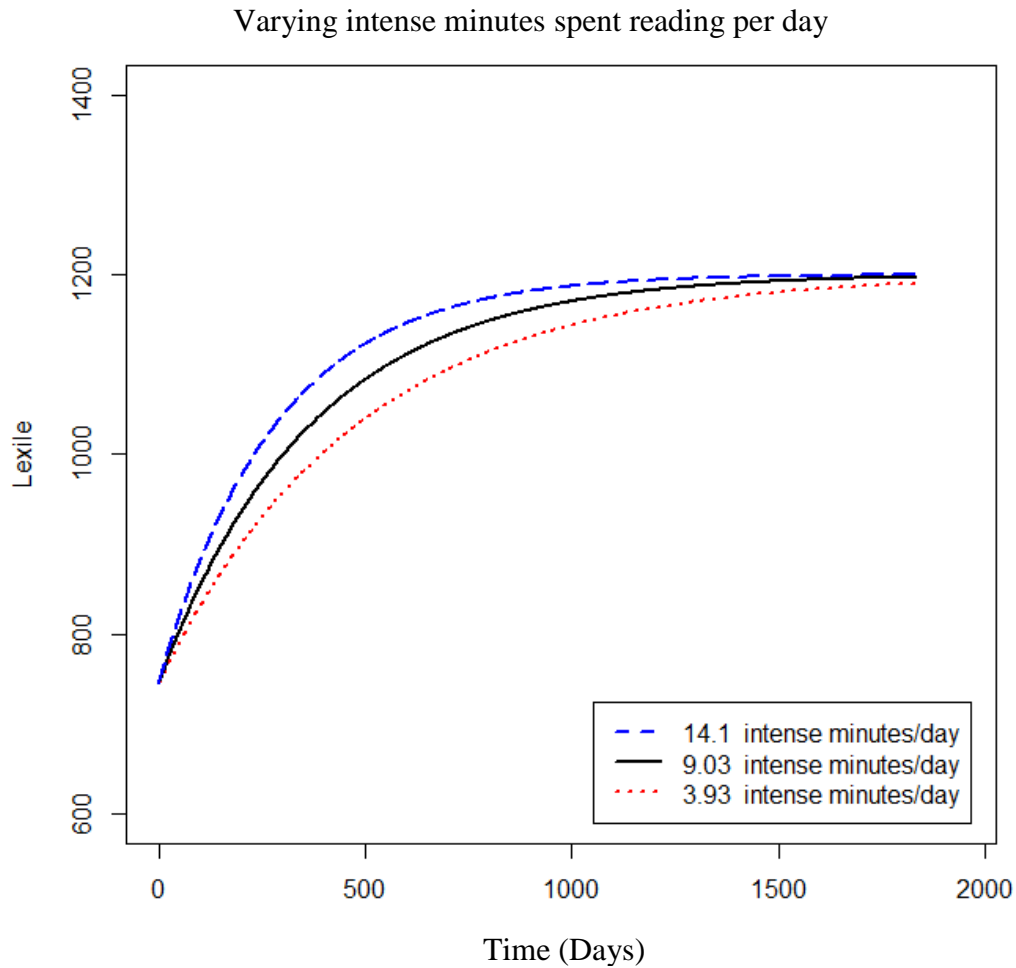


Figure 9. Plot of three prototypical participants, varying only intense minutes spent reading per day.

H5-a: The standard deviation of elapsed days between reading experiences will negatively relate to the rate of change in reading ability over time

H5-b: The standard deviation of elapsed days between reading experiences will negatively relate to the asymptote

ElapsedDaysSD represented the standard deviation of elapsed days between reading experiences. Participants with small *ElapsedDaysSD* values had their days of reading spaced

more evenly than participants with larger *ElapsedDaysSD*. *ElapsedDaysSD* was not a statistically significant predictor of the rate parameter, suggesting it did not contribute to model fit. Therefore, hypothesis H5-a was not supported.

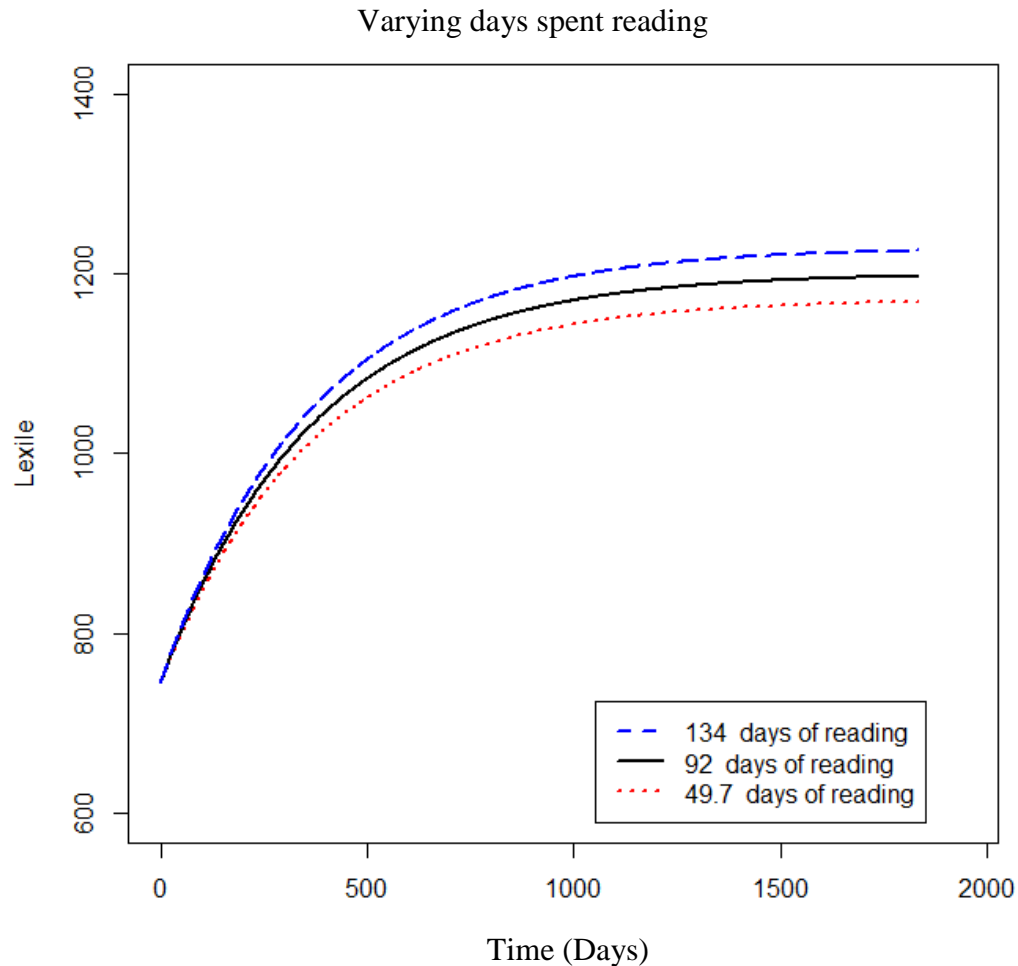


Figure 10. Plot of three prototypical participants, varying only number of days of reading.

ElapsedDaysSD was a statistically significant predictor of the asymptote parameter. The positive sign of the fixed effect suggested that participants who had their reading practice more unevenly spaced experienced a higher asymptote than participants with more evenly spaced practice. Visual inspection of the influence of *ElapsedDaysSD* confirmed this

finding (Figure 11). Therefore, hypothesis H5-b was not supported and the result was unexpected.

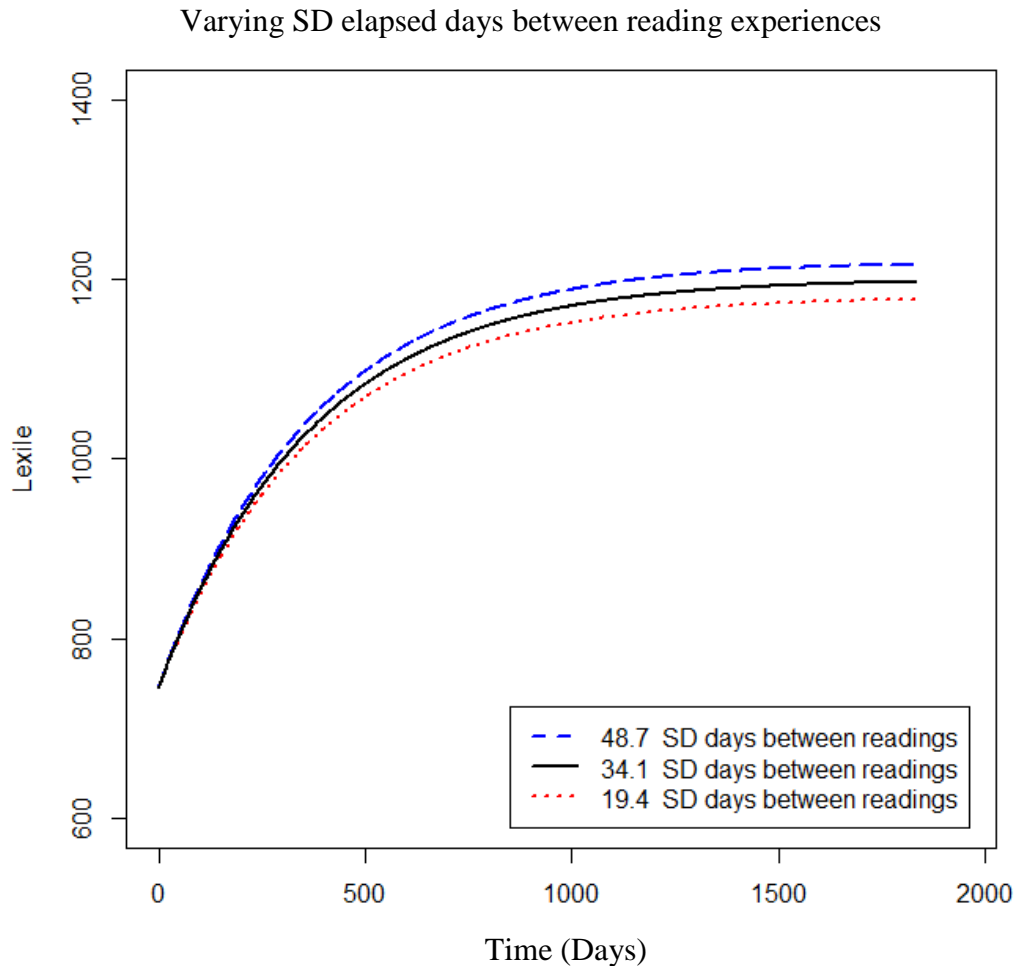


Figure 11. Plot of three prototypical participants, varying only standard deviation of elapsed days between reading experiences

H6-a: The initial status in reading ability over time will be positively related to the *initial reading ability* of the participant

H6-b: The rate of change in reading ability over time will be negatively related to the *initial reading ability* of the participant

H6-c: The asymptote will be positively related to the *initial reading ability* of the participant. Participants who start at a higher measure will finish at a higher measure

InitialLexile represented the estimated reading ability of participants when they first started using *Learning Oasis*. *InitialLexile* was a statistically significant predictor of the initial status, and the positive sign of the fixed effect suggested that participants who had a higher *InitialLexile* had a higher initial status. Therefore, hypothesis H6-a was retained.

InitialLexile was a statistically non-significant predictor of the rate parameter, and was removed from the final model during simplification. Hypothesis H6-b was not supported.

InitialLexile was a statistically significant predictor of the asymptote term. The sign of the fixed effect suggested that participants with a higher *InitialLexile* experienced a higher asymptote, as evident in the graphical representation (Figure 12). Hypothesis H6-c was retained.

H7-a: The initial status in reading ability over time will be positively related to the *initial grade of the participant*

H7-b: The rate of change in reading ability over time will be negatively related to the *initial grade of the participant*

H7-c: The asymptote will be positively related to the *initial grade of the participant*

InitialGrade reflected the grade of a participant when *Learning Oasis* was first accessed. *InitialGrade* was a statistically significant predictor of initial status, rate parameter, and asymptote in the full model. However, *InitialGrade* was removed from the model to reduce the potential influence of multicollinearity. Hypotheses H7-a, H7-b, H7-c could not be examined using the data and model for this study.

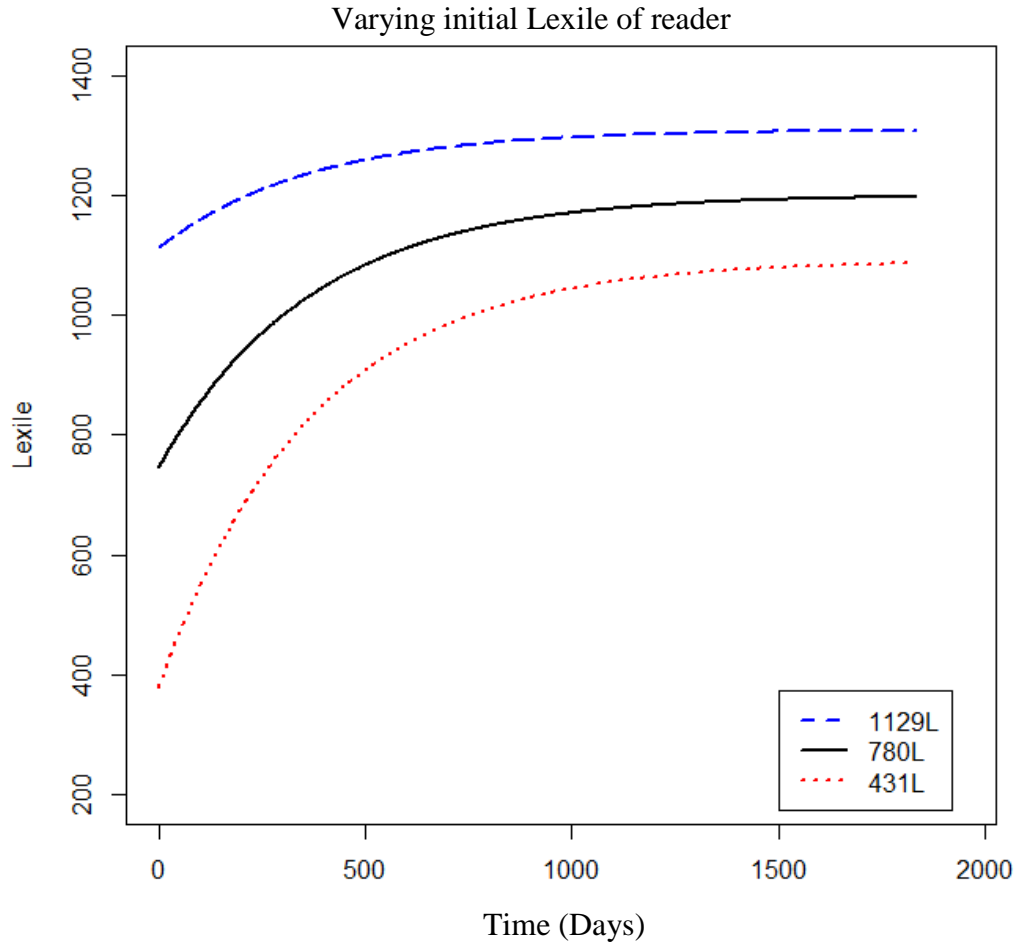


Figure 12. Plot of three prototypical participants, varying only initial Lexile reading level.

H8-a: The socio-economic status (SES) of a participant will be related to initial reading ability. Participants with lower SES, as measured by receiving free lunch, will have a lower initial intercept than participants with higher SES, as measured by paying for lunch.

H8-b: The SES of a participant will be related to rate of change in reading ability over time. As predicted by the reading literature, participants with lower SES will have lower rates of change than participants with higher SES.

H8-c: The SES of a participant will be related to the asymptote. Participants with lower SES will have a lower asymptote than participants with higher SES.

SES was represented by the lunch status of a participant (i.e., received free lunch, or paid for lunch). *SES* was a statistically non-significant predictor of initial status and it was removed from the model during simplification. Hypothesis H8-a was not supported.

SES was a statistically significant predictor of both the rate parameter and asymptote. The sign of the fixed effects suggested that participants that received free lunch grew more slowly and had a lower asymptote than participants that paid for lunch. Visual inspection of the influence of *SES* supported this finding (Figure 13). Hypotheses H8-b and H8-c were retained.

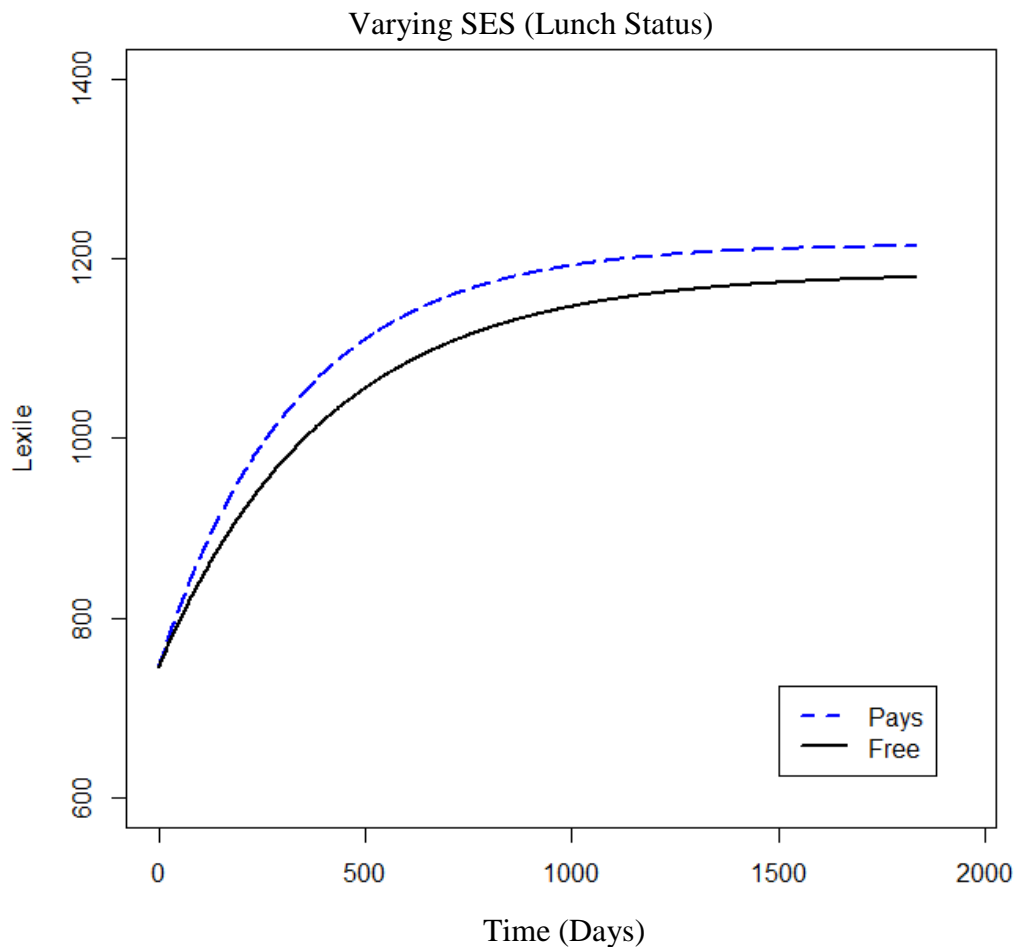


Figure 13. Plot of two prototypical participants, varying lunch status.

Deliberate Practice as a Whole

To examine how the statistically significant predictors of deliberate practice contributed to change in reading ability as a group, three graphs were created with three prototypical participants each. On each graph, all three prototypical participants had the same *InitialLexile* and *SES* and varied by the amount of *Words*, *IntenseMinPerDay*, *DaysUsed*, and *ElapsedDaysSD*. The mean (i.e., solid line) represented a prototypical participant who completed the mean amount of deliberate practice, denoted by the mean of *Words*, *IntenseMinPerDay*, *DaysUsed*, and *ElapsedDaysSD* respectively. The “low” prototypical participant (i.e., dotted line) represented a participant one standard deviation below the mean on all statistically significant deliberate practice predictors. The “high” prototypical participant represented a participant one standard deviation above the mean on all statistically significant deliberate practice predictors. While the fixed effect for *ElapsedDaysSD* was opposite the hypothesized effect, it remained in the model as a part of the package of deliberate practice as it was statistically significant in the final model. Each graph depicted prototypical participants with different *InitialLexile* estimates: 431L (Figure 14), 780L (Figure 15), and 1129L (Figure 16).

Examination of Assumptions

The assumptions of functional form, normality, and homoscedasticity were examined using graphical means. Functional form was investigated by creating empirical plots of random participants with negative exponential fits superimposed and examining the fit curves for all participants (Appendix A). A negative exponential model appeared to be a reasonable fit to the sampled participants.

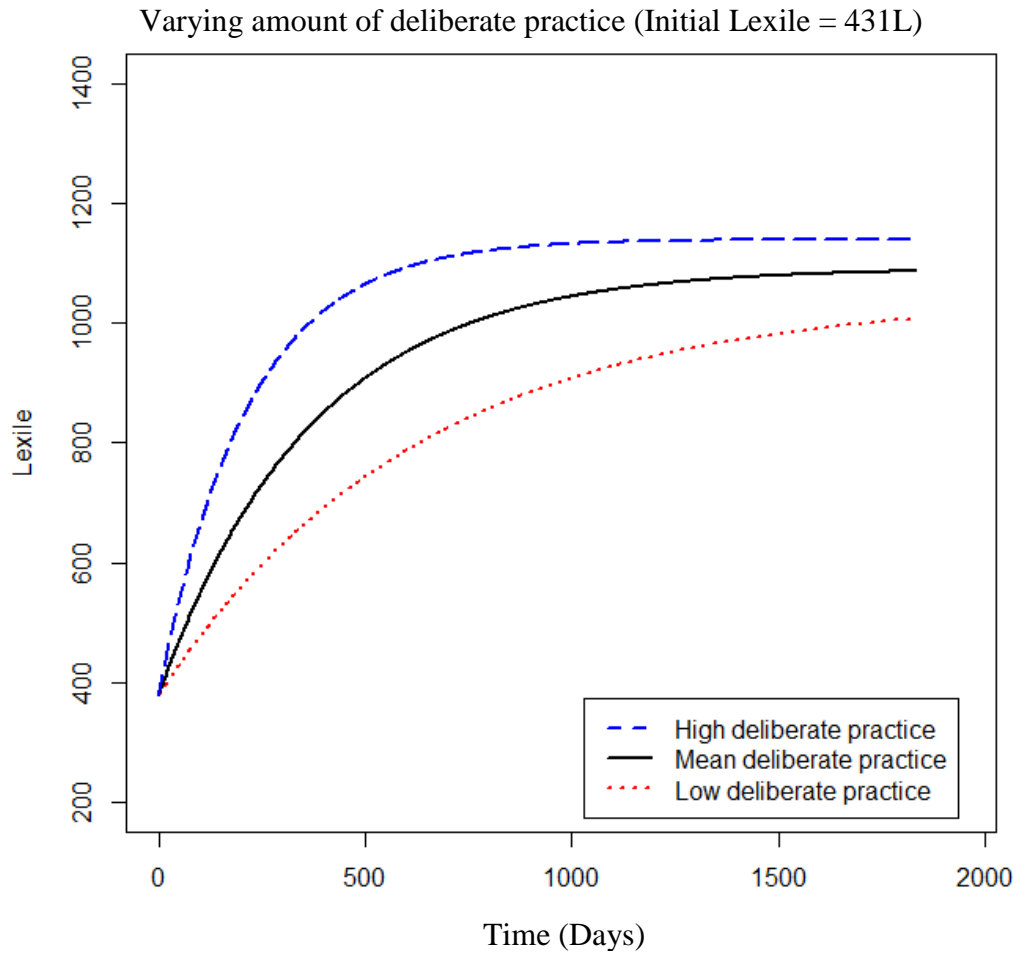


Figure 13. Plot of three prototypical participants with *InitialLexile* of 431L and varying in amount *Words*, *IntenseMinPerDay*, *DaysUsed*, and *ElapsedDaysSD*.

Varying amount of deliberate practice (Initial Lexile = 780L)

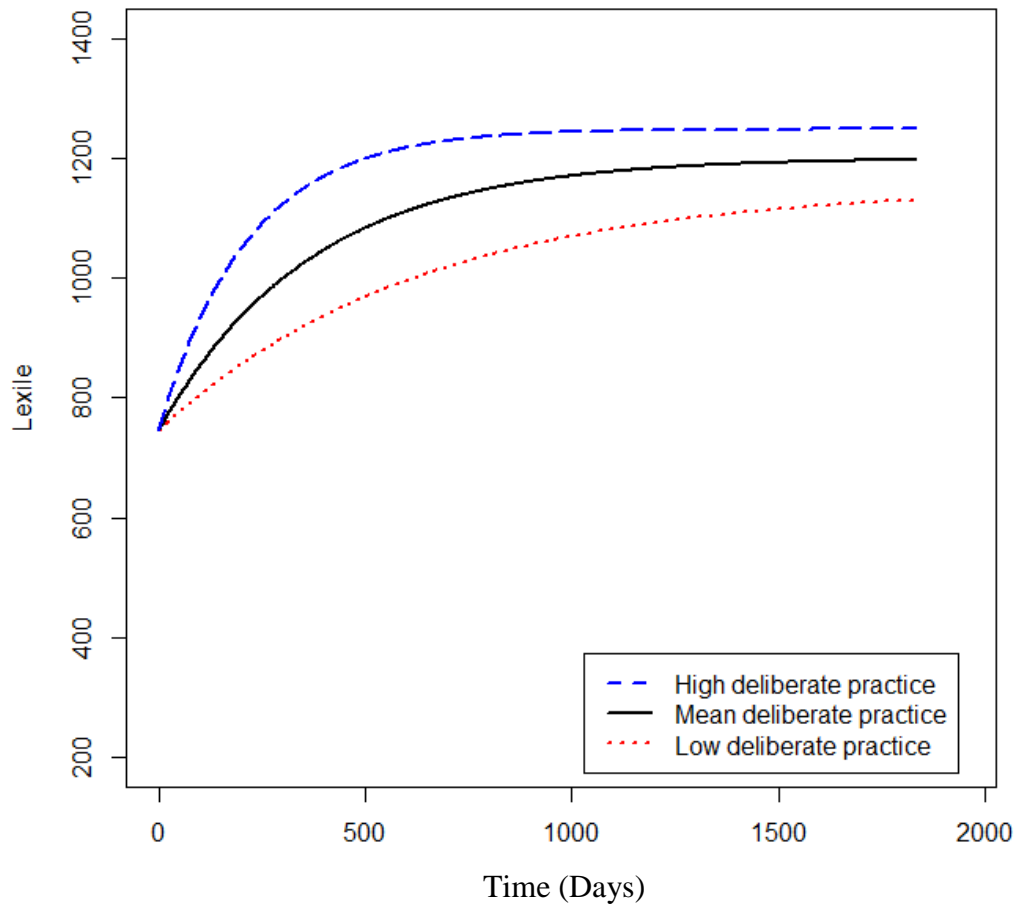


Figure 14. Plot of three prototypical participants with *InitialLexile* of 780L and varying in amount *Words*, *IntenseMinPerDay*, *DaysUsed*, and *ElapsedDaysSD*

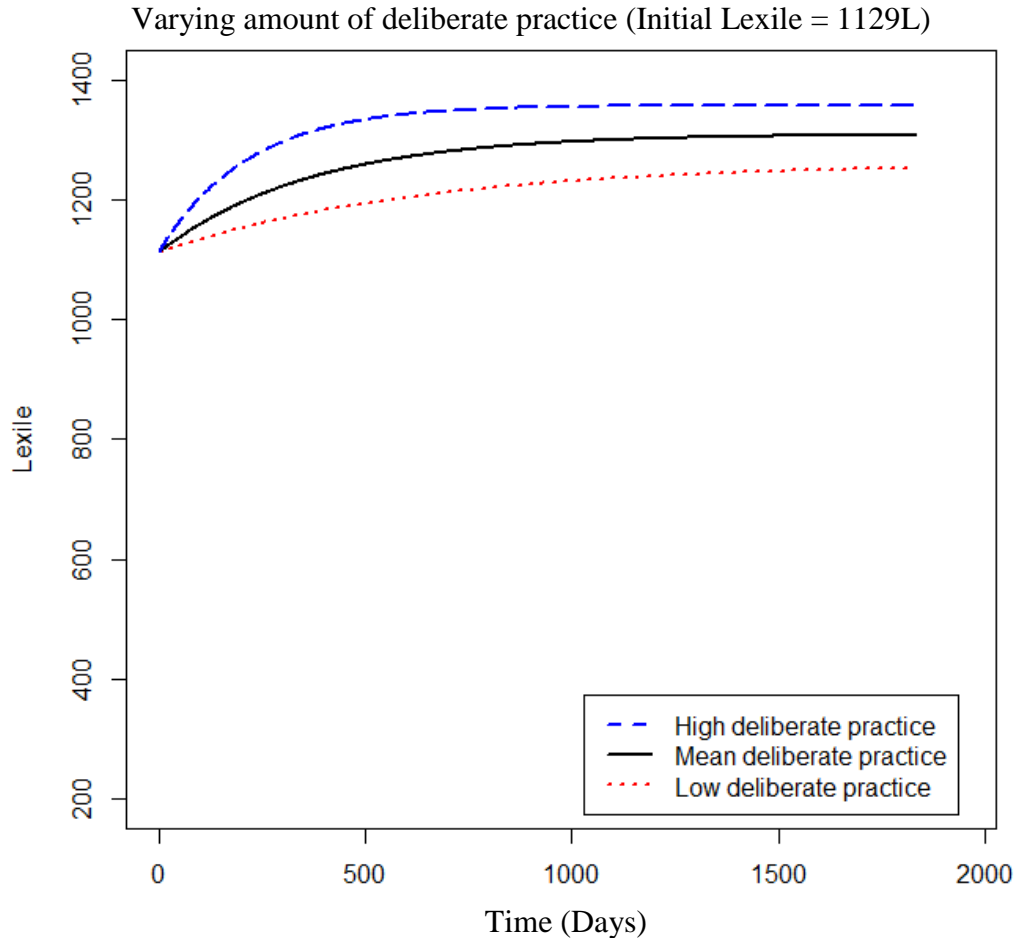


Figure 15. Plot of three prototypical participants with *InitialLexile* of 1129L and varying in amount *Words*, *IntenseMinPerDay*, *DaysUsed*, and *ElapsedDaysSD*

It was assumed that all residuals at both level-1 and level-2 were normally distributed. To analyze the tenability of this assumption, normal probability plots for each raw residual (i.e., one at level-1 and three at level-2) were plotted (Appendix C). Departures from linearity would have suggested a violation of the normality assumption; however, all four plots looked roughly linear, suggesting the assumption of normality was not violated. Plots of standardized residuals by participant identifier were also created to identify extreme cases (Appendix C). As the vast majority of points were located within 2 standard deviations of center, the assumptions of normality appeared reasonable.

It was assumed that there were equal variances of the level-1 and level-2 residuals at each level for every predictor. This assumption of homoscedasticity was examined by plotting raw residuals against predictors at each level (i.e., level-1 residuals against the level-1 predictor, level-2 residuals against each level-2 predictor). These plots were created for *InitialLexile* at level-1 (Appendix C), and each level-2 predictor: *Words*, *DaysUsed*, *IntenseMinPerDay*, *ElapsedDaysSD*, *InitialLexile*, and *SES* (Appendix C). Based on visual inspection of these plots, it was determined that homoscedasticity was a reasonable assumption, with only two level-2 predictors possibly showing slight heteroscedasticity (i.e., *Words*, *IntenseMinPerDay*). The assumptions of functional form, normality, and homoscedasticity appeared reasonable.

Summary of Results

The hypotheses for the present study were expanded to include hypotheses relating to the prediction of the asymptote by predictors-of-interest. The outcomes of the hypotheses related to the rate parameter were mixed. As expected, *Words* (H1-a), *IntenseMinPerDay* (H3-a), and *SES* (H8-b) were all positively related to the rate parameter. In contrast, *DaysUsed* (H4-a), *ElapsedDaysSD* (H5-a), and *InitialLexile* (H6-a) did not appear to be related to the rate parameter. *DaysUsed* (H4-b), *InitialLexile* (H6-b), and *SES* (H8-c) were shown to positively relate to the asymptotic term, as predicted. However, *Words* (H1-b) and *IntenseMinPerDay* (H3-b) were not related to the asymptote and the influence of *ElapsedDaysSD* (H5-b) was opposite of the hypothesis. Hypotheses related to *Items* (H2) and *InitialGrade* (H7) were explored, but potential collinearity led to the variables being dropped from the final model. This made it difficult to interpret their unique effects.

CHAPTER 5

DISCUSSION

Many students are not prepared for life after high school as their reading ability is insufficient to meet the literacy demands of college and career (Common Core State Standards, 2010; MetaMetrics, 2008; Williamson, 2008; Wirt et al., 2004). To help students succeed in the post-secondary world, the gap between participants' reading ability at the conclusion of high school and the complexity of the text encountered after graduation must be closed (Common Core State Standards, 2010; Williamson, 2008). The adage "practice makes perfect" has been a popular perspective as both researchers and educators have supported the notion that increased reading volume (i.e., words read) leads to improvement in reading ability (Allington, 1977, 1980, 1983, 1984a; Anderson et. al., 1988; Cunningham & Stanovich, 1998; Gambrell, 1984, 2007; Hiebert, 1983; Knapp, 1995; Krashen, 2004; Meyer & Wardrop, 1994; Stanovich, 2000; Stanovich et. al., 1996; Thurlow et. al., 1984; Vaughn et. al., 1998; Wu & Samuels, 2004). Despite the ubiquity of support for reading volume facilitating the development of reading ability, the National Reading Panel (NRP; National Institute of Child Health and Human Development [NICHD], 2000) questioned the importance of reading quantity, citing a lack of compelling empirical evidence.

Regardless of the opinion of the NRP, Allington (2009) believed that reading volume influenced the development of reading; however, he observed a lack of research into the

amount and type of practice necessary to nurture reading ability. Deliberate practice, proposed by Ericsson et al. (1993), consists of a series of principles that could potentially describe the type of practice necessary to enhance the development of reading. The present study was designed to explore the relationship between deliberate practice and reading ability. An educational technology, *Learning Oasis*, provided a data collection mechanism and digital learning environment that provided opportunities for participants to engage in deliberate practice.

The remainder of this chapter is divided into five sections. The first section offers a discussion of deliberate practice and how the results of the present study are related to theories of reading development. The next section describes the role technology played in this study in overcoming the methodological limitations of past reading volume research. The limitations of the study are described next, followed by a discussion of future research. Finally, implications for educators, educational technology designers, and researchers are presented.

Deliberate Practice and Reading Ability

Deliberate practice requires activity that is specifically designed to challenge the learner and improve performance. Deliberate practice can be characterized by five principles: (1) targeted activity that is designed to appropriately challenge the learner, (2) real-time corrective feedback that provides an indicator of performance, (3) distributed practice over a long period of time, (4) intensive practice that does not require the learner to concentrate beyond their limits, and (5) self-directed practice when a teacher or coach is unavailable (Ericsson, 1996a, 1996b, 2002, 2004, 2006a, 2006b; Ericsson et. al, 1993). Traditionally, practice sessions were organized and coordinated by a teacher or coach; however, for the

present study, *Learning Oasis* provided participants with opportunities to engage in deliberate practice.

The present study required a systematic way to identify and quantify the amount of deliberate practice completed by a participant. As there was no accepted procedure for representing each principle of deliberate practice, operational definitions based on the literature were created for this study. Subsequent sections present a discussion of each component of deliberate practice, as defined for the present study. For each component, the operational definition is briefly reviewed followed by a discussion of the results and how the findings are related to theories of reading. The final sections describe the influence of the control variables (i.e., initial Lexile and initial grade of reader, SES) and present commentary about deliberate practice in the context of an educational technology.

Targeted practice. For this study, targeted practice was defined as the amount of conditional reading volume. Conditional reading volume was operationalized as the number of words read inside *Learning Oasis*. *Learning Oasis* allowed readers to choose articles $\pm 100L$ around their ability, resulting in an expected comprehension between 66% and 82% (Stenner, 2003). This targeting was consistent with past reading research that used The Lexile Framework for Reading to match readers to text (e.g., Kim, 2006, 2007; Kim & White, 2008; Kim & Guryan, 2010).

The results of model fitting suggested that conditional reading volume was positively related to the rate of change in participant reading ability. The magnitude of the fixed effect was relatively small (i.e., $2.38e-06$) as the parameter was measured in words which varied from 9,173 to 625,279 ($M = 151,574$, $SD = 113,117$). Interpretation was challenging due to the rate parameter appearing in the exponential term; however, visual inspection of the graph

depicting the influence of targeted words read supported the finding that participants who read more words grew at a faster rate than participants who read fewer words (Figure 7). This finding was consistent with connectionist theories of reading (Adams, 1994; Rumelhart & McClelland, 1986a; Seidenberg & McClelland, 1989). From a connectionist perspective, exposure to appropriately challenging words strengthens areas of the neural network that are less-well developed. This finding was also consistent with the avalanche of literature supporting the notion that reading volume is a critical component to the developing reader (Allington, 1977, 1980, 1983, 1984a; Anderson, et. al., 1988; Cunningham & Stanovich, 1998; Gambrell, 1984, 2007; Hiebert, 1983; Knapp, 1995; Krashen, 2004; Meyer & Wardrop, 1994; Stanovich, 2000; Stanovich, et. al., 1996; Thurlow, et. al., 1984; Vaughn, et. al., 1998; Wu & Samuels, 2004).

While conditional reading volume was a statistically significant predictor of how quickly participants grew, it did not predict the asymptote. This suggested that the number of words read was not a contributing factor to the maximum reading level achieved by a participant. Said differently, two participants could read different numbers of targeted words, and still reach the same peak level; participants that read fewer words would just take longer to reach the asymptote. These findings suggested that conditional reading volume alone was insufficient to model participant change over time.

Real-time corrective feedback. Real-time corrective feedback was quantified by the number of automatically-generated semantic cloze items completed by a reader. Every reading experience inside *Learning Oasis* contained cloze items and each item was accompanied by color-based feedback (i.e., green or red) that was determined by the

accuracy of the participant response (i.e., correct or incorrect). The number of items in a particular passage was a function of passage length (i.e., longer passages had more items).

While the variable representing number of items was a statistically significant predictor of both the rate parameter and asymptote, it was removed from the final model due to potential multicollinearity. Given the relationship between number of words and items, this potential collinearity was not unexpected. Unfortunately, number of words read and items taken were irrevocably confounded as all reading experiences in *Learning Oasis* contained items and the number of items was largely a function of passage length. The collinearity between these variables prevented a clear interpretation as to which component was the more influential predictor in predicting a participant growth trajectory.

While the predictor of real-time corrective feedback was removed from the model, participants reading inside *Learning Oasis* received immediate feedback while reading. A participant received an item approximately every 65 words. So, while the potential influence of immediate feedback was not explicitly represented in the final model and therefore evidence could not be gathered regarding this hypothesis, it remains possible that the feedback a participant receives positively affects change in reading ability. This influence would be consistent with existing research into the relationship between feedback and learning (Epstein, et. al., 2001; Epstein, et al., 2002; Epstein & Brosvic, 2002) and connectionist theories of reading in which feedback provided critical information that informed the formation and strengthening of the neural network of the reader (Adams, 1994; Rumelhart & McClelland, 1986a; Seidenberg & McClelland, 1989).

Intensive practice. The amount of intensive practice was represented by the number of minutes spent reading per day that fell within the acceptable reading rate range. The

acceptable range, measured in standard-length words per minute (Wpm), was based on Carver's (1972, 1992b, 2000) theories about the gears of reading. For this study, participants reading in the *learning gear* and the *reading gear* were considered to be intensely reading. Participant reading experiences that were too slow (i.e., *memorization gear*), or too fast (i.e., *scanning gear* or *skimming gear*) were not included in the intense minutes per day computation. A critical property of measuring reading rate using Wpm was that the complexity of the text would not influence reading rate as the number of words was standardized (Carver, 1983, 1990).

Results of model fitting suggested that intense daily activity was a statistically significant predictor of the rate parameter. The magnitude of the fixed effect, 0.052, is represented in the unit minutes per day. The number of intense minutes spent reading per day varied from 2.04 to 33.64 ($M = 9.03$, $SD = 5.09$) and participants that read more intensely each day showed greater growth than participants that read less intensely (Figure 8). This finding was consistent with LaBerge and Samuels' (1974) conceptualization of internal attention, which required participants to focus on a reading task to improve ability.

While the number of minutes of intense reading per day was a statistically significant predictor of how quickly participants grew, it did not predict the asymptote. This suggested that the number of intense minutes of reading per day was not a contributing factor to the maximum reading level achieved. Two participants could read intensely for different numbers of minutes each day, and still reach the same peak level. It would just take participants that read for fewer intense minutes longer to reach the asymptote than participants that read more intensely per day.

Distributed practice. Distributed practice was represented using two values. First, the number of days of reading represented how many days a participant read at least one passage. The second representation of distributed practice was the standard deviation of days elapsed between reading experiences. This second quantity was designed to represent how equally days of reading were spaced. Smaller deviations would reflect more equal spacing than larger deviations.

After fitting the multilevel model, it was determined that number of days reading was a statistically significant predictor of the asymptote; however, it was a statistically non-significant predictor of the rate parameter. The magnitude of the fixed effect, 0.673, is represented in the unit days of reading; the number of days of reading varied from 7 to 284 ($M = 92.01$, $SD = 42.33$). This finding suggested that participants that read on more days reached a higher reading level (i.e., asymptote) than participants that read on fewer days (Figure 9).

The positive relationship between days spent reading and growth was consistent with the literature on spacing effects of practice and connectionist theories of reading. Increased frequency of practice over time has been shown to lead to more learning than repetitions that occurred in short succession (Dempster, 1988; Toppino & Schneider, 1999). It has been theorized that separating practice temporally provides the time necessary for expert knowledge structures to develop (Chi et. al., 1989; Lesgold et al., 1988; Schooler, 1990). From a connectionist reading theorist perspective, spacing exposure to text over an extended period of time provides the neural network of a reader the necessary time to form and strengthen (Adams, 1994; Rumelhart & McClelland, 1986a; Seidenberg & McClelland, 1989).

The second indicator of distributed practice, standard deviation of elapsed days between reading experiences, was a statistically significant predictor of the asymptote and a statistically non-significant predictor of the rate parameter. The fixed effect was measured in the unit standard deviation of days between reading encounters. The range of standard deviations between days of reading varied between 8.4 and 149.0 ($M = 34.1$, $SD = 14.6$). The magnitude of the fixed effect, 1.341, suggested that reading less-evenly over time led to a higher maximum reading level (i.e., asymptote).

The positive effect of the second distributed practice predictor, standard deviation of days between reading experience, was unexpected, contrary to the result suggested by the expertise literature. There are at least two possible explanations for the lack of expected contribution. First, visual inspection of the graph depicting the influence of standard deviation between days showed that the predictor had a relatively small influence on the asymptote (Figure 10). This suggested that the distribution of reading experiences over time might not influence the asymptote in actuality, and that the statistically significant findings were a result of artifacts specific to the data in this study. Second, the operational definition designed to capture “how spread out” the reading encounters were could have been ill-conceived. Given the results, it was possible that using the standard deviation of days elapsed between reading encounters was not a reasonable representation of the spacing between reading experiences.

Self-directed. All activity completed inside *Learning Oasis* was driven by the learner. Participants chose when to read (i.e., 24 hours per day, 365 days per year) and what to read (e.g. searches based on keywords, categories, publication, passage length). Given that participants were never immersed in non-self-directed learning, the final model did not

contain a predictor for the self-directed nature of the reading practice. While the effect of self-directed practice versus other types of practice (e.g., practice selected by a third party a priori) could not be assessed, participants in this study were required to self-direct learning to use *Learning Oasis*. Therefore, the practice completed by participants was aligned with the principles of deliberate practice.

Control predictors. The original model included three control predictors. Initial grade, initial Lexile, and SES (i.e., lunch status) were all hypothesized to influence the initial status, rate of change, and asymptote. While initial grade was a statistically significant predictor of initial status, the rate parameter, and asymptote, it was removed from the model due to potential multicollinearity. Given that participants who are in higher grades typically possess more initial ability, grow less rapidly, and reach a higher reading level than participants in lower grades, as evidenced by grade-based metrics (Nelson, et. al., 2011; Francis, et. al., 1996; Stenner, 2003; Stenner, et. al., 2006, 2007; Stenner & Stone, 2004), this potential multicollinearity was not surprising.

Initial Lexile was a statistically significant predictor of both initial status and asymptote. Participants that started at a higher Lexile measure had a higher initial status and higher asymptote. While initial Lexile was not a statistically significant predictor of the rate parameter, examination of the sign of the non-significant parameter estimate suggested that participants with a higher initial Lexile grew less rapidly. The influence of initial Lexile on participant growth was best exhibited visually (Figure 11). Visual inspection of prototypical plots showed that participants who started at a lower initial Lexile level started at a lower initial status, appeared to grow more rapidly, and reached a lower asymptote than participants that started at a higher initial Lexile. These findings were consistent with both the existing

literature on growth in reading ability (Francis, et. al., 1996; Stenner, et. al, 2006, 2007; Williamson, et. al., 1991) and the Matthew Effect (Stanovich, 1986, 2000). That is, early achievement in reading, as represented by initial Lexile in this research, traditionally leads to quicker and easier acquisition of reading ability later.

SES was a statistically significant predictor of the rate parameter and asymptote. Participants that paid for lunch grew more rapidly and reached a higher peak than participants that received free lunch. Visual inspection of prototypical plots supported this finding, highlighting that on average, participants receiving free lunch did not close the gap with higher SES participants over the course of the study (Figure 12). The predictive power of SES on the rate parameter and asymptote was consistent with existing research on the relationship between SES and participant reading (Bradley et. al., 2001; Noble et. al., 2006; Fryer & Levitt, 2004; Rathbun & West, 2004; White, 1982; Whitehurst, 1997).

The fixed effect coefficient for SES predicting initial status, 9.80, while not statistically significant, suggested that participants with lower SES started at a lower initial status. The removal of SES as a predictor of initial status due to lack of predictive power was surprising as the literature on the influence of SES on reading suggested that participants with lower SES would have a lower initial status (Bradley et. al., 2001; Noble et. al., 2006; Fryer & Levitt, 2004; Rathbun & West, 2004; White, 1982; Whitehurst, 1997). A potential explanation for this unexpected finding is that the initial status of participants with lower SES was actually not statistically different from the initial status of participants with higher SES in this particular sample. This could occur, despite the literature-based evidence to the contrary, due to existing programs and supports in the district designed to close the gap between participants of different SES conditions.

Deliberate practice. The results of this study suggested that an educational technology could be designed to foster the development of expertise in reading through deliberate practice. While one component of deliberate practice was removed from the final model due to potential multicollinearity (i.e., items), the remaining components of deliberate practice contributed to the prediction of the rate parameter (i.e., words read, intense minutes spent reading per day) and asymptote (i.e., days used, standard deviation of elapsed days between reading experiences). Examination of the entire “package” of deliberate practice, as operationalized in this study, showed that, on average, participants that engaged in more deliberate practice grew more rapidly and reached a higher ability level than participants that completed less deliberate practice (Figures 13, 14, 15).

The Role of Educational Technology

Previous research into the effect of reading volume suffered from methodological limitations in quantifying the amount of reading completed and estimating changes in participant reading ability. To measure volume of reading, researchers have traditionally relied on self-report (e.g., questionnaires, diaries) estimates of reader activity (e.g., Ennis, 1965; Nell, 1988; Stanovich & West, 1989; Wagner & Stanovich, 1996; Walberg & Tsai, 1984). Unfortunately, these self-report measures required commitment and were susceptible to social pressure (Allington, 2009; Wagner & Stanovich, 1996). The second methodological limitation was quantifying change in reading ability over time. Researchers have often been limited to using a small number of measurement occasions, such as a pretest and posttest (e.g., Kim, 2006, 2007; Kim & Guryan, 2010; Kim & White, 2008). Such a small number of measurement occasions did not provide information about the growth trajectory (e.g., slope, intercept, functional form) and only resulted in limited estimates of change (Rogosa et. al.,

1982; Willett, 1988). The methodological limitations of past reading volume research were overcome in the present study by using an educational technology that provided ongoing records of participant activity and continuous measurement throughout the course of the study.

Learning Oasis provided a more precise estimate of reading activity than traditional self-report methods. Each reading encounter inside *Learning Oasis* was tracked and stored, providing a detailed account of reader activity. The average participant had access to *Learning Oasis* for 1,422 calendar days. This suggested that the present study was truly longitudinal in nature, extending for almost four full years. During that time, the average participant read 212.2 articles which contained 151,574 total words. This level of information was far more detailed and accurate than the traditional questionnaires or reading journals historically employed to track reading activity.

On average, participants using *Learning Oasis* answered 2,324 automatically-generated semantic cloze items. This volume of items was two orders of magnitude greater than the traditional 36-56 item high-stakes reading assessment participants in many states take once per year (Table 5). By querying participants with cloze items during each reading experience, *Learning Oasis* provided ongoing estimates of participant reading ability throughout the course of the study. The average number of measurement occasions, 22.3, provided an order-of-magnitude improvement over the traditional pre/post-test measures common in past reading volume research. Even participants with minimal *Learning Oasis* usage had at least three measurement occasions, and some participants had more than 100 measurement occasions.

While the frequency of measurement was greater than traditional reading volume research, the spacing of the measurement occasions was also more flexible than past reading volume research. Instead of the pre-defined measurement schedule of past research (e.g., at start and end of study), participants were continuously measured over the course of the study based on how often they read inside *Learning Oasis*. The average participant was in the system for 1,422 days and over that time, measurement was constantly occurring. Based on the means, the average participant was measured every 63.77 calendar days (i.e., school days weekends, school holidays, and summer break).

Limitations

Learning Oasis was a critical component to the present study; however, its architecture limited possible inferences. These limitations were largely related to the data collected and the restrictions the system placed on learners. *Learning Oasis* was not designed to vary the learning environment in ways that would allow for experimental designs; several of the key indicators of deliberate practice were hard-wired into the system, and could not be varied for this research.

Participants reading inside *Learning Oasis* all received targeted text based on their ability. Participants reading between 400L and 1400L only read text that was $\pm 100L$ around their reading ability. Participants reading below 400L received text from BR (i.e., Beginning Reader) to 500L while participants reading above 1400L engaged text that was greater than 1300L (i.e., no upper limit). These relaxed targeting requirements were necessary because *Learning Oasis* had fewer passages at the very low (e.g., < 500L) and very high (e.g., > 1300L) ends of the Lexile scale. There was no variability in the level of targeting, so all participants received text targeted in the same manner, depending on their reading ability,

throughout the course of the study. This reality precluded exploration of the impact of different levels of targeting on student change in reading ability. Moreover, the relaxed targeting for very low and very high readers suggested that, by design, these participants did not read text potentially as well-targeted as participants not reading at the extremes.

In addition to reading targeted text, all participants received automatically-generated cloze items. As the number of items was a function of article length, it was not possible to examine the influence of words read and items taken independently as the number of items completed was removed from the model due to potential multicollinearity. This was an important limitation as one of the components of deliberate practice (i.e., immediate feedback) could not be explicitly included in the model and therefore the unique effect could not be determined. The item generation protocol itself could also be a limitation. If Bormuth's (1967) assertion that removing every fifth word was ideal, the feedback received by participants inside *Learning Oasis* did not occur often enough (i.e., item every 65 words). A final limitation related to the definition of immediate feedback was that the only feedback indicator selected for inclusion was items answered. Additional types of feedback were not represented in the operationalization of immediate feedback.

The approach used to measure level of intensity was based on Carver's research on reading rate (Carver, 1983, 1992b; Rayner, 1975; Zuber & Wetzel, 1981). While this approach provided a useful conceptual framework grounded in existing research, the particular reading rates at each grade-level were based on data collected nearly 50 years ago. It was possible that these ranges did not reflect the current reality of participants reading on a computer screen. Unfortunately, contemporary silent reading norms were not available (Hiebert, Samuels, & Rasinski, 2012).

The self-directed nature of *Learning Oasis* resulted in no operational contrast in the amount or quality of self-direction. All participants were given choice in selecting when and what to read. This limitation resulted in a final model that could not contain self-direction as a predictor, and therefore could not directly evaluate the effect of this aspect of deliberate practice. While not included in the multilevel model, the choice in reading material afforded to readers could have played a critical role in influencing the amount of reading activity and student motivation to read. Research suggests that providing students with choice in what to read increases effort and commitment while reading and leads to larger amounts of reading activity (Guthrie & Wigfield, 2000; Worthy & McKool, 1996).

Learning Oasis provided a powerful deliberate practice delivery platform; however, the learning environment precluded a thorough investigation of the relative influence of each component of deliberate practice compared to the others. It was not possible to definitively assess which components of deliberate practice were more beneficial to participants. The present study provided insight into this, but the study was not designed or executed in such a way that the deliberate practice components could be disentangled.

As a supplemental reading program, *Learning Oasis* complemented existing reading instruction in the classroom (e.g., teacher-led instruction, other reading interventions); however, the influence of the classroom context on this study was not represented in the model. This limitation is important as educators integrated *Learning Oasis* into their classroom in a variety of ways. Some teachers used their classroom computers (e.g., typically three or four machines) and had students “cycle through”, taking turns reading during morning work or centers. Other educators reserved a school computer lab or laptop cart, providing an opportunity for the entire class of students to engage *Learning Oasis* at the

same time. Educators in higher grades often supplemented in-class computer lab time with assignments that required students to read outside of class. To maximize opportunities for students to access *Learning Oasis*, these educators ensured the school computer labs were open before school, during lunch and study halls, and after school. In many instances, teachers would combine approaches (e.g., using classroom computers daily, and reserving the computer lab once per week). It is possible that the frequency of participant usage of *Learning Oasis* was a function of individual educator decisions (e.g., time devoted to class-wide usage in computer labs). These instructional decisions may have greatly influenced the spacing of deliberate practice received by study participants. Neither the model of implementation nor other classroom influences (e.g., other interventions) was investigated in this research.

The literature suggests that a variety of teacher factors can influence the quality and extent of an educational technology's integration into a classroom. Teacher age, amount of teaching experience, perception of technology, familiarity with technology, and time in a particular school have all been shown to influence how a particular teacher integrates educational technology in the classroom (Cuban, 2001; O'Dwyer, Russell, & Bebell, 2006; Russell, Bebell, & O'Dwyer, 2005; Russell, Bebell, O'Dwyer, & O'Connor, 2003; Russell, O'Dwyer, Bebell, & Tao, 2007; Wood, Mueller, Willoughby, Specht, & Deyoung, 2007). Educators more facile with technology might have been more able to scaffold the use of the educational technology for students than educators with less familiarity with technology. These teacher-level effects were not considered in this study, suggesting that variability in teachers (e.g., age, experience, perception and familiarity with technology) could have influenced the findings. It is possible that teachers more able to integrate educational

technology into their classrooms provided participants with better modeling in how to use the technology and more opportunities to engage in the deliberate practice delivered by *Learning Oasis*.

The impact of the classroom environment on student motivation and learning was also not included in this study. Different goal orientations (i.e., mastery, performance) at the classroom-level have been shown to lead to differences in student-level engagement (Church, Elliot & Gable, 2001; Greene, Miller, Crowson, Duke & Akey, 2004; Meece, Anderman, & Anderman, 2006; Nolen, 2001; Nolen & Haladyna, 1990; Roeser, Midgley & Urdan, 1996). Differences in goal orientations from classroom-to-classroom could have influenced the findings of this research. Meece et. al. found that students in classrooms that emphasized mastery, understanding, and the improvement of skills were more likely to show improved motivation and learning than students in classrooms that focused on raw performance and competition for grades. It is possible that participants in classrooms with mastery-based goal orientations were more willing to engage in *Learning Oasis* activity than participants in performance-based goal oriented classrooms.

This study was conducted in a single school district in suburban Mississippi, so the generalizability of the results to other contexts could be limited. The participants either chose to use *Learning Oasis* or were asked to read by their teacher, suggesting that a self-selection bias could have influenced the findings. While the sample was representative of the entire school district, based on demographic breakdown, the demographics of the sample suggested that differences between the sample and the state and nation existed (Table 4). This study had a much larger percentage of African American participants (i.e., 40%) than the national average (i.e., 15%). In contrast, the percentage of Hispanic participants in the current study

was less than the national average (i.e., 23% nationally, 4% in this study). While the percentage of participants receiving free lunch (i.e., 49%) was less than the average in Mississippi (i.e., 62%), it was greater than the national average (i.e., 35%). Given the magnitude of these differences in ethnicity and lunch status, this research could reasonably be expected to generalize to locations with demographic profiles roughly similar to this study.

The non-experimental nature of the present study precluded claims of causality. All participants in the school district had access to *Learning Oasis*, and a control group was not created. Moreover, activity that occurred outside the technology over the course of the study was not captured or incorporated during data modeling. The potential influence of external influences on the results of this study might be seen in the statistical non-significance of SES in predicting initial status. Another example of data not included in the data modeling process was the number of words read by a participant outside *Learning Oasis*. Regardless of these limitations, the results of this study provide important insight that can inform future research.

Future Research

The present study provided an important foundation for future exploration into the relationship between deliberate practice and reading ability. This section highlights potentially beneficial avenues for additional research endeavors. While not a comprehensive list, the ideas presented in this section provide suggestions related to further exploration into deliberate practice, expansion into other potentially contributing factors, improvements in study design, and alternative modeling approaches.

Targeted practice is a pillar of deliberate practice; however, the optimal level of match between reader and text is unknown (Shanahan, 1983, 2011). Traditionally, research into the influence of targeting has focused predominantly on providing text to participants at their individual levels (e.g., Kim, 2006, 2007). Future research should explore how different levels of targeting between reader and text influence change in reading ability.

In addition to investigating the optimal level of targeting, researchers should explore how various types of feedback influence change in reading ability. Additional types of feedback could include: ongoing estimates of reader ability compared to the absolute standard of college and career readiness (Common Core State Standards, 2010) or auditory feedback (e.g., “nice job re-reading that difficult section of text”). The study environment of future research will likely impact the types of feedback explored. For example, if auditory feedback tied to specific activity is the feedback being investigated, an in-person teacher will likely provide more authentic feedback than a digital learning system.

Future research should also explore the spacing and distribution of reading activity that leads to optimal growth in reading ability. As supported by the expertise literature, the results of this study suggested that, on average, participants that read over a longer period of time reached a higher asymptote than participants that read over shorter timespans. However, in contrast to the expertise literature, the results suggested that less-even spacing of reading was associated with a higher predicted peak reading ability. Future researchers should explore the potential interaction between total time spent engaged in practice and the spacing of the practice; it is possible that the ideal spacing is dependent on the individual learner.

Self-directed practice is a critical component of deliberate practice that requires additional exploration. Past expertise research has framed self-directed practice using the

domain of inquiry (e.g., musicians rehearsing particularly challenging segments of a musical piece independent of an instructor) (Ericsson et. al., 1993). It is unclear how self-directed practice could be represented in the reading domain. Potential factors that could encapsulate self-directed learning might include providing a reader with a choice in when to read as well as what to read. Research into the influence of self-direction will likely be a function of the particular environment researched, as classroom settings offer different opportunities for self-direction than digital environments.

While the focus of the present study was the relationship between deliberate practice and reading ability, future research should explore potentially-predictive factors not included in this research. The reading-writing connection could be investigated (i.e., the relationship between writing and growth in reading ability) (Biancarosa & Snow, 2006; Graham & Perin, 2007). Alternatively, the influence of self-regulated learning (Zimmerman, 1990) on growth in reading ability could be explored. The impact of learning scaffolds (e.g., text-to-speech, strategy instruction) might also be an interesting area of inquiry.

Future researchers should also investigate the role of motivation in impacting student reading ability. As reading requires effortful activity, the motivation of a reader merits attention; according to Park (2011), reading motivation predicts reading performance beyond other literacy-related variables (e.g., reading outside school, SES, gender). Past research has explored the relationship between reading and motivation using a variety of theoretical perspectives. Research into reader attitude has shown that students with more positive attitudes towards reading were more motivated to read than students with less positive attitudes (Alexander & Filler, 1976; Mathewson, 1994; McKenna, Kear, & Ellsworth, 1995). Researchers using an engagement perspective integrate the cognitive, motivational, and

social aspects of reading, and have discovered that engaged readers are motivated to read for different purposes (e.g., for enjoyment) than less-engaged readers, use knowledge gained from previous reading, and participate in social interaction around reading (Baker, Afflerbach, & Reinking, 1996; Guthrie & Alvermann, 1999; Guthrie, McGough, Bennett, & Rice, 1996; Oldfather & Wigfield, 1996). Lastly, researchers in the achievement motivation field have argued that the efficacy beliefs, intrinsic and extrinsic motivation, and purposes for achievement play a central role in impacting reading activity (Bandura, 1997; Eccles, Wigfield, & Schiefele, 1998; Pintrich & Schunk, 1996; Wigfield, Eccles, & Rodriguez, 1998).

Future research will need a method to quantify the motivation of readers. A variety of instruments have been constructed to measure the various facets of motivation that could impact reading (e.g., self-efficacy, self-concept, attitudes towards reading). The Motivation for Reading Questionnaire (Wigfield & Guthrie, 1997), Motivation to Read Profile (Gambrell, Palmer, Codling, & Mazzoni, 1996), as well as the efforts of Chapman and Tunmer (1995) and McKenna et al., (1995) could all be used in future research into the motivation of readers. As self-report measures, these instruments could suffer from limitations similar to self-report measures of student reading activity (e.g., social pressure, overestimates of activity); however, Wigfield and Guthrie (1997) argued that students typically respond honestly.

Future exploration into the relationship between reading and motivation should investigate potential differences between readers (e.g., gender, race, developmental level). Previous research found that, compared to boys, girls typically possess stronger feelings of competence and value reading more (Baker & Wigfield, 1999; Eccles, Wigfield, Harold, &

Blumenfeld, 1993; Marsh, 1989; Wigfield, et al., 1997). Girls also showed higher self-efficacy in reading than boys (Wigfield & Guthrie, 1997). Past research suggests that differences in ethnicity also impact reader motivation: African American students exhibit more positive motivation related to reading than white students (Baker & Wigfield, 1999; McKenna et al., 1995; Stevenson, Chen, & Uttal, 1990). Lastly, developmental differences between readers should be explored. Previous research suggests that student motivation is negatively related to grade (i.e., motivation becomes less positive as students reach higher grades) (Eccles et al., 1998; Wigfield et al., 1998; Wigfield & Guthrie, 1997). Future research should not only investigate these differences (e.g., gender, race, developmental level), but the interaction between them.

Future research into the relationship between reading and motivation should also explore potential interactions with educational technology. Technology provides a powerful learning platform that can be tailored to individual student needs and it is possible that student motivation is impacted by the use of an educational technology. While past research suggests that technology can positively influence motivation (e.g., Lumley, 1991; Page, 2002; Swan, van't Hooft, Kratcoski & Unger, 2005), the impact of educational technology on motivation when receiving deliberate practice is not well understood. For example, it is possible that allowing readers to choose text that is high-interest and at their targeted reading level overcomes the traditionally observed gender or developmental differences in motivation. The potential interaction between reading motivation, deliberate practice, and educational technology could prove a fruitful area of future research.

If future research were to be powered by an educational technology, a factorial study design (Chakraborty, Collins, Strecher, & Murphy, 2009; Collins, Dziak, & Li, 2009; Dziak,

Nahum-Shani, Collins, 2012) could be incorporated directly into the technology, allowing assignment of participants to treatment conditions at the participant-level. Participants in a particular class could all experience a personalized learning environment, tailored to the specific treatment or comparison condition to which they were assigned. This path of inquiry could prove fruitful in identifying the ideal deliberate practice conditions as participants within the same class could be assigned to different experimental conditions, potentially controlling for any teacher effects. Another important outcome to this line of inquiry would be an understanding of the treatment package that leads to optimal growth in reading ability.

The present study explored three functional forms and treated the predictors of deliberate practice as time-invariant. Future research should explore alternative functional forms (e.g., discontinuous growth) and treat predictors of interest as time-variant.

Educational technologies, such as *Learning Oasis*, are critical for this type of research as the data demands can be large and require data collection on a variable schedule over a long period of time. Conducting research using an educational technology in the contemporary classroom poses unique challenges; the next section explores some of the implications for educators, educational technology designers, and researchers.

Implications for Educators, Educational Technology Designers, and Researchers

The results of this study provide useful information for educators, educational technology designers, and researchers. The practical utility of these implications varies depending on the adopted perspective (i.e., educator, technology designer, researcher); however, the foundational insight is that the principles of deliberate practice can facilitate the development of reading ability. The most important implications for each perspective are presented.

For educators, the influence of deliberate practice on reading ability as suggested by this study provides useful guidelines that can help supplement existing classroom instruction. First, exposing students to targeted text will facilitate growth in reading ability. This targeting can be accomplished in a variety of ways including: dividing students into reading groups based on ability and level of need, organizing the classroom library into “book bins” based on difficulty, teaching students how to use educational technology to find targeted text, or application of the “five finger rule”. The “five finger rule” is a targeting heuristic where students read the first page of a book, and count the number of words they do not know. If there are five or more words on the first page unknown to the reader, the book is considered too difficult (Boushey & Moser, 2006; Diller, 2003). In most classrooms, the manner of targeting will likely be a function of class-size, available tools, and range in reading ability of the students in the class.

The results of this study suggested that participants that read more intensely each day and read more often experienced greater growth than participants that read less often. Educators can help students improve as readers by providing the time necessary for students to read intensely over an extended period of days. Instructional approaches such as Sustained Silent Reading (SSR) and Drop Everything and Read (D.E.A.R.) have been used by classroom educators to encourage students to read on a daily basis; however, not all students understand how to read intensely (Adams, 1990; Anderson, 2000). Encouraging intense reading could be facilitated by teaching students critical reading strategies (e.g., remaining focused, how to apply repair strategies, thinking like a writer) (Adams, 1990) and monitoring activity during reading time.

Technology designed to supplement classroom instruction and assessment is more than the random combination of state-of-the-art technologies. As outlined in the NETP (2010), educational technology must be learner-centered, theory-based, and relevant. While the design of the technology (e.g., look-and-feel, user interaction paradigm, system architecture, underlying infrastructure) is critical to ensuring that students are engaged and willing to repeatedly use the technology, the system must be designed according to established educational theories. The results of this study suggested that the features of deliberate practice positively influenced change in reading ability over time. Given these findings, educational technology designers should consider incorporating the principles of deliberate practice into future educational technologies instead of relying on intuition, market trends, and the most cutting-edge technologies (Mayer, 2001, 2005). As evidenced by this study, the principles of deliberate practice offer the targeting, feedback, and self-direction necessary for students to remain engaged with an educational technology.

Conducting research using an educational technology offers unique challenges for researchers. A well-designed study, engineered into a bug-free technology is only the first obstacle to the successful administration of technology-enabled classroom-based research. The real challenge is ensuring sufficient implementation in the classroom. While technology can control key experimental conditions (e.g., type of feedback, level of text available), the most important factor lies outside the scope of the technology: available time. If a study is designed, for example, to have students engage an educational technology for 30 minutes per day, 3 times per week, treatment fidelity could be compromised if educators are unable or unwilling to accommodate the timing requirements. The results of this study showed that framing the technology as a tool (i.e., not a replacement for the educator), and letting

educators find the best models of implementation (e.g., used as part of seat-time work, weekly quiz grades, as homework), resulted in long-term usage across grades. Studies that require less longitudinal data could potentially have more stringent timing requirements as educators might be willing to accommodate stricter timing requirements if the duration is short. However, conducting research using technology requires a different manner of thinking as the researcher often gives up operational control when the technology goes to schools.

The final and most critical component to the successful deployment of educational technology is the educator. Personalized learning systems, such as *Learning Oasis*, can be considered *disruptive technologies* because they have the potential to fundamentally alter the educational landscape (Christensen, 2011). These technologies can extend the school day, increasing opportunities for learning beyond traditional school hours (Farbman & Kaplan, 2005; Pennington, 2006; Silva, 2007); however, it is important for educators to embrace the use of educational technology. Over the course of this study, educators integrated *Learning Oasis* into their classroom in a variety of ways. Some embraced the system, constantly monitoring student usage and performance. Others showed little interest in tracking student activity and were content to ensure their students had access to the technology. Without educator acceptance, the influence of educational technology could be severely limited.

Next Steps

The results of this study provided compelling evidence that an educational technology designed according to the principles of deliberate practice, such as *Learning Oasis*, could play a role in influencing student reading ability. Given these findings and the limitations of the present study, two paths should be explored moving forward: classroom-based scale-up

and expansion of the research. These directions are not mutually exclusive, and should be pursued in parallel.

The first path, classroom-based scale-up, involves trying to increase usage of educational technologies designed according to the principles of deliberate practice (Ericsson, 1996a, 2002, 2004, 2006b; Ericsson et al., 1993). As evidenced by this research, aspects of deliberate practice could be uniquely delivered to students via an educational technology (e.g., text targeted to the specific reading ability of students, access to the learning environment over extended periods of time). *Learning Oasis* provided participants with an environment designed to meet the challenges issued by the NETP (2010). That is, providing learners with a relevant and personalized learning experience. Participants were immersed in content-area non-fiction text and were able to choose what text to engage. *Learning Oasis* was designed to provide a blended experience where assessment was interwoven and embedded with instruction; the findings of this study support the existing literature that suggests this blending can improve student learning (Black & Wiliam, 1998; Butler, 2010).

While certain components of deliberate practice (e.g., immediate feedback) could not be explicitly modeled due to nuances in the present study, the results suggested that participants that engaged in more deliberate practice, as defined by this study, grew more rapidly and reached higher peak ability than participants who completed less deliberate practice. Causal claims relating deliberate practice to reading ability could not be made based on this research. The evidence does suggest that deliberate practice delivered via educational technology could contribute to increased reading ability. To delay the scale-up of such

educational technologies until a stronger case for causality can be made would result in withholding an educational intervention that research suggests may benefit students.

Successful scale-up of an educational technology requires more than a research-based learner-centered technology. Deploying and supporting a student-focused technology requires that the educational technology be scalable and that the support accompanying the deployment be robust. Developing an educational technology that is web browser-based increases potential usage as learners can access the application from a variety of ubiquitous devices (e.g., smart phones, tablets, e-readers, laptops, desktops) without requiring a download and installation of software. Hosting the application in the cloud allows the technology to scale based on demand (i.e., more server instances and resources when increased numbers of students access the application). The support accompanying the deployment of an educational technology requires ongoing professional development and support for educators and students. This professional development and support can be accomplished using both in-person and virtual methods (e.g., online conference calls and meetings, email). Both a robust technology that can accommodate increasingly large user demand along with a well-conceived support plan are necessary elements to successful scale-up of an educational technology.

While the current study could not make causal claims about the impact of deliberate practice on reading ability, future researchers should begin to pursue this line of inquiry. Research should not cease or become suspended since the educational technology is being deployed to a larger audience. The power of a technology-based program such as *Learning Oasis* is that new features, informed by the ongoing research results, can be engineered and released to all students using the system without causing disruption (e.g., no additional

professional development, re-installation of the software). The U.S. Department of Education (USDOE, 2013) recommended an expanding evidence approach to technology-enabled education research whereby evidence is collected continuously as it becomes available. This information is then used for iterative improvement to digital learning systems. Through expanding evidence approaches, factorial designs (e.g., assigning students within the same classroom to different experimental conditions that could include varying level of targeting, type of immediate feedback, or level of self-direction) (Chakraborty, Collins, Strecher, & Murphy, 2009; Collins, Dziak, & Li, 2009; Dziak, Nahum-Shani, Collins, 2012), and thoughtful use of technology as a research platform, ongoing research should continue to explore the relationship between deliberate practice and reading ability.

To prepare students for the literacy challenges of college and career, as defined by the Common Core State Standards (2010) and others (e.g., MetaMetrics, 2008; Williamson, 2008), requires novel approaches to student learning (NETP, 2010). Educators at all levels (i.e., classroom, school, district, state) must start thinking about how technology-based instructional systems, such as *Learning Oasis*, can help their students achieve college and career readiness. The results of this research suggest that technology designed according to the tenants of deliberate practice can help prepare students for life after high school (i.e., more deliberate practice better prepares a student for the literacy demands of college and career, as evidenced by a higher ability estimate). While the discussion about where students need to be, how to get them there, and the evolving nature of technology-based assessment, instruction, and research is still ongoing (e.g., Pearson, 2013; SBAC, 2013; PARCC, 2013; Williamson, Fitzgerald, & Stenner, 2013; USDOE, 2013), the results of this research suggest

that educational technology will play an important role in helping student achieve college and career readiness.

Conclusion

This study was designed to explore the relationship between deliberate practice and reading ability. An educational technology, *Learning Oasis*, played a central role in immersing students in deliberate practice and monitoring changes in reading ability. The results suggested that, controlling for initial reading ability and SES, participants that engaged in more deliberate practice grew at a faster rate and achieved a higher estimated ability than participants that completed less deliberate practice. These findings provided an important first step in understanding how educational technology can be designed to facilitate optimal growth in participant reading ability. Deliberate practice embedded inside an educational technology offers a promising solution that could help prepare students to meet and exceed the literacy demands of college and career.

Appendix A:

Supplemental Exploratory Analysis

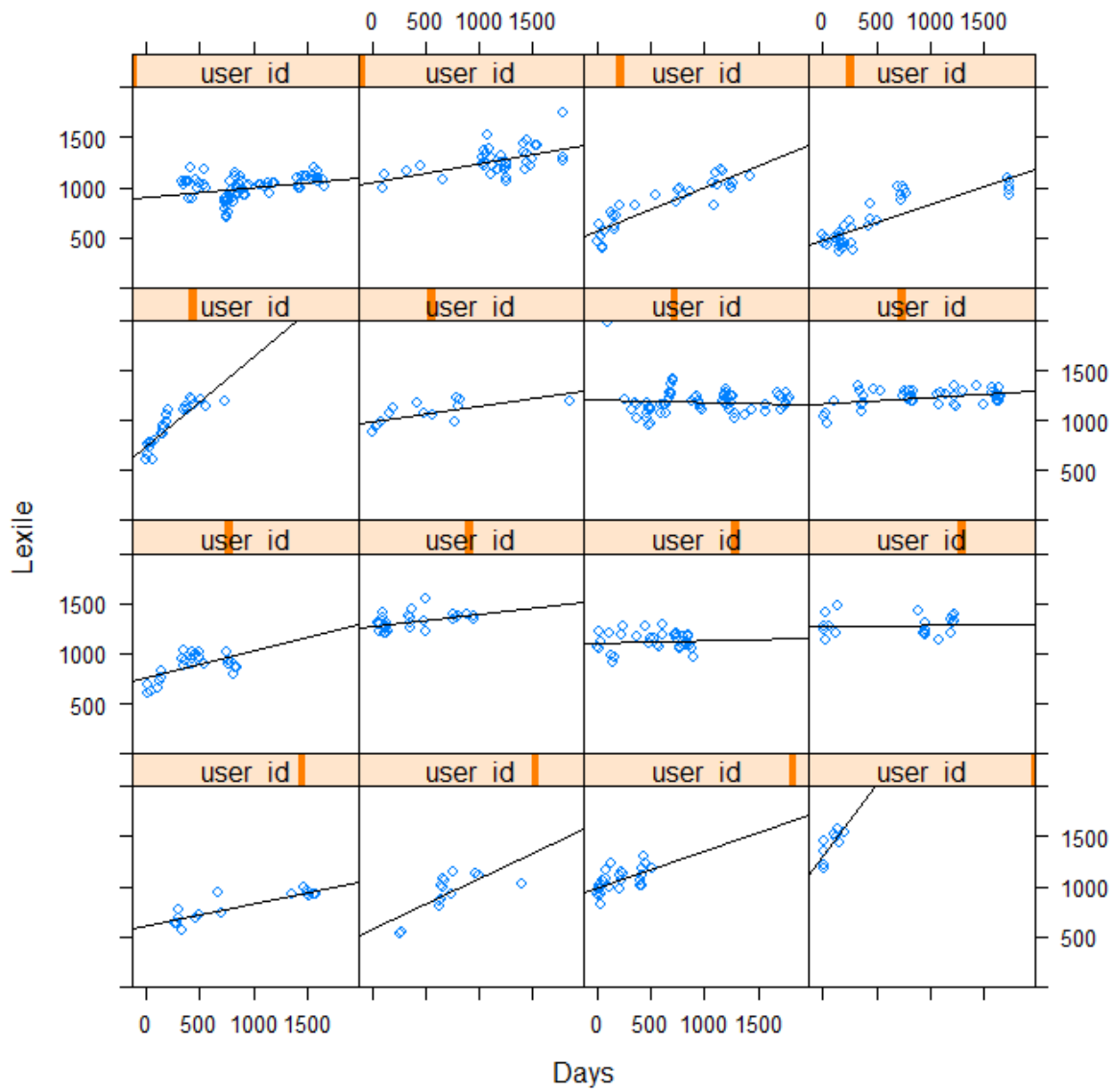


Figure 17. Empirical growth plots with superimposed OLS linear trajectories: Random participants from entire sample ($N = 1,369$)

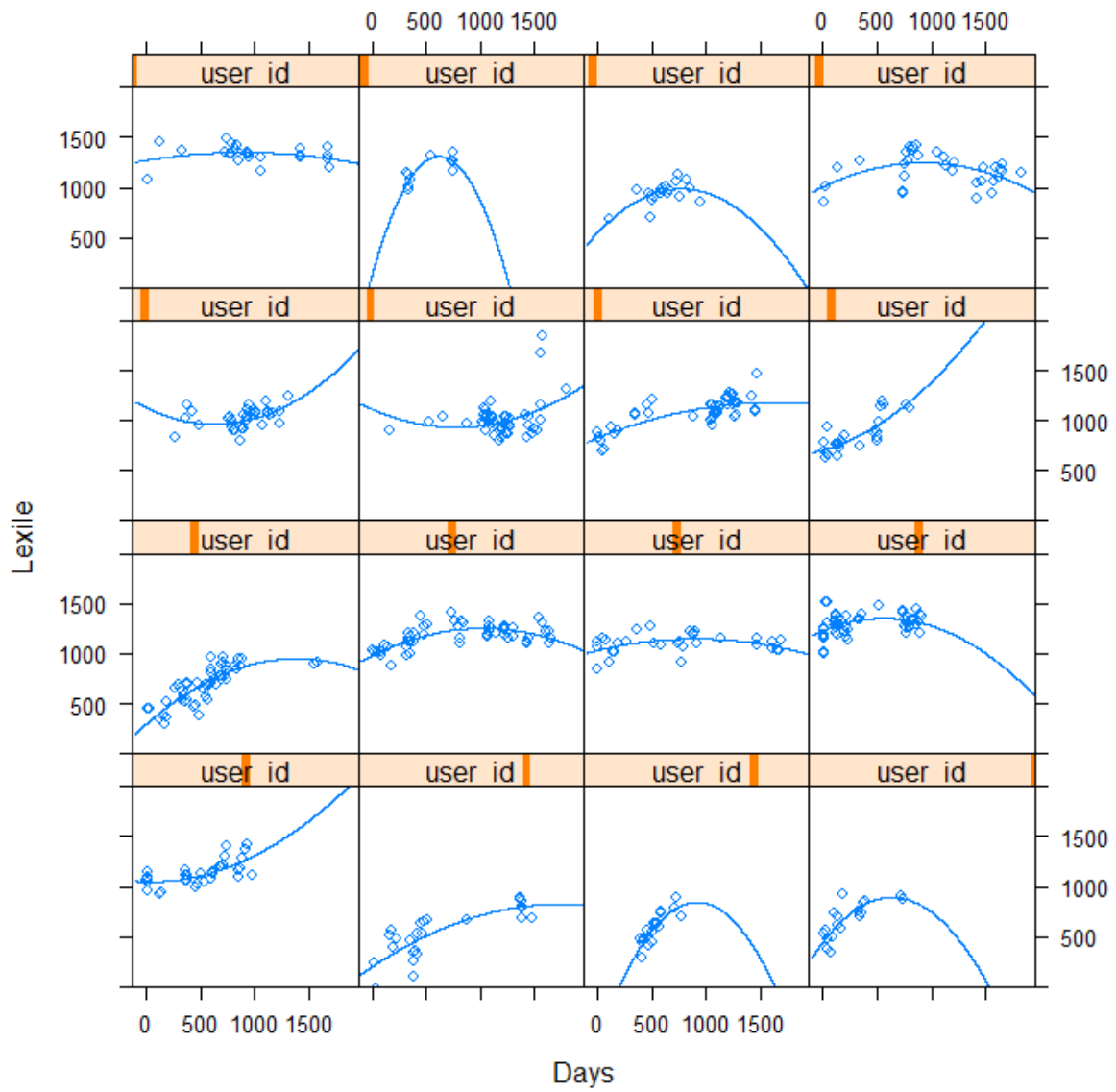


Figure 18. Empirical growth plots with superimposed OLS quadratic trajectories: Random participants from entire sample ($N = 1,369$)

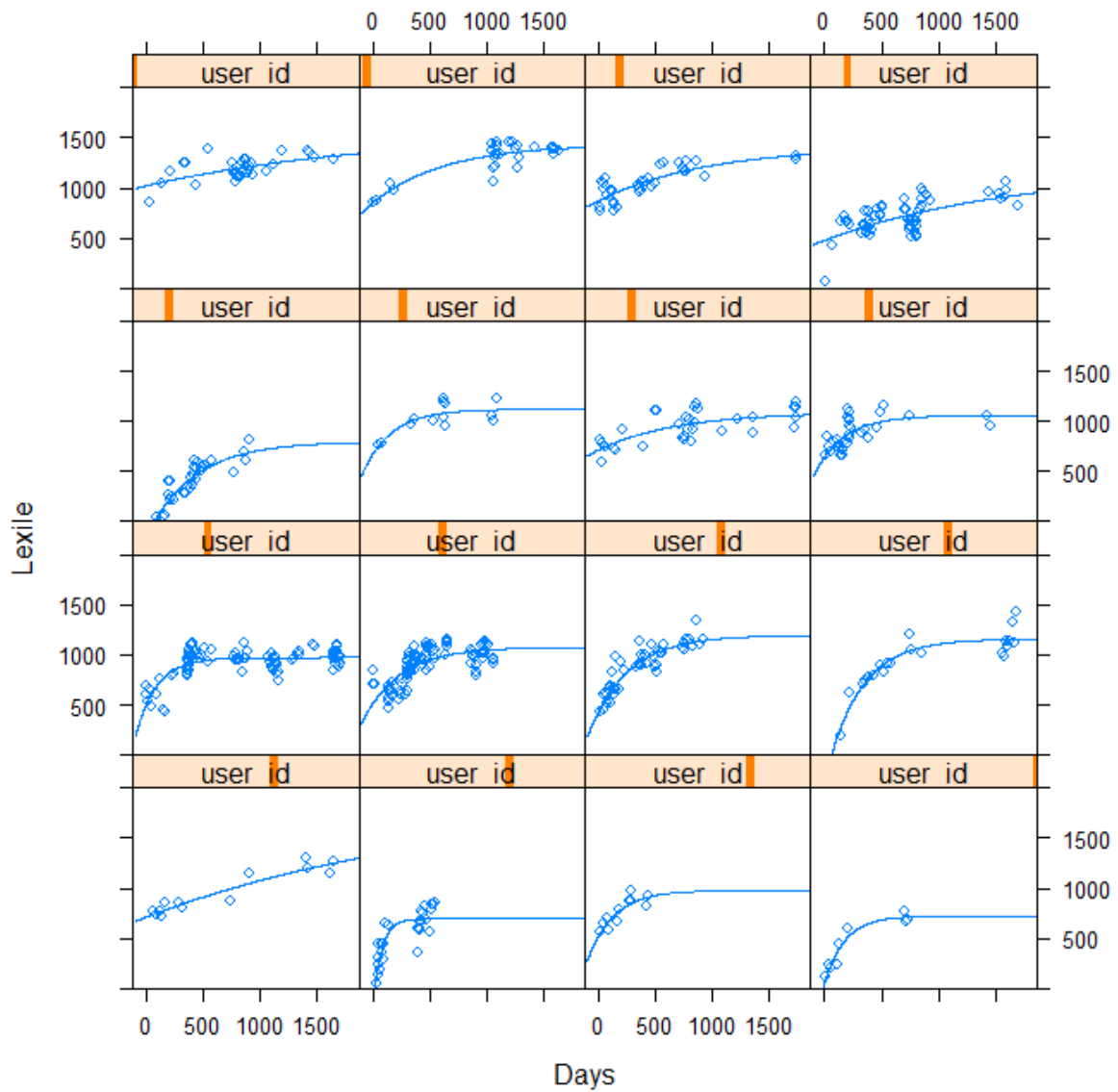


Figure 19. Empirical growth plots with superimposed Nonlinear Least Squares negative exponential trajectories: Random participants from entire sample ($N = 1,369$)

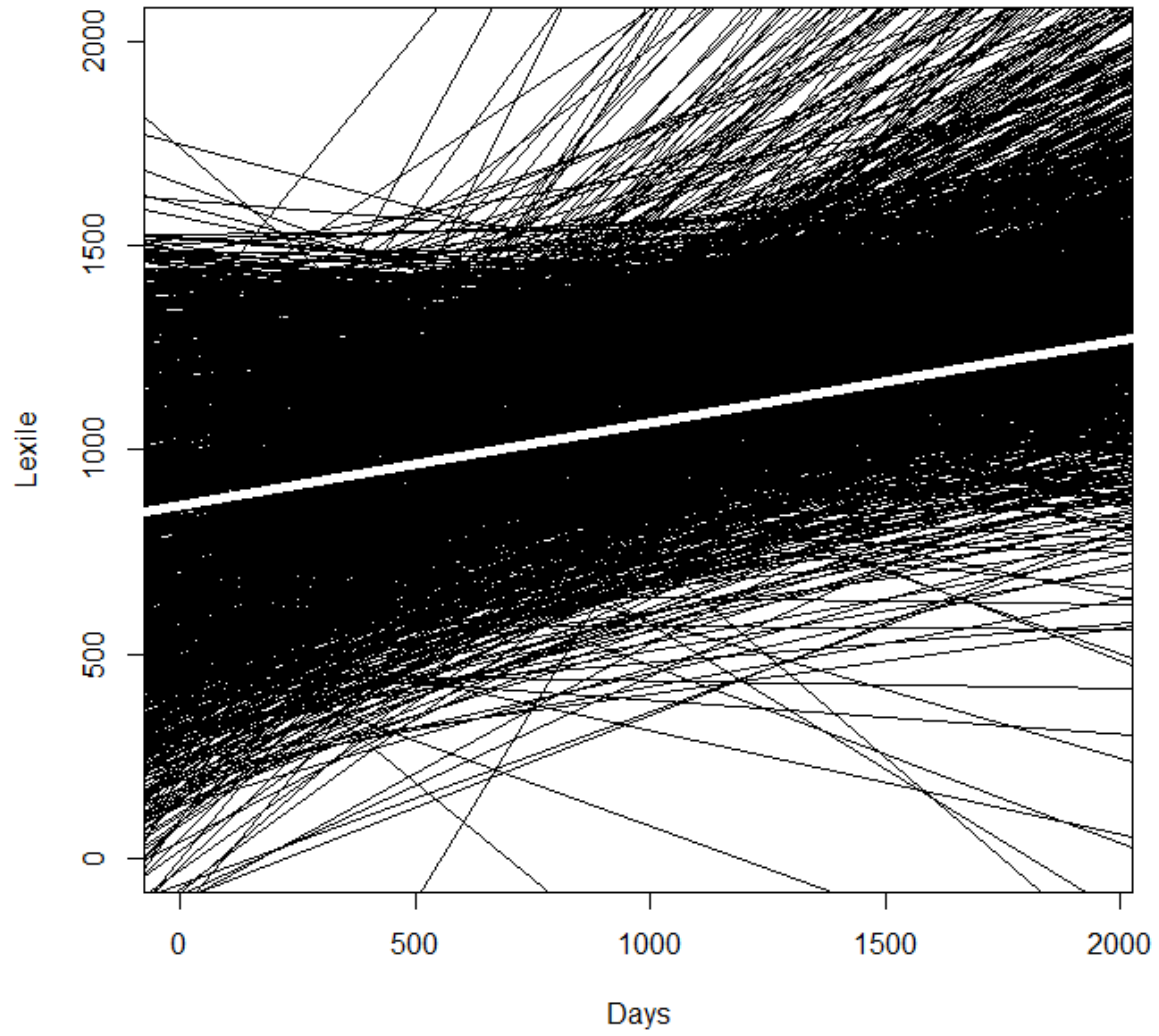


Figure 20. Collection of fitted OLS linear trajectories for all participants in sample ($N=1,369$). White line represents the average change trajectory for the entire group

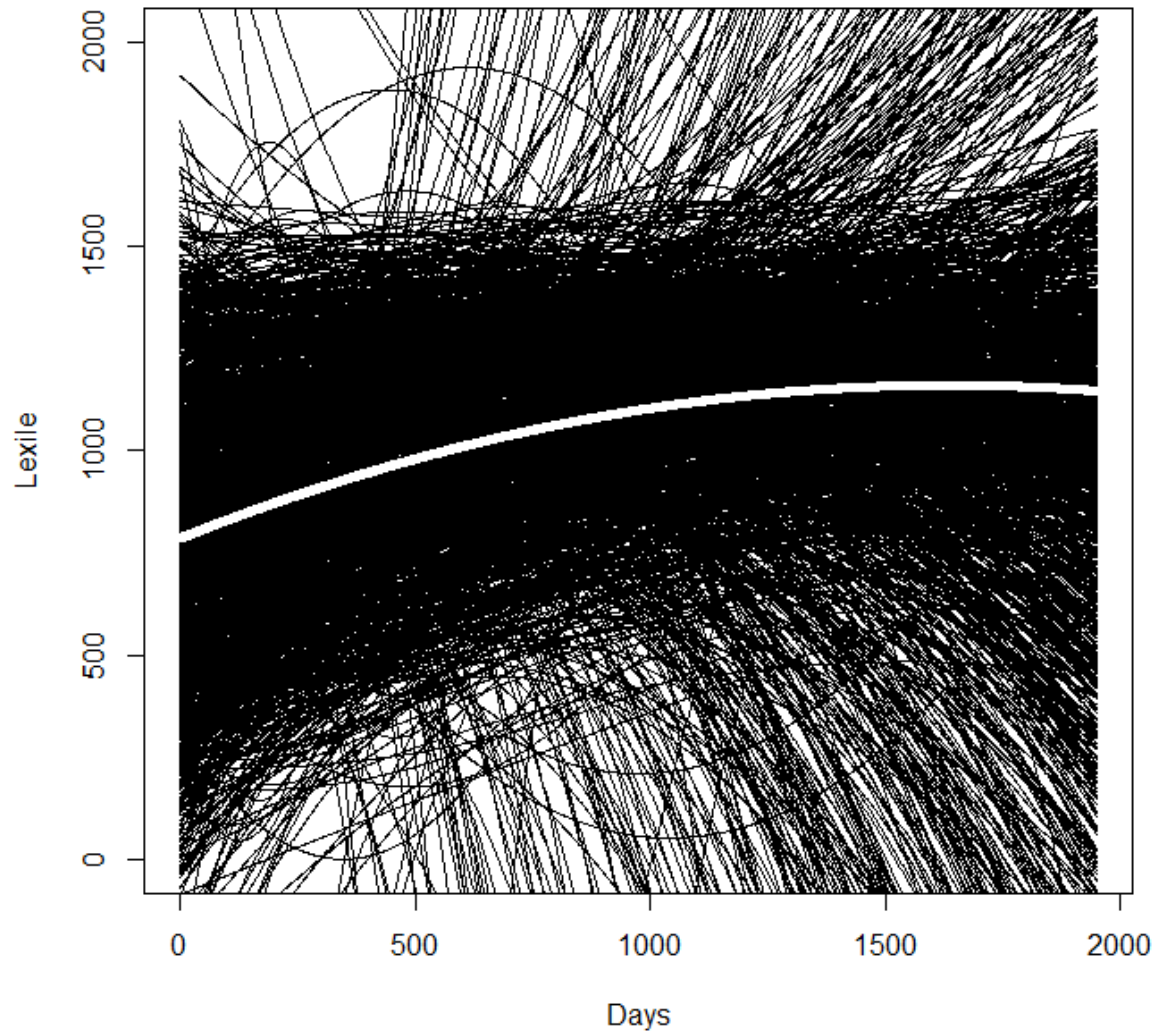


Figure 21. Collection of fitted OLS quadratic trajectories for all participants in sample ($N=1,369$). White line represents the average change trajectory for the entire group

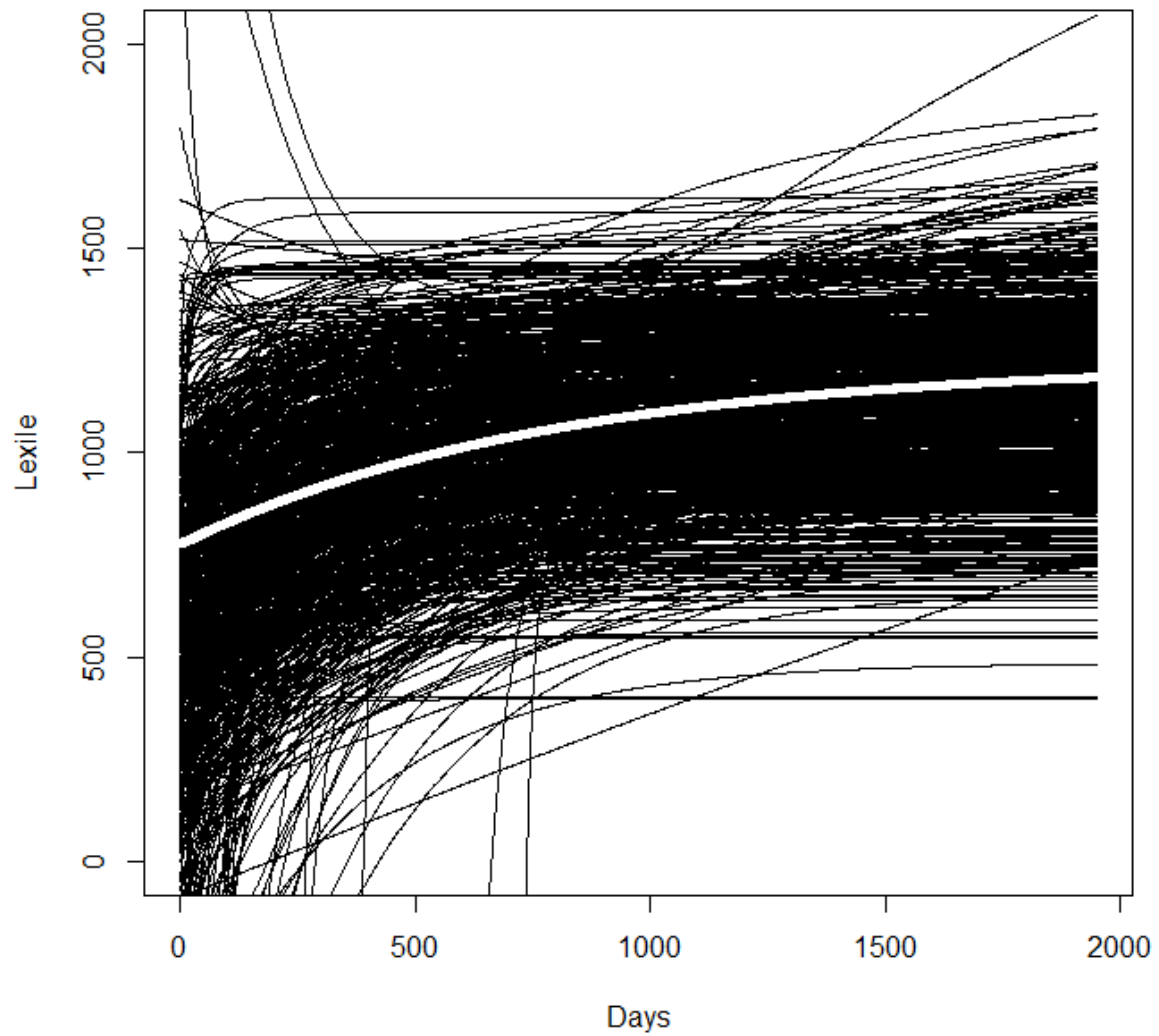


Figure 22. Collection of fitted Nonlinear Least Squares negative exponential trajectories for all participants in sample ($N = 1,369$). White line represents the average change trajectory for the entire group

Table 19. Observed variation in fitted OLS linear trajectories: Fitted initial status (Lexile), rate of change (Lexile per day), correlation between initial status and rate of change, and R^2 statistics for sample of participants stratified by predictors-of-interest

	n	Initial Status (IS)	Rate Parameter (RP)	r (IS-RP)	R^2
Overall	1,369	838.0 (354.1)	0.31 (0.31)	-0.69	0.45 (0.28)
<u>Initial Grade</u>					
Grade 1	29	381.4 (235.9)	0.57 (0.32)	-0.61	0.63 (0.21)
Grade 2	311	491.1 (214.3)	0.53 (0.32)	-0.43	0.63 (0.23)
Grade 3	196	631.7 (202.1)	0.48 (0.27)	-0.26	0.60 (0.20)
Grade 4	142	726.8 (247.2)	0.40 (0.27)	-0.46	0.57 (0.20)
Grade 5	116	914.6 (201.1)	0.23 (0.23)	-0.59	0.49 (0.25)
Grade 6	152	980.6 (217.5)	0.17 (0.15)	-0.56	0.38 (0.26)
Grade 7	95	1054.0 (260.3)	0.17 (0.17)	-0.59	0.28 (0.23)
Grade 8	116	1185.0 (169.0)	0.09 (0.16)	-0.48	0.19 (0.20)
Grade 9	103	1230.0 (181.5)	0.05 (0.16)	-0.66	0.19 (0.21)
Grade 10	92	1250.0 (183.7)	0.07 (0.19)	-0.57	0.20 (0.21)
Grade 11	17	1300.0 (177.8)	0.19 (0.43)	-0.53	0.25 (0.28)
<u>Initial Reading Ability</u>					
Q1: (-89.7L, 521.2L]	343	415.8 (203.1)	0.53 (0.33)	-0.47	0.62 (0.22)
Q2: (521.2L, 771.7L]	342	706.6 (148.3)	0.41 (0.27)	-0.37	0.55 (0.23)
Q3: (771.7L, 1030.1L]	342	969.6 (159.7)	0.23 (0.21)	-0.57	0.41 (0.27)
Q4: (1030.1L, 1975.8L]	342	1259.0 (149.8)	0.07 (0.20)	-0.55	0.22 (0.23)
<u>Lunch Status</u>					
Receives Free	670	746.8 (339.0)	0.33 (0.29)	-0.67	0.47 (0.28)
Pays	699	925.5 (346.3)	0.29 (0.33)	-0.71	0.43 (0.29)

Table 20. Observed variation in fitted OLS quadratic trajectories: Fitted initial status (Lexile), rate of change (Lexile per day), correlation between initial status, rate of change, and quadratic term, and R^2 statistics for sample of participants stratified by predictors-of-interest

	n	Initial Status (IS)	Rate Param. (RP)	Quadratic (C)	r IS-RP, IS-C, RP-C	R^2
Overall	1,369	746.6 (884.6)	0.71 (1.81)	-0.00026 (0.002)	-0.84, 0.38 -0.75	0.58 (0.27)
<u>Initial Grade</u>						
Grade 1	29	339.6 (240.7)	0.70 (0.98)	-0.00022 (0.002)	-0.39, 0.09, -0.82	0.70 (0.20)
Grade 2	311	420.3 (270.3)	0.85 (1.26)	-0.00027 (0.003)	-0.48, 0.26, -0.90	0.73 (0.20)
Grade 3	196	522.6 (559.9)	1.08 (1.75)	-0.00062 (0.002)	-0.85, 0.60, -0.89	0.71 (0.18)
Grade 4	142	517.1 (1671.0)	1.26 (3.70)	-0.00093 (0.003)	-0.98, 0.75, -0.84	0.69 (0.19)
Grade 5	116	700.4 (1475.2)	0.83 (2.09)	-0.00044 (0.001)	-0.96, 0.64, -0.82	0.64 (0.23)
Grade 6	152	918.1 (253.7)	0.43 (0.75)	-0.00016 (0.0007)	-0.60, 0.44, -0.94	0.50 (0.24)
Grade 7	95	1015.6 (323.7)	0.31 (0.90)	-0.000068 (0.0010)	-0.52, 0.26, -0.90	0.44 (0.25)
Grade 8	116	1137.7 (309.9)	0.31 (1.01)	-0.00034 (0.002)	-0.71, 0.19, -0.73	0.33 (0.23)
Grade 9	103	1246.6 (713.4)	0.14 (1.20)	-0.000067 (0.001)	-0.92, 0.56, -0.80	0.32 (0.24)
Grade 10	92	1113.4 (1048.5)	0.43 (1.60)	-0.00052 (0.002)	-0.83, 0.19, -0.63	0.37 (0.25)
Grade 11	17	1505.2 (1278.3)	0.80 (2.83)	-0.00204 (0.005)	-0.83, 0.39, -0.79	0.37 (0.28)
<u>Initial Reading Ability</u>						
Q1: (-89.7L, 521.2L]	343	256.6 (1077.2)	1.24 (2.61)	-0.00069 (0.002)	-0.93, 0.62, -0.82	0.74 (0.19)
Q2: (521.2L, 771.7L]	342	624.6 (275.7)	0.81 (1.16)	-0.00039 (0.002)	-0.54, 0.22, -0.85	0.67 (0.21)
Q3: (771.7L, 1030.1L]	342	869.2 (910.5)	0.56 (1.60)	-0.00023 (0.002)	-0.84, 0.31, -0.73	0.54 (0.24)
Q4: (1030.1L, 1975.8L]	342	1237.4 (746.6)	0.22 (1.37)	-0.00031 (0.002)	-0.81, 0.23, -0.66	0.35 (0.26)
<u>Lunch Status</u>						
Receives Free	670	630.7 (1129.3)	0.81 (2.30)	-0.00050 (0.002)	-0.90, 0.50, -0.76	0.59 (0.26)
Pays	699	857.8 (534.8)	0.61 (1.16)	-0.00031 (0.002)	-0.61, 0.16, -0.80	0.57 (0.28)

Appendix B:

Supplemental Descriptive Statistics

Table 21. Means and standard deviations for reading activity completed by participants, by initial grade

	n	Articles Read	Number of Measures	Minutes Spent Reading	Words Read	Items	Intense Minutes per Day	Total Days Used	SD Days Elapsed bw Readings
Overall	1,369	212.2 (102.3)	22.3 (15.4)	745.8 (389.3)	151,574 (113,117)	2,324 (1,213)	9.03 (5.09)	92.01 (42.33)	34.1 (14.6)
Initial Grade									
Grade 1	29	229.3 (161.7)	18.0 (15.6)	658.8 (336.6)	60,460 (43,321)	1,999 (1,410)	6.72 (1.54)	102.50 (52.81)	31.2 (30.3)
Grade 2	311	197.6 (98.5)	16.5 (10.4)	567.8 (268.2)	61,380 (37,314)	1,718 (877)	6.53 (1.93)	91.51 (45.75)	34.5 (12.6)
Grade 3	196	214.8 (99.4)	20.8 (11.9)	686.2 (321.5)	89,000 (43,721)	2,040 (933)	6.09 (1.67)	114.00 (44.66)	44.88 (17.37)
Grade 4	142	229.1 (99.6)	22.0 (13.0)	641.1 (288.5)	110,000 (52,192)	2,184 (999)	5.29 (1.37)	122.30 (46.92)	30.86 (12.80)
Grade 5	116	182.2 (66.9)	15.8 (8.4)	494.2 (178.6)	103,400 (39,036)	1,703 (600)	4.78 (1.24)	105.50 (30.64)	26.45 (11.58)
Grade 6	152	218.2 (100.0)	26.5 (15.6)	870.6 (459.8)	204,100 (106,782)	2,660 (1,182)	9.25 (3.74)	91.27 (29.64)	30.44 (13.40)
Grade 7	95	219.0 (78.8)	24.6 (15.4)	857.0 (382.1)	211,900 (84,950)	2,748 (977)	11.44 (3.35)	73.62 (23.76)	35.32 (12.29)
Grade 8	116	223.1 (113.2)	32.1 (20.2)	1,023.0 (388.9)	272,400 (99,546)	3,165 (1,356)	13.28 (3.79)	78.89 (27.95)	34.56 (11.41)
Grade 9	103	221.2 (107.1)	28.3 (17.6)	995.4 (396.1)	269,400 (98,382)	3,108 (1,282)	16.72 (4.34)	60.07 (21.03)	37.04 (12.57)
Grade 10	92	237.6 (127.3)	28.9 (22.4)	1,028.0 (455.2)	297,000 (124,211)	3,224 (1,626)	16.86 (4.59)	61.67 (23.83)	26.52 (8.67)
Grade 11	17	212.2 (102.3)	22.3 (15.4)	745.8 (389.3)	151,600 (113,118)	2,324 (1,213)	9.03 (5.09)	92.01 (42.33)	34.05 (14.62)

Table 22. Means and standard deviations for reading activity completed by participants, stratified by initial reading ability and lunch status

	n	Articles Read	Number of Measures	Minutes Spent Reading	Words Read	Items	Intense Minutes per Day	Total Days Used	SD Days Elapsed bw Readings
Overall	1,369	212.2 (102.3)	22.3 (15.4)	745.8 (389.3)	151,600 (113,117)	2,324 (1,213)	9.03 (5.1)	92.01 (42.30)	34.1 (14.6)
Initial Reading Ability									
Q1	343	228.1 (118.9)	20.0 (13.0)	610.6 (305.0)	76,890 (50,104)	2,090 (1,091)	6.21 (1.92)	102.70 (50.88)	35.4 (16.1)
Q2	342	210.0 (96.2)	21.85 (14.4)	640.6 (322.9)	104,300 (73,985)	2,064 (1,084)	6.59 (2.70)	101.70 (44.32)	36.2 (16.3)
Q3	342	210.6 (95.6)	25.1 (17.7)	773.6 (409.3)	175,300 (118,920)	2,464 (1,400)	8.79 (4.37)	93.04 (34.16)	31.4 (13.2)
Q4	342	200.1 (95.0)	22.3 (15.6)	958.7 (408.8)	250,100 (105,115)	2,678 (1,143)	14.53 (5.44)	70.62 (28.16)	33.2 (12.0)
Lunch Status									
Receives Free	670	216.8 (108.1)	23.4 (15.9)	685.6 (362.2)	135,100 (105,451)	2,326 (1,260)	8.18 (4.43)	92.04 (43.41)	34.6 (16.2)
Pays	699	207.8 (96.3)	21.3 (14.8)	803.5 (405.4)	167,400 (117,930)	2,322 (1,168)	9.84 (5.54)	91.99 (41.30)	33.5 (12.9)

Appendix C:

Figures Supporting Examination of Assumptions

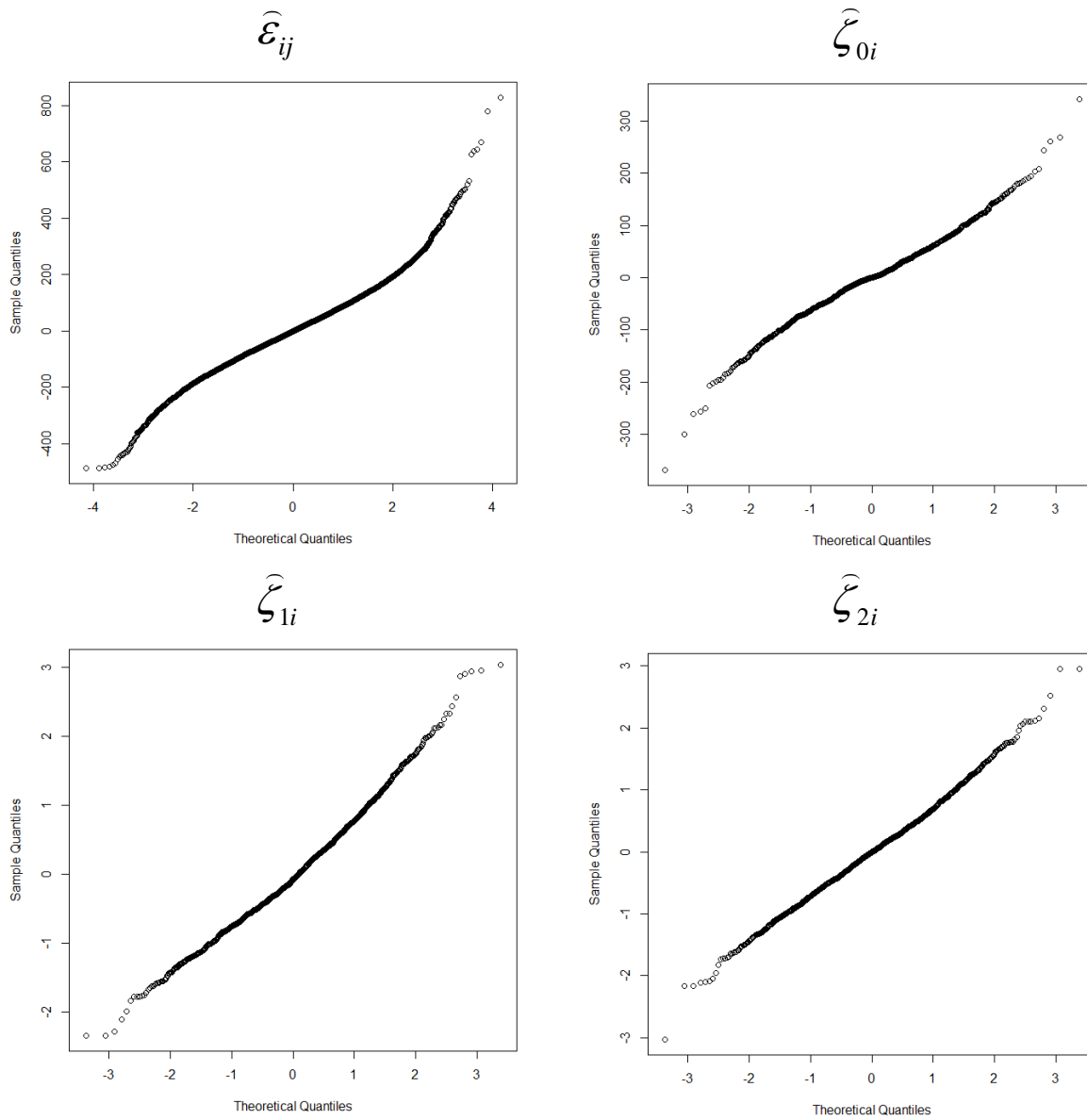


Figure 23. Examining normality assumptions in the multilevel model for change: Normal probability plots for the raw residuals at level-1 and level-2.

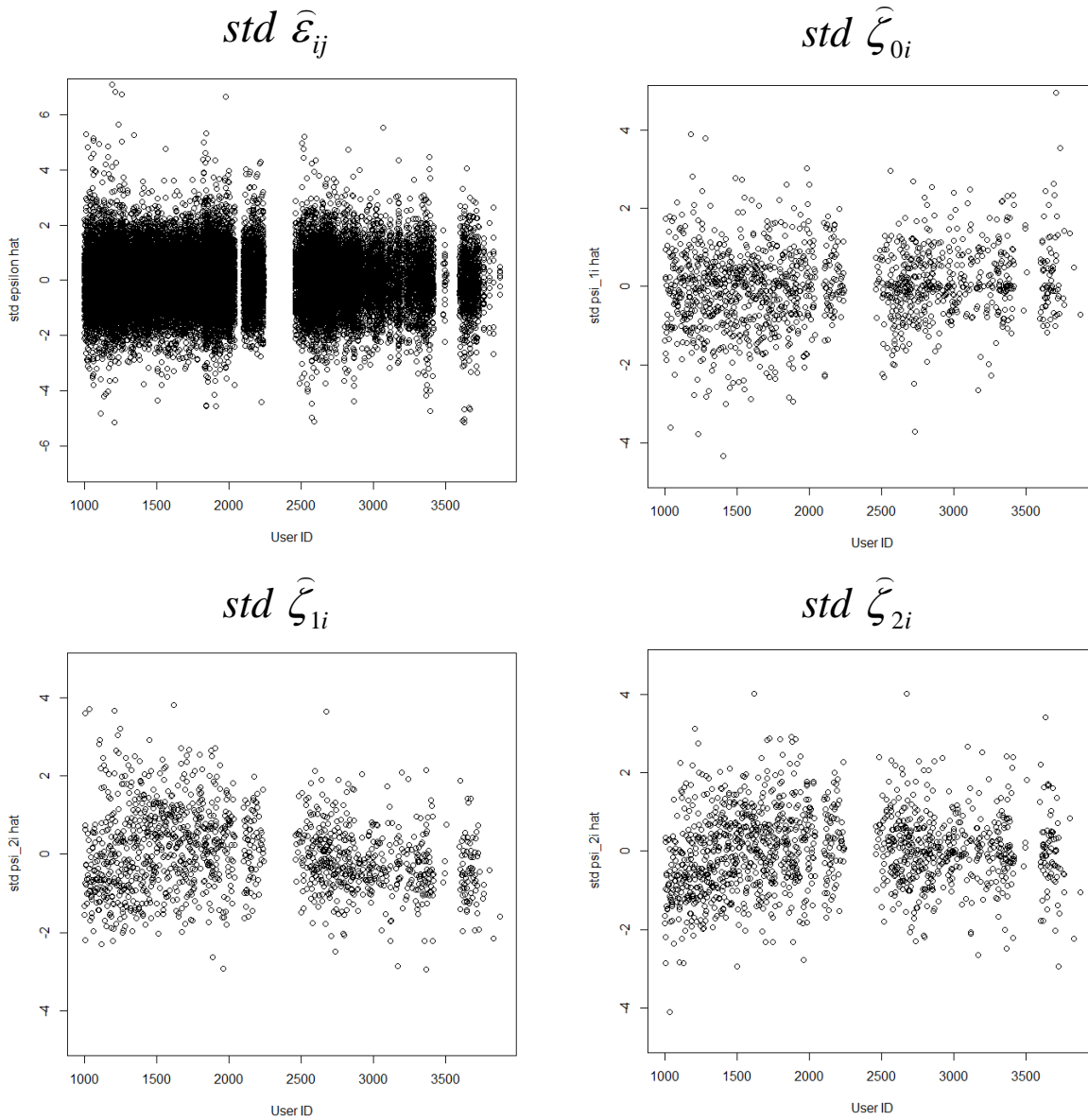


Figure 24. Examining normality assumptions in the multilevel model for change: Plots of standardized residuals at level-1 and level-2 versus participant ID numbers.

$$\hat{\epsilon}_{ij}$$

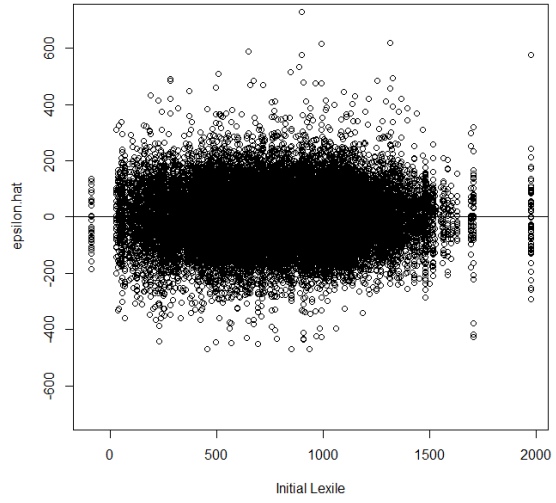


Figure 25. Examining the homoscedasticity assumptions in the multilevel model for change: Raw level-1 residuals vs. level-1 predictors for Initial Lexile

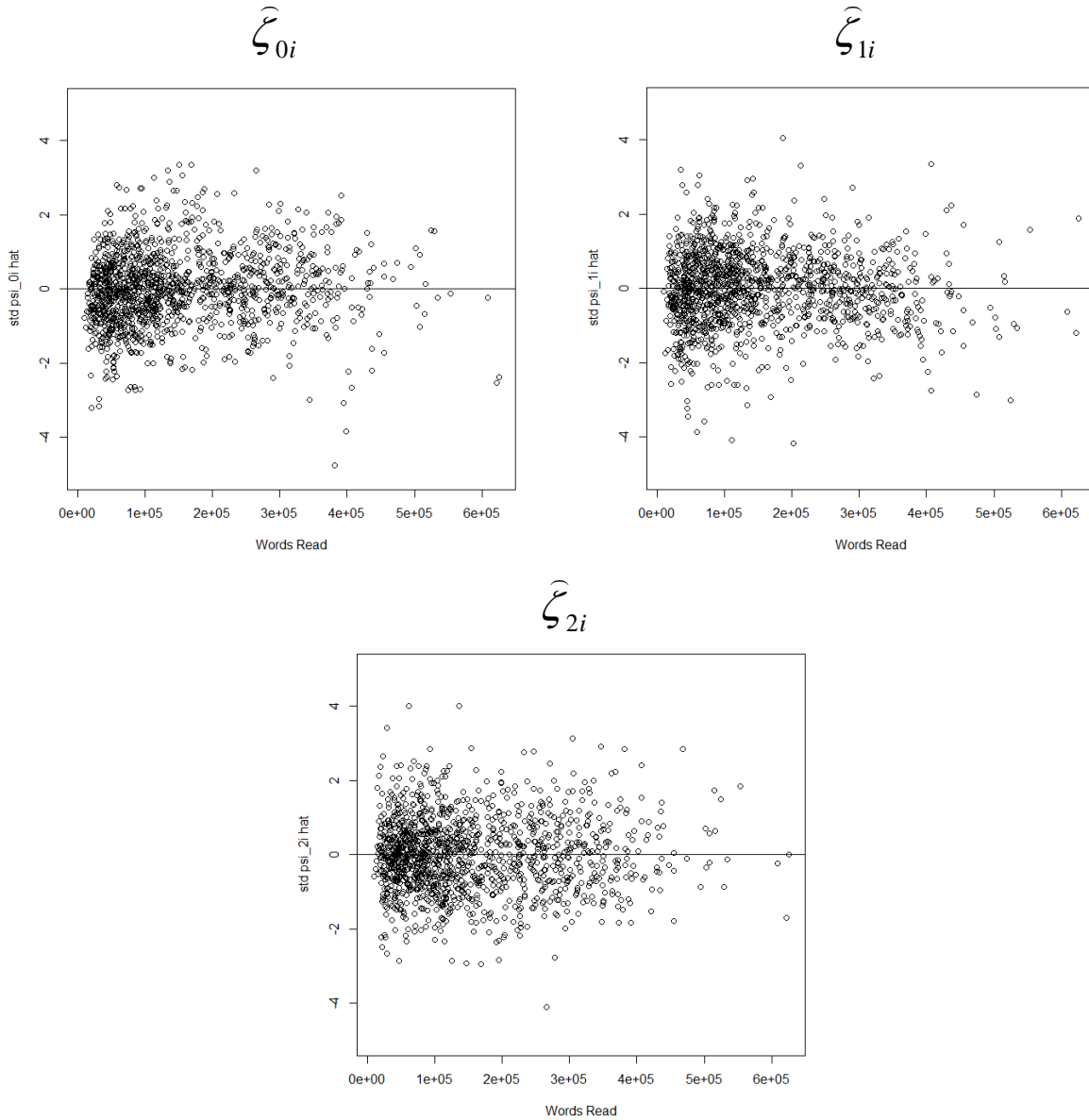


Figure 26. Examining the homoscedasticity assumptions in the multilevel model for change: Raw level-2 residuals vs. level-2 predictor *Words*

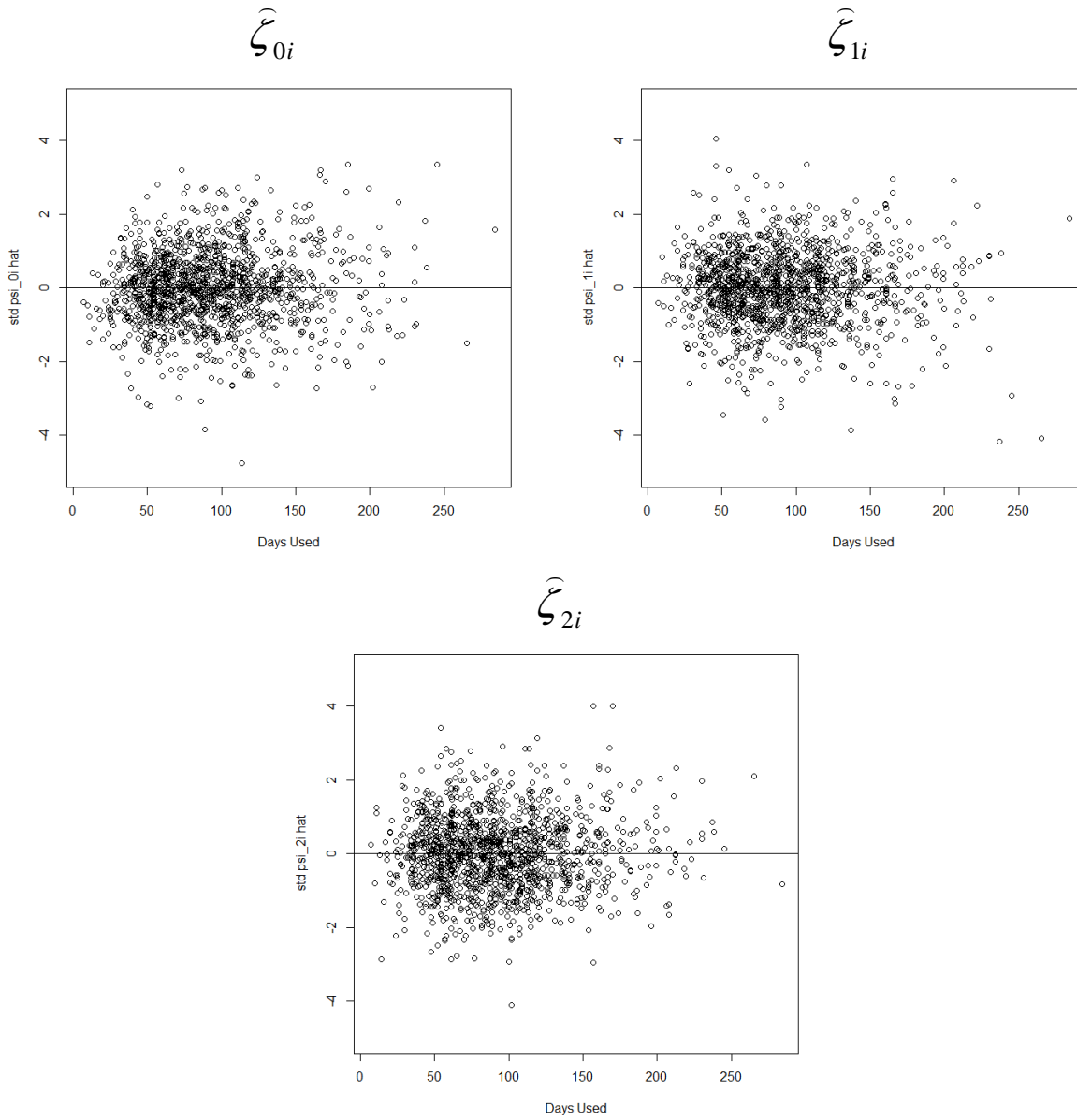


Figure 27. Examining the homoscedasticity assumptions in the multilevel model for change: Raw level-2 residuals vs. level-2 predictor *DaysUsed*

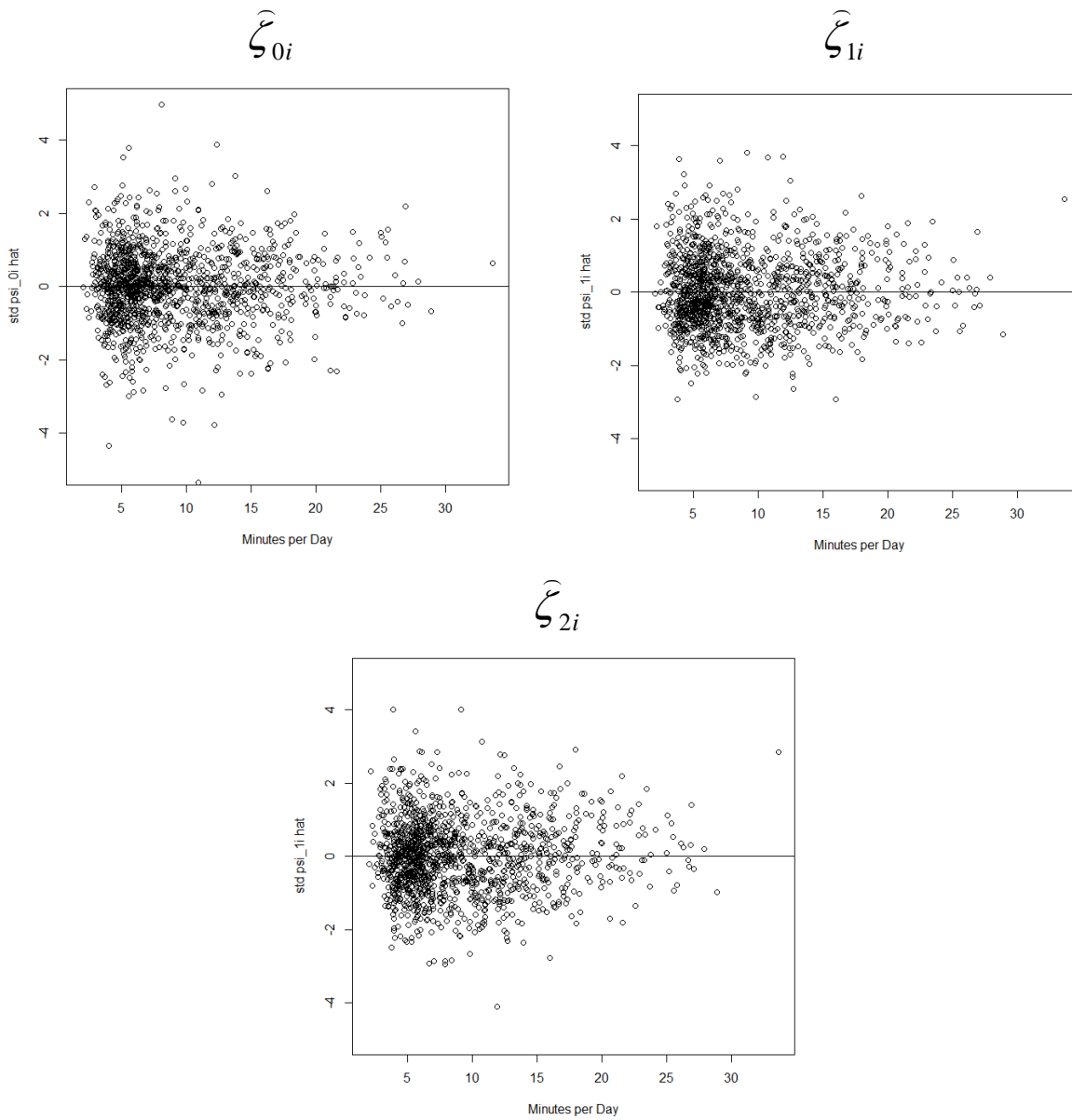


Figure 28. Examining the homoscedasticity assumptions in the multilevel model for change: Raw level-2 residuals vs. level-2 predictor *IntenseMinPerDay*

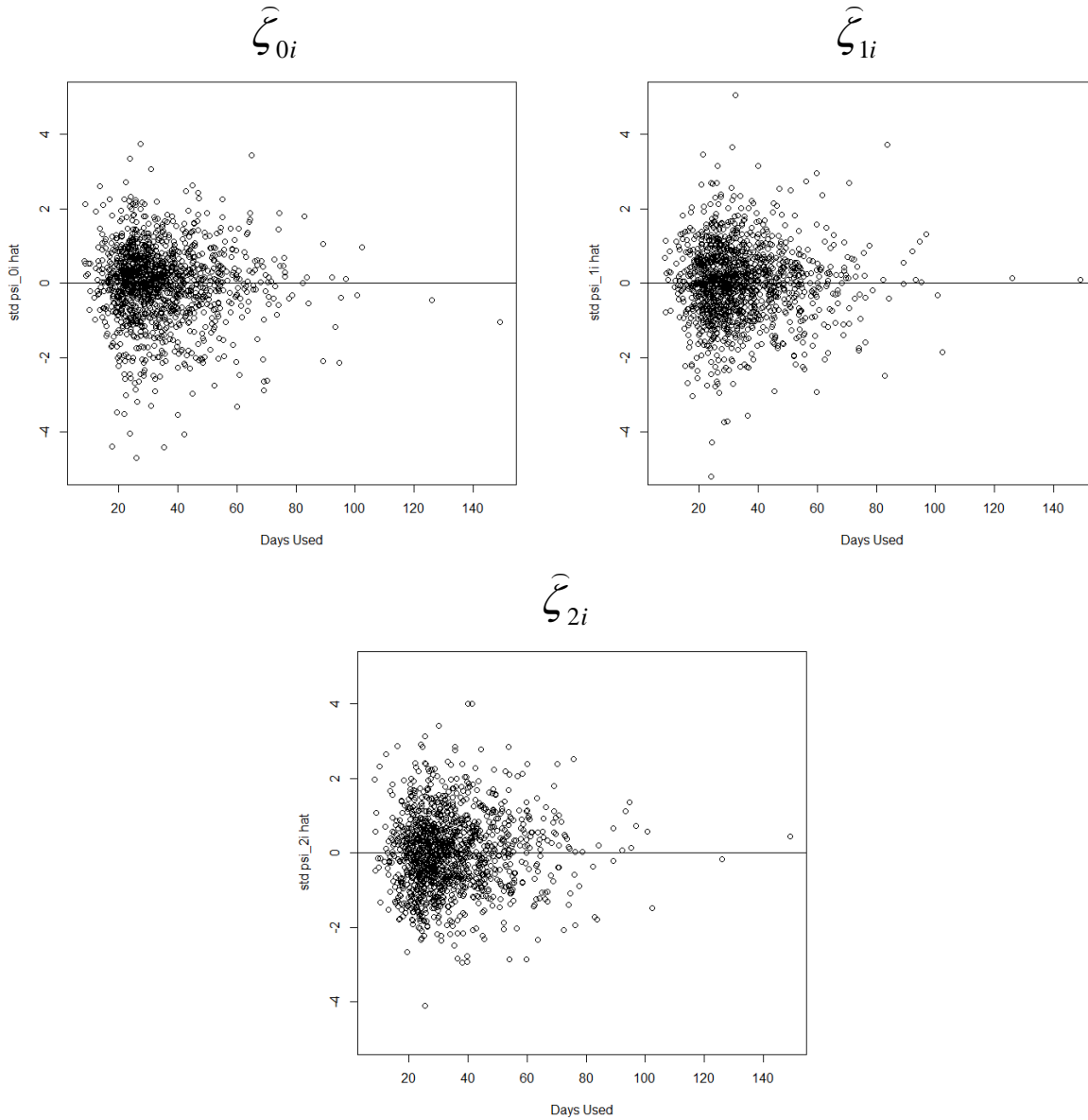


Figure 29. Examining the homoscedasticity assumptions in the multilevel model for change: Raw level-2 residuals vs. level-2 predictor *ElapsedDaysSD*

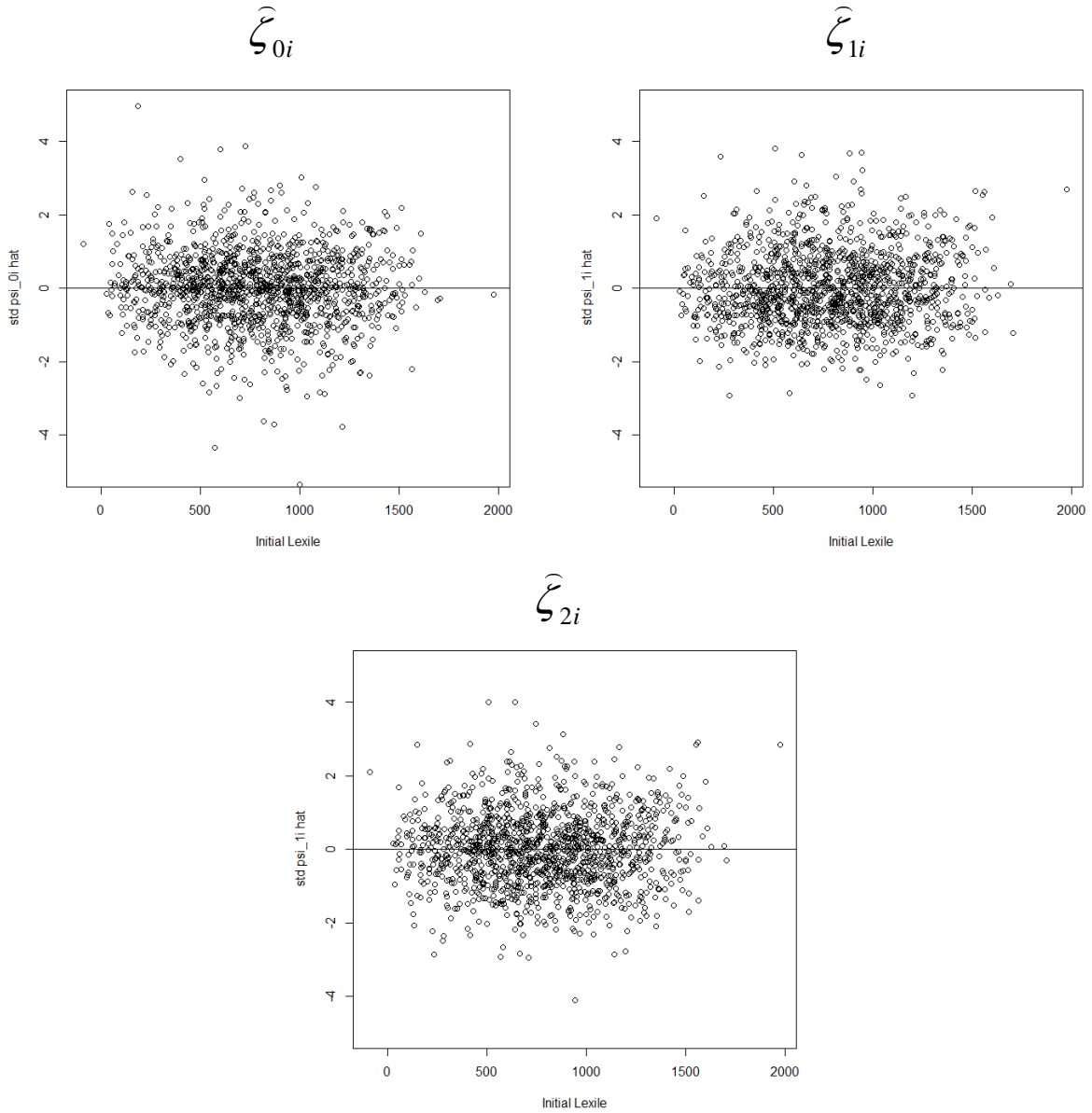


Figure 30. Examining the homoscedasticity assumptions in the multilevel model for change: Raw level-2 residuals vs. level-2 predictor *InitialLexile*

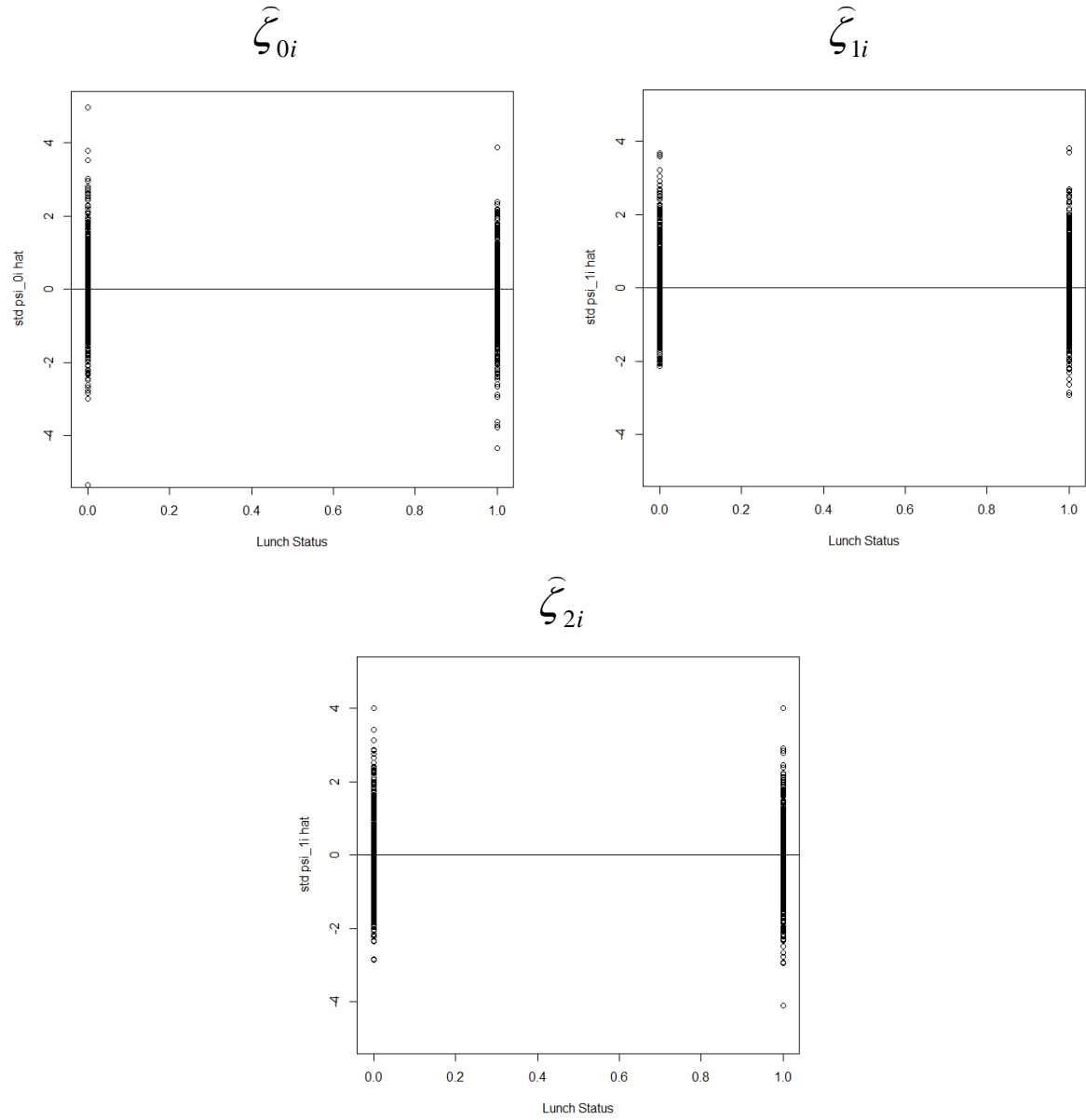


Figure 31. Examining the homoscedasticity assumptions in the multilevel model for change: Raw level-2 residuals vs. level-2 predictor *SES*

References

- Achieve, Inc. (2005). *Rising to the challenge: Are high school graduates prepared for college and work?* Washington, DC.
- Ackerman, P. L., & Beier, M. E. (2006). Methods of studying the structure of expertise: Psychometric approaches. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 147-166). Cambridge, UK: Cambridge University Press.
- ACT. (2005). *Crisis at the core: Preparing all students for college and work*. Iowa City. Retrieved from http://www.act.org/path/policy/pdf/crisis_report.pdf
- ACT. (2012). Retrieved from <http://actstudent.org/faq/answers/numbquestions.html>
- Adams, K. H. (1993). *A history of professional writing instruction in American colleges*. Dallas: Southern Methodist Press.
- Adams, M. J. (1979). Some differences between good and poor readers. In M. L. Kamil, & A. J. Moe (Eds.), *Reading research: Studies and applications* (pp. 140-144). Clemson, SC: National Reading Conference.
- Adams, M. J. (1981). What good is orthographic redundancy? In O. J. Tzeng, & H. Singer (Eds.), *Perception of print* (pp. 197-221). Hillsdale, NJ: Erlbaum.
- Adams, M. J. (1990). *Beginning to read*. Cambridge, MA: MIT Press.
- Adams, M. J. (1994). Modeling the connections between word recognition and reading. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed., pp. 838-863). Newark, DE: International Reading Association.
- Adelson, B. (1981). Problem solving and the development of abstract categories in programming. *Memory and Cognition*, 9, 422-433.
- Akaike, H. (1974). Information theory as an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory* (pp. 267-288). Akademiai Kiado, Budapest, Hungary.
- Alaska Department of Education and Early Development. (2012). Retrieved from <http://www.eed.state.ak.us/tls/assessment/>
- Alderson, J. C. (1979b). The cloze procedure and proficiency in English as a foreign language. *TESOL Quarterly*, 13(2), 219-227.

- Alexander, P. A. (1997). Mapping the multidimensional nature of domain learning: The interplay. In M. L. Maehr, & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (pp. 213-250). Greenwich, CT: JAI Press.
- Alexander, P. A. (2003a). The development of expertise: The journey from acclimation to proficiency. *Educational Researcher*, 32(8), 10-14.
- Alexander, P. A. (2003b). Profiling the developing reader: The interplay of knowledge, interest, and strategic processing. In C. M. Fairbanks, J. Worthy, B. Maloch, J. V. Hoffman, & D. L. Schallert (Eds.), *52nd Yearbook of the National Reading Conference* (pp. 47-65). Oak Creek, WI: National Reading Conference.
- Alexander, J. E., & Filler, R. C. (1976). *Attitudes and reading*. Newark, DE: International Reading Association.
- Alexander, P. A., & Fox, E. (2004). A historical perspective on reading research and practice. In R. B. Ruddell, & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 33-68). Newark, DE: International Reading Association.
- Alexander, P. A., Kulikowich, J. M., & Schulze, S. K. (1994). The influence of topic knowledge, domain knowledge, and interest on the comprehension of scientific exposition. *Learning and Individual Differences*, 6, 379-397.
- Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research*, 61, 315-343.
- Alexander, P. A., Sperl, C. T., Buehl, M. M., Fives, H., & Chiu, S. (2004). Modeling domain learning: Profiles from the field of special education. *Journal of Educational Psychology*, 96(3), 545-557.
- Allen, L., Cipielewski, J., & Stanovich, K. E. (1992). Multiple indicators of children's reading habits and attitudes: Construct validity and cognitive correlates. *Journal of Educational Psychology*, 84, 489-503.
- Allington, R. L. (1977). If they don't read much, how are they ever gonna get good? *Journal of Reading*, 21, 57-61.
- Allington, R. L. (1980). Poor readers don't get much in reading groups. *Language Arts*, 57(8), 872-877.
- Allington, R. L. (1983). The reading instruction provided readers of differing abilities. *Elementary School Journal*, 83, 548-559.

- Allington, R. L. (1984a). Content coverage and contextual reading in reading groups. *Journal of Reading Behavior, 16*(1), 85-96.
- Allington, R. L. (1984b). Oral reading. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 829-864). New York: Longman.
- Allington, R. L. (2009). If they don't read much...30 years later. In E. H. Hiebert (Ed.), *Reading more, reading better* (pp. 30-54). New York: Guilford Press.
- Amirault, R. J., & Branson, R. K. (2006). Educators and expertise: A brief history of theories and models. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 69-86). Cambridge, UK: Cambridge University Press.
- Anderson, C. (2000). Sustained silent reading: Try it, you'll like it! *Reading Teacher, 54*(3), 258-259.
- Anderson, J. R. (Ed.). (1981). *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum.
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369-406.
- Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review, 94*, 192-210.
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing*. New York: Longman.
- Anderson, R. C., & Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. In B. Hutson (Ed.), *Advances in reading/language research* (pp. 231-256). Greenwich, CT: JAI Press.
- Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading. In P. D. Pearson (Ed.), *Handbook of reading research* (Vol. 1, pp. 185-224). New York: Longman.
- Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. (1989). *Becoming a nation of readers: The report of the Commission on Reading*. Washington, DC: National Institute of Education.
- Anderson, R. C., Wilson, P. T., & Fielding, L. C. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly, 23*, 285-303.

- Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review*, 105, 678-723.
- Arizona Department of Arizona. (2012). Retrieved from <http://www.azed.gov/standards-development-assessment/>
- Armbruster, B. B., Lehr, F., & Osborn, J. (2001). *Put reading first: The research building blocks for teaching children to read*. Washington, DC: Center for Improvement of Early Reading Achievement.
- Aud, S., Fox, M. A., KewelRamani, A. (2010). *Status and trends in the education of racial groups* (NCES 2010-015). U.S. Department of Education, National Center for Education Statistics. Washington, D.C.: U.S. Government Printing Office.
- Baddeley, A. D. (1979). Working memory and reading. In P. Kolers, E. Wrolstad, & H. Bouma (Eds.), *Processing of visible language* (Vol. 1). New York: Plenum.
- Baddeley, A. D. (2000). Short-term and working memory. In E. Tulving, & F. Craik (Eds.), *The Oxford handbook of memory* (pp. 77-92). Oxford, UK: Oxford University Press.
- Baddeley, A. D. (2002). Is working memory still working? *European Psychologist*, 7(2), 85-97.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York: Academic Press.
- Bachman, L. F. (1985). Performance on cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly*, 19, 535-555.
- Baker, L., Afflerbach, P., & Reinking, D. (1996). Developing engaged readers in school and home communities: An overview. In L. Baker, P. Afflerbach, & D. Reinking (Eds.), *Developing engaged readers in school and home communities* (pp. xiii-xxvii). Mahwah, NJ: Erlbaum.
- Baker, L. & Wigfield, A. (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading achievement. *Reading Research Quarterly*, 34(4), 452-477.
- Bandura, Albert (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Barron, B. (2006). Interest and self-sustained learning as catalysts of development: A learning ecology perspective. *Human Development*, 49(4), 153-224.
- Bartlett, F. (1958). *Thinking*. New York: Basic Books.

- Bartlett, F. C. (1932). *Remembering: A study of experimental and social psychology*. Cambridge, UK: Cambridge University Press.
- Bean, T.W., & Rigoni, N. (2001). The impact of a multicultural novel on intergenerational dialogue journal discussion. *National Reading Conference Handbook*, 50. 127-137.
- Bedard, J., & Chi, M. H. (1993). Expertise in auditing. *Auditing*, 12, 1-25.
- Benner, P. A. (1984). *From novice to expert: Excellence and power in clinical nursing practice*. Menlo Park: Addison-Wesley.
- Benner, P. A., Tanner, C. A., & Chesla, C. A. (1996a). Entering the field: Advanced beginner practice. In P. A. Benner, C. A. Tanner, & C. A. Chesla (Eds.), *Expertise in nursing practice* (pp. 48-77). New York: Springer Publishing Company.
- Benner, P. A., Tanner, C. A., & Chesla, C. A. (1996b). The competence stage: A time of analysis, planning and confrontation. In P. A. Benner, C. A. Tanner, & C. A. Chesla (Eds.), *Expertise in nursing practice* (pp. 78-113). New York: Springer Publishing Company.
- Benner, P. A., Tanner, C. A., & Chesla, C. A. (1996c). Proficiency: A transition to expertise. In P. A. Benner, C. A. Tanner, & C. A. Chesla (Eds.), *Expertise in nursing practice* (pp. 114-141). New York: Springer Publishing Company.
- Benner, P. A., Tanner, C. A., & Chesla, C. A. (1996d). Expert practice. In P. A. Benner, C. A. Tanner, & C. A. Chesla (Eds.), *Expertise in nursing practice* (pp. 142-169). New York: Springer Publishing Company.
- Bereiter, C. (1963). Some persisting dilemmas in the measurement of change. In C. W. Harris (Ed.), *Problems in measurement of change* (pp. 3-20). Madison, WI: University of Wisconsin Press.
- Bereiter, C. (1991). Implications of connectionism for thinking about rules. *Educational Researcher*, 20(3), 10-16.
- Bergeron, B. S. (1990). What does the term Whole Language mean?: Constructing a definition from the literature. *Journal of Reading Behavior*, 22(4), 301-329.
- Bernard, H. R. (2000). *Social research methods: Qualitative and quantitative methods*. Thousand Oaks, CA: Sage.
- Berninger, V. W., Abbott, R. D., Brooksher, R., Lemos, Z., Ogier, S., Zook, D., & Mostafapour, E. (2000). A connectionist approach to making the predictability of English orthography explicit to at-risk beginning readers: Evidence for alternative, effective strategies. *Developmental Neuropsychology*, 17, 241-271.

- Best, J. L., & Kahn, J. V. (1998). *Research in education* (8th ed.). Needham Height, MA: Allyn & Bacon.
- Betts, E. A. (1946). *Foundations of reading instruction with emphasis on differentiated guidance*. New York: American Book Company.
- Biancarosa, C., & Snow, C. E. (2006). Reading next: A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York (2nd ed.). Washington, DC: Alliance for Excellent Education.
- Biemiller, A. (1970). The development of the use of graphic and contextual information as children learn to read. *Reading Research Quarterly*, 6, 75-96.
- Bigge, M. L., & Shermis, S. S. (1992). *Learning theories for teachers* (5th ed.). New York: Harper Collins.
- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. London: King's College.
- Blanchard, J., Rottenberg, C., & Jones, J. (1989). *Foundational literature in elementary reading methodology textbooks*. Tempe, AZ: Arizona State University, College of Education.
- Bloom, B. (1985). *Developing talent in young people*. New York: Ballentine.
- Bock, R. D. (1983). The discrete Bayesian. In H. Wainer, & S. Messick (Eds.), *Principles of modern psychological measurement: A festschrift for Frederick M. Lord*. Hillsdale, NJ: Erlbaum.
- Bock, R. D., & Moore, E. G. (1986). *Advantage and disadvantage: A profile of American youth*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Boice, R. (1994). *How writers journey to comfort and fluency: A psychological adventure*. Westport, CT: Praeger.
- Boice, R. (1997). Which is more productive, writing in binge patterns of creative illness or in moderation. *Written Communications*, 14, 435-459.
- Borich, G. D., & Tombari, M. L. (1997). *Educational psychology: A contemporary approach* (2nd ed.). New York: Addison-Wesley Educational Publishers.
- Bormuth, J. R. (1966). Readability. *Reading Research Quarterly*, 7, 79-132.
- Bormuth, J. R. (1967). Comparable cloze and multiple choice comprehension tests scores. *Journal of Reading*, 10, 291-299.

- Boushey, G., & Moser, J. (2006). *The daily five*. Stenhouse Publishing.
- Bracken, B. A., & McCallum, R. S. (1993). *Woodcock-Johnson Psycho-Educational Battery Revised*. Brandon, VT: Clinical Psychology Publishing Co.
- Bradley, R. H., Corwyn, R. F., Burchinal, M., Pipes-McAdoo, H., Garcia-Coll, C. (2001). The home environments of children in the United States, Part I. Variations by age, ethnicity, and poverty status. *Child Development, 72*(6), 1844-1867.
- Bransford, J. D., Barron, B., Pea, R., Meltzoff, A., Kuhl, P., Bell, P., Stevens, R., Schwartz, D., Vye, N., Reeves, B., Roschelle, J., & Sabelli, N. (2006). Foundations and opportunities for an interdisciplinary science of learning. In K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 19-34). New York: Cambridge University Press.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn*. Washington, DC: National Academy Press.
- Brown, A. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin, 109*, 204-223.
- Brown, A., & MacNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5*, 325-337.
- Brown, G. D. (1987). Resolving inconsistency: A computational model of word naming. *Journal of Memory and Language, 23*, 1-23.
- Brown, G. D. (1997). Connectionism, phonology, reading, and regularity in developmental dyslexia. *Brain and Language, 59*, 207-235.
- Brown, J. D. (1983). A closer look at cloze: Validity and reliability. In J. W. Oller, Jr. (Ed.) *Issues in Language Testing Research* (pp. 237-250). Rowley, MA: Newbury House.
- Brumbaugh, R. S., & Lawrence, N. M. (1963). *Philosophers on education: Six essays on the foundations of western thought*. Boston: Houghton Mifflin.
- Brumbaugh, R. S., & Lawrence, N. M. (1985). *Philosophical themes in modern education*. Landham, MD: Houghton Mifflin.
- Bryan, W. L., & Harter, N. (1899). Studies on the telegraphic language: The acquisition of a hierarchy of habits. *Psychological Review, 6*, 345-375.
- Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin, 101*, 147-158.

- Bureau of Labor Statistics. (2007). *The 30 fastest growing occupations covered in the 2008-2009 Occupational Outlook Handbook*. Retrieved from <http://www.bls.gov/news.release/ooh.t01.htm>
- Burns, G. L., & Kondrick, P. A. (1998). Psychological behaviorism's reading therapy program: Parents as reading therapists for their children's disability. *Journal of Learning Disabilities, 31*(3), 278-285.
- Burns, P. C., Roe, B. D., & Ross, E. P. (1999). *Teaching reading in today's elementary schools* (7th ed.). Boston: Houghton Mifflin.
- Butler, A. C. (2010). Repeated testing procedures produce superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology, 36*(5), 1118-1133.
- Butterworth, B. (2006). Mathematical expertise. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 553-568). Cambridge, UK: Cambridge University Press.
- Byrnes, J. P. (2001). *Minds, brains, and learning: Understanding the psychological and educational relevance of neuroscientific research*. New York: Guilford Press.
- California Department of Education. (2012). Retrieved from <http://www.cde.ca.gov/ta/tg/sr/blueprints.asp>
- Carnegie Learning. (2012). Retrieved from Carnegie Learning: <http://www.carnegielearning.com/secondary-solutions/adaptive-math/>
- Carpenter, S. K., Cepeda, N. J., Rohrer, D., Kang, S. H. K., & Pashler, H. (2012). Using spacing to enhance diverse forms of learning: Review of recent research and implications for instruction. *Educational Psychology Review*. Advance online copy.
- Carroll, J. B., Davies, P., & Richman, B. (1971). *Word frequency book*. Boston, MA: Houghton Mifflin.
- Carver, R. P. (1972). Speed readers don't read: they skim. *Psychology Today, 22*-30.
- Carver, R. P. (1974). Measuring the primary effect of reading: Reading storage technique, understanding judgments and cloze. *Journal of Reading Behavior, 6*, 249-274.
- Carver, R. P. (1976). Word length, prose difficulty, and reading rate. *Journal of Reading Behavior, 8*, 193-204.
- Carver, R. P. (1977). Toward a theory of reading and rauding. *Reading Research Quarterly, 13*, 8-63.

- 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 Reading Behavior*, 21(2), 155-166.
- Carver, R. P. (1990). *Reading rate: A review of research and theory*. New York: Academic Press.
- Carver, R. P. (1992a). Commentary: Effect of prediction activities, prior knowledge, and text type upon amount comprehended: Using reading theory to critique schema theory research. *Reading Research Quarterly*, 27(2), 165-174.
- Carver, R. P. (1992b). Reading rate: Theory, research, and practical implications. *Journal of Reading*, 36(2), 84-95.
- Carver, R. P. (2000). *The causes of high and low reading achievement*. Mahwah, NJ: Erlbaum.
- Carver, R. P., & Leibert, R. E. (1995). The effect of reading library books at different levels of difficulty upon gain in reading ability. *Reading Research Quarterly*, 30(1), 26-48.
- Chakraborty, B., Collins, L.M., Strecher, V.J., & Murphy, S.A. (2009). Developing multicomponent interventions using fractional factorial designs. *Statistics in Medicine*, 28(21), 2687-2708.
- Chapman, J. W., & Tunmer, W. E. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of Educational Psychology*, 87, 154-167.
- Charness, N. (1976). Memory for chess positions: Resistance to interference. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 641-653.
- Charness, N. (1979). Components of skill in bridge. *Canadian Journal of Psychology*, 33, 1-50.
- Charness, W. G., Krampe, R. T., & Mayer, U. (1996). The role of practice and coaching in entrepreneurial skill domains: An international comparison of life-span chess skill acquisition. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 51-80). Mahwah, NJ: Erlbaum.
- Charness, W. G., Tuffiash, M. I., Krampe, R., Reingold, E., & Vasyukova, E. (2005). The role of deliberate practice in chess expertise. *Applied Cognitive Psychology*, 19, 151-165.

- Chase, W. G. (Ed.). (1973). *Visual information processing*. New York: Academic Press.
- Chase, W. G. (1983). Spatial representations of taxi drivers. In D. R. Rogers, & J. H. Slobada (Eds.), *Acquisition of symbolic skills* (pp. 391-405). New York: Plenum.
- Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.
- Chavez-Oller, M. A., Chihara, T., Weaver, K. A., & Oller, J. W., Jr. (1985). When are cloze items sensitive to constraints across sentences? *Language Learning*, 35, 181-206.
- Chi, M. T. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 73-96). Hillsdale, NJ: Erlbaum.
- Chi, M. T. (1985). Interactive roles of knowledge and strategies in the development of organized sorting and recall. In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills: Research and open questions* (Vol. 2, pp. 457-483). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Chi, M. T. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 21-30). Cambridge, UK: Cambridge University Press.
- Chi, M. T., Bassok, M., Lewis, M., Reinmann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13, 145-182.
- Chi, M. T., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152.
- Chi, M. T., Glaser, R., & Farr, M. J. (Eds.). (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.
- Chihara, T., Oller, J. W., Jr., Weaver, K. A., & Chavez-Oller, M. A. (1977). Are cloze items sensitive to constraints across sentences? *Language Learning*, 27, 63-73.
- Christensen, C. M. (2011). *The innovator's dilemma: The revolutionary book that will change the way you do business*. New York, Harper Business.
- Church, M. A., Elliot, A., Gable, S. (2001). Perceptions of classroom environment, achievement goals, and achievement outcomes. *Journal of Educational Psychology*, 93, 43-54.
- Clancey, W. J., & Shortliffe, E. H. (1984). *Readings in medical artificial intelligence: The first decade*. Reading, MA: Addison-Wesley.

- Clarke, E. F. (1988). Generative principles in music performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 1-26). London: Oxford University Press.
- Clay, M. M. (2000). *Running records for classroom teachers*. Portsmouth, NH: Heinemann.
- Collins, L. M., Dziak, J. J., & Li, R. (2009). Design of experiments with multiple independent variables: A resource management perspective on complete and reduced factorial designs. *Psychological Methods*, 14(3), 202-224.
- Common Core State Standards. (2010). *Common Core State Standards for English: language arts & literacy in history/social studies, science, and technical subjects*. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers.
- Connecticut State Department of Education. (2012). Retrieved from <http://www.csde.state.ct.us/public/cedar/assessment/cmt/cmt.htm>
- Cowan, N., Chen, Z., & Rouder, J. (2004). Constant capacity in an immediate serial-recall task: A logical sequel to Miller (1956). *Psychological Science*, 15(9), 634-640.
- Cowley, M. (Ed.). (1958). *Writers at work: The Paris review interviews* (Vol. 1). New York: Viking Press.
- Cowley, M. (1959). *Writers at work: The Paris review interviews*. New York: Viking.
- Crain, S., & Shankweiler, D. (1988). Syntactic complexity and reading acquisition. In A. Davidson, & G. M. Green (Eds.), *Linguistic complexity and text comprehension: Readability issues reconsidered*. Hillsdale, NJ: Erlbaum Associates.
- Cronbach, L. J., & Furby, L. (1970). How should we measure "change" - or should we? *Psychological Bulletin*, 74, 68-80.
- Crowder, R. G., & Wagner, R. K. (1992). *The psychology of reading*. New York: Oxford University Press.
- CTB. (2012). Retrieved from <http://www.ctb.com/ctb.com/control/productFamilyViewAction?productFamilyId=449&p=products>
- Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920*. New York: Teachers College Press.
- Cuban, L. (2001). *Oversold & underused: Computers in the classroom*. Cambridge, MA: Harvard University Press.

- Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology, 82*, 733-740.
- Cunningham, A. E., & Stanovich, K. E. (1998). The impact of print exposure on word recognition. In J. Metsala, & L. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 235-262). Mahwah, NJ: Erlbaum.
- Dawes, R. M. (1994). *House of cards: Psychology and psychotherapy built on myth*. New York: Free Press.
- Dede, C. (2009). Immersive interfaces for engagement and learning. *Science, 323*, 66-69.
- Delaware Department of Education. (2012). Retrieved from http://www.doe.k12.de.us/aab/DSTP_gen_info/default.shtml
- Dempster, F. M. (1988). A case study in the failure to apply the results of psychological research. *American Psychologist, 43*, 627-634.
- Derogatis, L. R. (2000). The SCL-90-R and the Brief Symptom Inventory (BSI) in Primary Care. In M. E. Maruish (Ed.), *Handbook of psychological assessment in primary care settings*. Mahwah, NJ: Erlbaum.
- Dieterle, E. (2009). Neomillennial learning styles and River City. *Children, Youth, and Environments, 19*(1), 345-278.
- Diller, D. (2003). *Literacy work stations: Making centers work*. Stenhouse Publishers.
- Discovery. (2012). Retrieved from Discovery Education: <http://www.discoveryeducation.com/administrators/curricular-resources/science-techbook/#/Features>
- Doane, S. M., Pellegrino, J. W., & Klatzky, R. L. (1990). Expertise in computer operating system: Conceptualization and performance. *Human-Computer Interaction, 5*, 267-304.
- Dohm, A., & Shniper, L. (2007). *Occupational employment projections to 2016*. Washington, DC: U.S. Department of Labor, Bureau of Labor Statistics.
- Donovan, C., 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*, 309-333.

- Dreyfus, H. L. (1997). Intuitive, deliberative, and calculative models of expert performance. In C. E. Zsombok, & G. Klein (Eds.), *Naturalistic decision making* (pp. 17-28). Mahwah, NJ: Lawrence Erlbaum Associates.
- Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine: The power of human intuition and expertise in the age of the computer*. Oxford: Basil Blackwell.
- Dreyfus, H. L., & Dreyfus, S. E. (2005). Peripheral vision: Expertise in real world contexts. *Organizational Studies*, 26(5), 779-792.
- Dreyfus, S. E., & Dreyfus, H. L. (1980). *A five-stage model of the mental activities involved in directed skill acquisition*. Unpublished Report supported by the Air Force Office of Scientific Research No. Contract F49620-79-C-0063: University of California at Berkeley.
- Droop, M., & Verhoeven, L. (1998). Background knowledge, linguistic complexity, and second-language reading comprehension. *Journal of Literacy Research*, 30(2), 253-271.
- Dunphy, B. C., & Williamson, S. L. (2004). In pursuit of excellence. *Advances in Health Sciences Education*, 9, 107-127.
- Dziak, J., Nahum-Shani, I. R., Collins, L. M. (2012). Multilevel factorial experiments for developing behavioral interventions: Power, sample size, and resource considerations. *Psychological Methods*, 17, 153-175. Advance online publication.
- Eccles, J. S., Wigfield, A., Harold, R. D., & Blumenfeld, P. B. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, 64, 830-847.
- Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In N. Eisenberg (Ed.), *Handbook of child psychology* (Vol. IV, 5th ed, pp. 1017-1095). New York: John Wiley.
- Egan, D. E., & Schwartz, B. J. (1979). Chunking in recall of symbolic drawings. *Memory and Cognition*, 7, 149-158.
- Eggen, P., & Kauchak, D. (2004). *Educational psychology: Windows on the classroom* (6th ed.). Columbus, OH: Prentice Hall.
- Ehri, L. C. (1998). Word reading by sight and by analogy in beginning readers. In C. Hulme, & R. Joshi (Eds.), *Reading and spelling: Development and disorders* (pp. 87-111). Mahwah, NJ: Lawrence Erlbaum Associates.

- Eisenstadt, M., & Kareev, Y. (1979). Aspects of human problem solving: The use of internal representations. In D. A. Norman, & D. E. Rumelhart (Eds.), *Exploration in cognition*. San Francisco: W. H. Freeman.
- Ekwall, E. E. (1976). Informal reading inventories: The instructional level. *The Reading Teacher*, 29(7), 662-665.
- Endsley, M. (2006). Expertise and situation awareness. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 633-652). Cambridge, UK: Cambridge University Press.
- Engle, R. W., & Bukstel, L. (1978). Memory processes among bridge players of differing expertise. *American Journal of Psychology*, 91, 673-689.
- Ennis, P. H. (1965). *Adult book reading in the United States (National Opinion Research Center Report No. 105)*. Chicago: University of Chicago Press.
- Epstein, M. L., & Brosvic, G. M. (2002). Immediate Feedback Assessment Technique: Multiple-choice test that "behaves" like an essay examination. *Psychological Reports*, 90(1), 226.
- Epstein, M. L., Epstein, B. B., & Brosvic, G. M. (2001). Immediate feedback during academic testing. *Psychological Report*, 88, 889-894.
- Epstein, M. L., Lazarus, A. D., Calvano, T. B., Matthews, K. A., Hendel, R. A., Epstein, B. B., & Brosvic, G. M. (2002). Immediate Feedback Assessment Technique promotes learning and corrects inaccurate first responses. *Psychological Record*, 52(2), 187-201.
- Ericsson, K. A. (1996a). The acquisition of expert performance: An introduction to some of the issues. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 1-50). Mahwah, NJ: Erlbaum.
- Ericsson, K. A. (Ed.). (1996b). *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games*. Mahwah, NJ: Erlbaum.
- Ericsson, K. A. (2002). Attaining excellence through deliberate practice: Insights from the study of expert performance. In M. Ferrari (Ed.), *The pursuit of excellence in education* (pp. 21-55). Hillsdale, NJ: Erlbaum.
- Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine*, 10, S70-S81.

- Ericsson, K. A. (2006a). Introduction to Cambridge handbook of expertise and expert performance: Its development, organization, and content. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 3-19). Cambridge, UK: Cambridge University Press.
- Ericsson, K. A. (2006b). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 683-703). Cambridge, UK: Cambridge University Press.
- Ericsson, K. A., Charness, N., Feltovich, P., & Hoffman, R. R. (Eds.). (2006). *Cambridge handbook of expertise and expert performance*. Cambridge, UK: Cambridge University Press.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*(2), 211-245.
- Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363-406.
- Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence on maximal adaptations on task constraints. *Annual Review of Psychology*, *47*, 273-305.
- Ericsson, K. A., Patel, V. L., & Kintsch, W. (2000). How experts' adaptations to representative task demands account for the expertise effect of memory recall: Comment on Vicente and Wang (1998). *Psychological Review*, *107*, 578-592.
- Ericsson, K. A., & Smith, J. A. (1991a). Prospects and limits in the empirical study of expertise: An introduction. In K. A. Ericsson, & J. A. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 1-38). Cambridge, UK: Cambridge University Press.
- Ericsson, K. A., & Smith, J. A. (Eds.). (1991b). *Toward a general theory of expertise: Prospects and limits*. Cambridge, UK: Cambridge University Press.
- Estes, W.K. (1977). On the interaction of perception and memory in reading. In D. LaBerge and S. J. Samuels (Eds.) *Basic process in reading: Processes and comprehension*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- ETS. (2012). Retrieved from http://www.ets.org/gre/revised_general/about/content/pbt/
- Farbman, D., & Kaplan, C. (2005). *Time for a change: The promise of extended-time schools for promoting student achievement*. Boston, MA: Massachusetts 2020. Retrieved from http://www.educationsector.org/usr_doc/OntheClock.pdf

- Feltovich, P., Ford, K. M., & Hoffman, R. R. (Eds.). (1997). *Expertise in context: Human and machine*. Menlo Park: AAAI Press.
- Feltovich, P., Johnson, P. E., Moller, J., & Swanson, D. (1984). The role and development of medical knowledge in diagnostic reasoning. In W. Clancey, & E. Shortliffe (Eds.), *Readings in medical artificial intelligence: The first decade* (pp. 275-319). Reading, MA: Addison Wesley.
- Feltovich, P., Prietula, M., & Ericsson, K. A. (2006). Studies of expertise from psychological perspectives. In K. A. Ericsson, P. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 41-67). Cambridge, UK: Cambridge University Press.
- Feltovich, P., Spiro, R., & Coulson, R. (1997). Issues of expert flexibility in contexts characterized by complexity and change. In P. Feltovich, K. Ford, & R. R. Hoffman (Eds.), *Expertise in context: Human and machine*. Menlo Park, CA: AAAI Press.
- Fitts, P. M. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press.
- Fitts, P. M. (1967). *Human performance*. Belmont, CA: Brooks/Cole.
- Flavell, J. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, *34*, 906-911.
- Florida Department of Education. (2012). Retrieved from <http://fcats.fldoe.org/handbk/fcathandbook.asp>
- Fountas, I. C., & Pinnell, G. S. (1996). *Guided reading: Good first teaching for all children*. Portsmouth, NH: Heinemann.
- Fountas, I. C., & Pinnell, G. S. (2001). *Guiding readers and writers: Teaching comprehension, genre, and content literacy*. Portsmouth, NH: Heinemann.
- Fox, E. (2009). The role of reader characteristics in processing and learning from informational text. *Review of Educational Research*, *79*(1), 197-361.
- Fox, J. (2008). *Applied regression analysis and generalized linear models*. Sage.
- Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curve analysis. *Journal of Educational Psychology*, *88*(1), 3-17.
- Frederiksen, J. R. (1981). Understanding anaphora: Rules used by readers in assigning pronominal referents. *Discourse Processes*, *4*, 323-347.

- Fresch, M. J. (1995). Self-selected books of beginning readers: Standing before the smorgasbord. In M. Sorensen, & B. Lehman (Eds.), *Teaching with children's books: Paths to literature-based instruction* (pp. 121-128). Urbana, IL: National Council of Teachers of English.
- Fryer, R., & Levitt, S. (2004). Understanding the Black-White test score gap in the first two years of school. *The Review of Economics and Statistics*, 86, 447-464.
- Gambrell, L. B. (1984). How much time do children spend reading during reading instruction? In J. A. Niles, & L. A. Harris (Eds.), *Changing perspectives on research in reading/language processing and instruction* (pp. 127-135). Rochester, NY: National Reading Conference.
- Gambrell, L. B. (2007). Reading: Does practice makes perfect? *Reading Today*, 24(6), 16.
- Gambrell, L. B., Palmer, B. M., Codling, R. M., & Mazzoni, S. A. (1996). Assessing motivation to read. *The Reading Teacher*, 49, 2-19.
- Garner, R. &. (1989). Metacognition: Answered and unanswered questions. *Educational Psychologist*, 24, 143-158.
- Garner, R. (1987). *Metacognition and reading comprehension*. Norwood, NJ: Ablex.
- Gelman, A., & Rubin, D. B. (1995). Avoiding model selection in Bayesian social research. *Sociological Methodology*, 25, 165-195.
- Georgia Department of Education. (2012). Retrieved from <http://www.doe.k12.ga.us/Curriculum-Instruction-and-Assessment/Assessment/Pages/CRCT.aspx>
- Glaser, R. (1984). Education and thinking: the role of knowledge. *American Psychologist*, 39, 93-104.
- Glaser, R., & Chi, M. T. (1988). Overview. In M. T. Chi, R. Glaser, & J. Farr (Eds.), *The nature of expertise* (pp. xv-xxviii). Hillsdale, NJ: Erlbaum.
- Gobet, F., & Charness, N. (2006). Expertise in chess. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 523-538). Cambridge, UK: Cambridge University Press.
- Goldstein, H. (1995). *Multilevel statistical models* (2nd ed.). New York: Halstead Press.
- Good, R. H., & Kaminski, R. A. (2003). *DIBELS oral reading fluency passages for first through third grade (Technical report No. 10)*. Eugene, OR: University of Oregon.

- Goodman, K. S. (1965a). A linguistic study of cues and miscues in reading. *Elementary English*, 42, 639-643.
- Goodman, K. S. (1965b). Dialect barriers to reading comprehension. *Elementary English*, 42, 853-860.
- Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist*, 6, 126-135.
- Goodman, K. S. (1968). *The psycholinguistic nature of the reading process*. Detroit, MI: Wayne State University Press.
- Gough, P. B. (1972). One second of reading. In J. F. Kavanagh, & I. G. Mattingly (Eds.), *Language by ear and eye: The relationships between speech and reading*. Cambridge, MA: MIT Press.
- Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools - A report to Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education.
- Graham, S. E., Singer, J. D., & Willett, J. B. (2009). Modeling individual change over time. In R. E. Millsap, & A. Maydeu-Olivares (Eds.), *The SAGE handbook of quantitative methods in psychology* (pp. 615-636). Thousand Oaks, CA: SAGE.
- Greaney, V. (1980). Factors related to amount and time of leisure time reading. *Reading Research Quarterly*, 15, 337-357.
- Greaney, V., & Hegarty, M. (1987). Correlates of leisure-time reading. *Journal of Research in Reading*, 10, 3-20.
- Green, L. (2002). *How popular musicians learn*. London: Ashgate.
- Greene, J. P., & Forster, G. (2003). *Public high school graduation and college readiness rates in the United States (Education Working Paper No. 3)*. New York: Center for Civic Innovation at The Manhattan Institute.
- Greene, B. A., Miller, R. B., Crowson, M., Duke, B., & Akey, K. (2004). Predicting high school students' cognitive engagement and achievement: Contributions of classroom perceptions and motivation. *Contemporary Educational Psychology*, 29, 499-517.
- Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. In D. C. Berliner, & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 15-46). New York: Macmillan.

- Gutek, G. L. (1972). *A history of Western educational experience*. Prospect Heights, IL: Waveland Press.
- Guthrie, J. T., & Alvermann, D. (Eds.) (1999). *Engagement in reading: Processes, practices, and policy implications*. New York: Teachers College Press.
- Guthrie, J. T., McGough, K., Bennett, L., & Rice, M. E. (1996). Concept-oriented reading instruction to develop motivational and cognitive aspects of reading. In L. Baker, P. Afflerbach, & D. Reinking (Eds.), *Developing engaged readers in school and home communities* (pp. 165-190). Mahwah, NJ: Erlbaum.
- Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (3rd Ed.). New York: Longman.
- Hanlon, S. T., Swartz, C. W., Stenner, A. J., Burdick, H., & Burdick, D. S. (2012). *Learning Oasis*. Retrieved from www.alearningoasis.com
- Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*, 491-528.
- Harrell, F. E. (2001). *Regression Modeling Strategies*. New York: Springer.
- Harris, K. R., & Graham, S. (1996). *Making the writing process work: Strategies for composition and self-regulation*. Cambridge, MA: Brookline.
- Harvey, S., & Goudvis, A. (2007). *Strategies that work: Teaching comprehension for understanding and engagement* (2nd ed.). Portland, ME: Stenhouse Publishers.
- Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher*, *59*(7), 636-644.
- Hatano, G., Miyake, Y., & Binks, M. B. (1977). Performance of expert abacus operators. *Cognition*, *5*, 57-71.
- Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, *13*, 512-521.
- Hayes, J. R. (1981). *The complete problem solver*. Philadelphia, PA: Franklin Institute Press.
- Hedeker, D. R., & Gibbons, R. D. (2006). *Longitudinal Data Analysis*. Chicago, IL: Wiley-Interscience.
- Hedeker, D. R., Gibbons, R. D., & Flay, B. R. (1994). Random regression models for clustered data: With an example from smoking prevention research. *Journal of Clinical and Consulting Psychology*, *62*, 757-765.

- Helsen, W. F., Starkes, J. L., & Hodges, N. J. (1998). Team sports and the theory of deliberate practice. *Journal of Sport and Exercise Psychology*, *20*, 12-34.
- Henry, J. (2000). *Writing workplace cultures: An archaeology of professional writing*. Carbondale, IL: Southern Illinois University Press.
- Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research*, *60*, 549-571.
- Hiebert, E. H. (1983). An examination of ability grouping for reading instruction. *Reading Research Quarterly*, *18*, 231-255.
- Hiebert, E. H., & Raphael, T. E. (1996). Psychological perspectives on literacy and extensions to educational practice. In D. C. Berliner, & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 550-602). New York: Simon & Schuster Macmillan.
- Hiebert, E. H., Samuels, S. J., & Rasinski, T. (2012). Comprehension-based silent reading rates: What do we know? What do we need to know? *Literacy Research and Instruction*, 110-124.
- Hinsley, D., Hayes, J., & Simon, H. A. (1978). From words to equations: Meaning and representation in algebra word problems. In M. Just, & P. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, NJ: Erlbaum.
- Hodges, K. K., Kline, J., Stern, L., Cytryn, L., & McNew, D. (1982). The development of a child assessment interview for children: The Child Assessment Schedule (CAS). *Journal of Child Psychology and Psychiatry*, *31*, 427-436.
- Hodges, N. J., & Starkes, J. L. (1996). Wrestling with the nature of expertise: A sport specific test of Ericsson, Krampe, and Tesch-Romer's (1993) "Deliberate Practice". *International Journal of Sport Psychology*, *27*, 400-424.
- Hodges, N. J., Starkes, J. L., & MacMahon, C. (2006). Expert performance in sport: A cognitive perspective of superior expert performance. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 471-488). Cambridge, UK: Cambridge University Press.
- Hoffman, R. R. (Ed.). (1992). *The psychology of expertise*. New York: Springer-Verlag.
- Horn, J., & Masunaga, H. (2006). A merging theory of expertise and intelligence. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 587-612). Cambridge, UK: Cambridge University Press.

- Hox, J. (2002). *Multilevel analysis : Techniques and applications*. Mahwah, New Jersey: Erlbaum
- Huey, E. B. (1908/1968). *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press.
- Hulme, C., Snowling, M., & Quinlan, P. (1991). Connectionism and learning to read: Steps towards a psychologically plausible model. *Reading and Writing, 3*, 59-168.
- Hunt, A. (2008). *Pragmatic thinking & learning: Refactor your wetware*. Lubbock, TX: Pragmatic Bookshelf.
- Idaho State Department of Education. (2012). Retrieved from <http://www.sde.idaho.gov/site/assessment/ISAT/>
- Illinois State Board of Education. (2012). Retrieved from <http://www.isbe.state.il.us/assessment/isat.htm>
- International Reading Association. (2004). *Standards for reading professionals*. Newark, DE.
- Iowa Department of Education. (2012). Retrieved from <http://itp.education.uiowa.edu/ia/TestDescriptions.aspx>
- Jarvella, R. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior, 10*, 409-416.
- Johns, J., & Berglund, R. (2006). *Fluency strategies and assessments*. Dubuque, IA: Kendall/Hunt.
- Johnson, M. S., & Kress, R. A. (1965). *Informal reading inventories*. Newark, DE: International Reading Association.
- Johnson, P. E., Duran, A. S., Hassebrock, F., Moller, J. H., Prietula, M., Feltovich, P., & Swanson, D. B. (1981). Expertise and error in diagnostic reasoning. *Cognitive Science, 5*, 135-283.
- Jonz, J. (1987). Textual cohesion and second language comprehension. *Language Learning, 37*, 409-38.
- Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Boston: Allyn & Bacon.
- Kaiser Family Foundation (2009). Generation M2: Media in the lives of 8- to 18-year-olds. Retrieved from: <http://www.kff.org/entmedia/upload/8010.pdf>

- Kame'enui, E. J., Simmons, D. C., Chard, D., & Dickson, S. (1997). Directed reading instruction. In S. A. Stahl, & A. Hayes (Eds.), *Instructional models in reading* (pp. 59-84). Mahwah, NJ: Erlbaum.
- Kansas State Department of Education. (2012). Retrieved from <http://www.ksde.org/Default.aspx?tabid=159>
- Kellogg, R. T. (1986). Writing method and productivity of science and engineering faculty. *Research in Higher Education, 25*, 147-163.
- Kellogg, R. T. (2006). Professional writing experience. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 389-402). Cambridge, UK: Cambridge University Press.
- Kim, J. S. (2006). Effects of a voluntary summer reading intervention on reading achievement: Results from a randomized trial. *Educational Evaluation and Policy Analysis, 28*, 335-355.
- Kim, J. S. (2007). The effects of voluntary summer reading intervention on reading activities and reading achievement. *Journal of Educational Psychology, 99*, 505-515.
- Kim, J. S., & Guryan, J. (2010). The efficacy of a voluntary summer book reading intervention for low-income Latino children from language minority families. *Journal of Educational Psychology, 102*, 20-31.
- Kim, J. S., & White, T. G. (2008). Scaffolding voluntary summer reading for children in grades 3 to 5: An experimental study. *Scientific Studies of Reading, 12*(1), 1-23.
- Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Erlbaum.
- Kintsch, W. (1985). Text processing: A psychological model. In T. A. van Dijk (Ed.), *Handbook of discourse analysis* (Vol. 2, pp. 231-243). London: Academic Press.
- Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review, 95*, 163-182.
- Kintsch, W. (1994). The role of knowledge in discourse comprehension: A construction-integration model. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed., pp. 951-995). Newark, DE: International Reading Association.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

- Kintsch, W. (2004) The Construction-Integration model of text comprehension and its implications for instruction. In R. Ruddell & N. Unrau (Eds.) *Theoretical Models and Processes of Reading*. 5th Edition, International Reading Association.
- Kintsch, W., & Bates, E. (1977). Recognition memory for statements from a classroom lecture. *Journal of Experimental Psychology: Human Learning and Memory*, 3, 150-159.
- Kintsch, W., & van Dijk, T. A. (1978). Towards a model of text comprehension and production. *Psychological Review*, 85, 363-394.
- Kintsch, W., Patel, V. L., & Ericsson, K. A. (1999). The role of long-term working memory in text comprehension. *Psychologia*, 42, 186-198.
- Klare, G. R. (1963). *The measurement of readability*. Ames, IA: Iowa State University Press.
- Kleiman, G. M. (1975). Speech recoding in reading. *Journal of Verbal Learning and Verbal Behavior*, 14, 323-339.
- Klein, G. A. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.
- Knapp, M. S. (1995). *Teaching for meaning in high-poverty classrooms*. New York: Teachers College Press.
- Kolers, P. A. (1970). Three stages in reading. In H. Levin, & J. T. Williams (Eds.), *Basic studies in reading*. New York: Basic.
- Kopta, J. A. (1971). The development of motor skills in orthopedic education. *Clinical Orthopedics*, 75, 80-85.
- Koschmann, T. D., LeBaron, C., Goodwin, C., & Feltovich, P. (2001). Dissecting common ground: Examining an instance of reference repair. *Proceedings of the 23rd Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.
- Koskinen, P. S. (1993). Motivating independent reading and writing in the primary grades through social cooperative literacy experiences. *The Reading Teacher*, 47(2), 162-164.
- Krashen, S. (2004). False claims about literacy development. *Educational Leadership*, 61(6), 18-21.
- Kreft, I. G. G. & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.

- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*, 1121-1134.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Kuiper, R., & Pesut, D. (2004). Promoting cognitive and metacognitive reflective reasoning in skilled nursing practice: Self-regularities learning theory. *Journal of Advanced Nursing*, *45*, 381-391.
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293-323.
- Langan, D., & Athanasou, J. (2005). Testing a model of domain learning in music therapy. *Journal of Music Therapy*, *42*(2), 296-312.
- Lawrence, J. A. (1988). Expertise on the bench: Modeling magistrates' judicial decision-making. In M. T. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp. 229-259). Hillsdale, NJ: Erlbaum.
- Lehmann, A. C. (2002). Effort and enjoyment in deliberate practice: A research note. In I. M. Hanken, S. G. Nielsen, & M. Nerland (Eds.), *Research in and for music education. Festschrift for Harald Jorgensen* (pp. 153-166). Oslo: Norwegian Academy of Music.
- Lehmann, A. C., & Gruber, H. (2006). Music. In K. A. Ericsson, N. Chaness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 457-470). Cambridge, UK: Cambridge University Press.
- Lesgold, A. H., Rubinson, P., Feltovich, R., Glaser, D., Klopfer, S., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing x-ray pictures. In M. T. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp. 311-342). Hillsdale, NJ: Erlbaum.
- Lesgold, A. M., & Resnick, L. (1982). How reading difficulties develop: Perspectives from a longitudinal study. In J. Das, R. Mulcahey, & A. Wall (Eds.), *Theory and research in learning disabilities* (pp. 155-187). New York: Plenum Press.
- Levy, B. A. (1977). Reading: Speech and meaning processes. *Journal of Verbal Learning and Verbal Behavior*, *16*, 623-628.
- Levy, B. A. (1978). Speech processes during reading. In A. M. Lesgold, S. W. Pellegrino, S. W. Fokkema, & R. Glaser (Eds.), *Cognitive psychology and instruction*. New York: Plenum.

- Liberman, I. Y., Mann, V. A., Shankweiler, D., & Westelman, M. (1982). Children's memory for recurring linguistic and non-linguistic material in relation to reading ability. *Cortex*, *18*, 367-375.
- Liberman, I. Y., Shankweiler, D., Orlando, C., Harris, K. S., & Berti, F. B. (1971). Letter confusions and reversals of sequence in the beginning reader: Implications of Orton's theory of developmental dyslexia. *Cortex*, *7*, 127-142.
- Linn, R. L., & Slinde, J. A. (1977). The determination of the significance of change between pre and posttesting periods. *Review of Educational Research*, *47*, 121-150.
- Logan, G. (1985). Skill and automaticity: Relations, implications, and future directions. *Psychological Review*, *95*, 492-527.
- Lumley, D. (1991). Improving student motivation. *Electronic Learning*, *11*(3), 14.
- MacKinnon, J. (1993). Becoming a rhetor: Developing writing ability in a mature, writing intensive organization. In R. Spilka (Ed.), *Writing in the workplace: New research perspectives* (pp. 41-55). Carbondale, IL: Southern Illinois University Press.
- Maine Department of Education. (2012). Retrieved from <http://www.maine.gov/education/necap/index.html>
- Manis, F., Seidenberg, M., Doi, L., McBride-Chang, C., & Peterson, A. (1996). On the basis of two subtypes of developmental dyslexia. *Cognition*, *58*, 157-195.
- Markham, P. L. (1985). The rational deletion cloze and global comprehension in German. *Language Learning*, *35*, 423-430.
- Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology*, *81*, 417-430.
- Massachusetts Department of Elementary and Secondary Education. (2012). Retrieved from <http://www.doe.mass.edu/mcas/tdd/ela.html?section=testdesign>
- Mathewson, G. C. (1994). Model of attitude influence upon reading and learning to read. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed., pp. 1131-1161). Newark, DE: International Reading Association.
- Matthews, M. R. (1994). *Science teaching: The role of history and philosophy in science*. New York: Routledge.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Mayer, R. E. (2003). *Learning and instruction*. Upper Saddle River, NJ: Prentice Hall.

- Mayer, R. E. (2005). Introduction to multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 1-16). New York: Cambridge University Press.
- Mayer, R. E. (2008). Foreword: Seven criteria for investigating children's learning. In T. Willoughby & E. Wood (Eds.), *Children's learning in a digital world* (pp.1-8). Malden, MA: Blackwell Publishing.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1-86.
- McClelland, J. L., & Johnston, J. C. (1977). The role of familiar units in perception of words and nonwords. *Perception and Psychophysics, 22*, 249-261.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1, an account of basic findings. *Psychological Review, 88*, 373-403.
- McClenaghan, B. A., & Ward, D. S. (1987). Health and physical education. In C. A. Maher, & S. G. Forman (Eds.), *A behavioral approach to education of children and youth (School Psychology Series)* (pp. 131-151). Hillsdale, NJ: Erlbaum.
- McKeithen, K. B., Reitman, J. S., Reuter, H. H., & Hirtle, S. C. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology, 13*, 307-325.
- McKenna, M. C., Kear, D. J., & Ellsworth, R. A. (1995). Children's attitudes toward reading: A national survey. *Reading Research Quarterly, 30*, 934-955.
- McKeown, M. G., Beck, I. L., Omanson, R. C., & Pople, M. T. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly, 20*, 522-535.
- McKinney, E. H., & Davis, K. .. (2004). Effects of deliberate practice on crisis decision performance. *Human Factors, 45*, 436-444.
- McKinney, E. H., & Davis, K. .. (2004). Effects of deliberate practice on crisis decision performance. *Human Factors, 45*, 436-444.
- McMorris, T. (2004). *Acquisition and performance of sports skills*. Hoboken, NJ: John Wiley & Sons.
- Meece, J. L., Anderman, E. M., & Anderman, L. H. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology, 57*, 487-503.

- Mellenbergh, G. J. (1989). Item bias and item response theory. *International Journal of Educational Research*, 13, 127-143.
- Menard, S. (2002). *Applied logistic regression*. Sage.
- Meredith, W., & Millsap, R. E. (1992). On the misuse of manifest variables in the detection of measurement bias. *Psychometrika*, 57, 289-311.
- Mesmer, H. A. (2008). *Tools for matching readers to texts*. New York: Guilford Press.
- MetaMetrics, Inc. (2008). *Texas Higher Education Coordinating Board: Text measurement and analysis*. Durham, NC: MetaMetrics, Inc.
- Mewhort, D. J., & Campbell, A. J. (1981). Toward a model of skilled reading: An analysis of performance in tachistoscopic tasks. In G. E. MacKinnon, & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (Vol. 3, pp. 39-118). New York: Academic.
- Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of psychological refractory-period phenomena. *Psychological Review*, 104, 749-791.
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90, 227-234.
- Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1974). Functions of phonemic and graphemic codes in visual word recognition. *Memory and Cognition*, 2, 309-321.
- Meyer, L. M., & Wardrop, J. L. (1994). Home and school influences on learning to read in kindergarten through second grade. In F. Lehr, & J. Osborn (Eds.), *Reading, language, and literacy: Instruction for the twenty-first century* (pp. 165-184). Hillsdale, NJ: Erlbaum.
- Miller, G. A. (1956). The magic number seven, plus or minus 2: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81-97.
- Miller, G. A., & Gildea, P. M. (1987). How children learn words. *Scientific American*, 257(3), 94-99.
- Miller, G. A., & Isard, S. (1963). Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, 2, 217-228.
- Miller, G. A., Bruner, J. S., & Postman, L. (1954). Familiarity of letter sequences and tachistoscopic identification. *Journal of Genetic Psychology*, 50, 129-139.

- Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart & Winston.
- Miller, J. R., & Kintsch, W. (1980). Readability and recall of short prose passages: A theoretical analysis. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 335-354.
- Millsap, R. E. (2010). Testing measurement invariance using item response theory in longitudinal data: An introduction. *Child Development Perspectives*, 4(1), 5-9.
- Mississippi Department of Education (2013). Mississippi Assessment and Accountability Reporting System. Retrieved from http://orsap.mde.k12.ms.us/MAARS/maarsMS_TestResultsProcessor.jsp?userSessionId=42&EmbargoAccess=0&DistrictId=2040&TestPanel=1&BrowserType=2
- Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology*, 85, 424-436.
- Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, 24, 237-270.
- Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, 20, 233-253.
- National Education Technology Plan (2010). *Transforming American Education: Learning Powered by Technology*. U.S. Department of Education, Office of Educational Technology. Retrieved from: <http://www.ed.gov/sites/default/files/netp2010.pdf>
- National Endowment for the Arts. (2007). *To read or not to read: A question of national consequence (Research Report No. 47)*. Washington, DC.
- 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.
- Nell, V. (1988). The psychology of reading for pleasure: Needs and gratification. *Reading Research Quarterly*, 23, 6-50.
- Nelson, J., Perfetti, C., Liben, D., & Liben, M. (2011). *Measures of text difficulty: Testing their predictive value for grade levels and student performance*.

- Newell, A. (1973). Artificial intelligence and the concept of mind. In R. C. Schank, & K. M. Colby (Eds.), *Computer models of language and thought* (pp. 1-60). San Francisco: W. H. Freeman.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.
- Noble, K. G., Farah, M. J., & McCandliss, B. D. (2006). Socioeconomic background modulates cognition-achievement relationships in reading. *Cognitive Development, 21*, 349-368.
- Nolen, S. B. (2001). Learning environment, motivation, and achievement in high school science. *Journal of Research in Science Teaching, 40*, 347-368.
- Nolen, S. B., & Haladyna, T. M. (1990). Personal and environmental influences on students' beliefs about effective study strategies. *Contemporary Educational Psychology, 15*, 116-130.
- Norris, D. (1994). A quantitative model of reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 1212-1232.
- North Carolina Department of Public Instruction. (2012). Retrieved from <http://www.dpi.state.nc.us/accountability/testing/eog/reading/>
- Obama, B. (2011). *Address before a joint session of the Congress on the state of the union, January 25, 2011*. Retrieved from GPO Access database: <http://www.gpo.gov/fdsys/pkg/CREC-2011-01-25/pdf/CREC-2011-01-25-pt1-PgH457-6.pdf>
- O'Dwyer, L. M., Russell, M., & Bebell, I. D. (2006). Identifying teacher, school and district characteristics associated with middle and high school teachers use of technology: A multilevel perspective. *Journal of Educational Computing Research, 33*(4), 369-393.
- Oldfather, P., & Wigfield, A. (1996). Children's motivations to read. In L. Baker, P. Afflerbach, & D. Reinking (Eds.), *Developing engaged readers in school and home communities* (pp. 89-113). Mahwah, NJ: Erlbaum.
- Page, M. (2002). Technology-enriched classrooms: Effects of low socioeconomic status. *Journal of Research on Technology in Education, 34*(4), 389-398.
- Paradis, J., Dobrin, D., & Miller, R. (1985). Writing at Exxon ITD: Notes on the writing environment of an R&D organization. In L. Odell, & D. Goswami (Eds.), *Writing in nonacademic settings* (pp. 281-307). New York: Guilford.

- Paris, S., & Winograd, P. (1990). How metacognition can promote academic learning and instruction. In B. Jones, & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 15-51). Hillsdale, NJ: Erlbaum.
- Park, Y. (2011). How motivational constructs interact to predict elementary students' reading performance: Examples from attitudes and self-concept in reading. *Learning and Individual Differences, 21*(4), 347-358.
- Partnership for Assessment of Readiness for College and Career (2013). About PARCC. Retrieved March 26, 2013: <http://www.parcconline.org/about-parcc>
- Patel, V. L., & Groen, G. J. (1991). The general and specific nature of medical expertise: A critical look. In K. A. Ericsson, & J. Smith (Eds.), *Toward a general theory of expertise* (pp. 93-125). Cambridge, MA: Cambridge University Press.
- Patterson, K. E., & Coltheart, V. (1987). Phonological processes in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*. Hillsdale, NJ: Erlbaum.
- Pearson, P. D. (2013). Research foundations of the Common Core State Standards in English language arts. In S. Neuman & L. Gambrell (Eds.), *Quality reading instruction in the age of the Common Core State Standards* (pp. 237-262). Newark, DE: International Reading Association.
- Pellegrino, J., Chudowsky, J. N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- Pennington, H. (2006). *Expanding learning time in high schools*. Washington, DC: Center for American Progress. Retrieved from http://www.americanprogress.org/issues/2006/10/pdf/extended_learning_report.pdf
- Pennsylvania Department of Education. (2012). Retrieved from http://www.portal.state.pa.us/portal/server.pt/community/pennsylvania_system_of_school_assessment_%28pssa%29/8757/resource_materials/507610
- Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.
- Perfetti, C. A., Bell, L. C., & Delaney, S. M. (1988). Automatic (prelexical) phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language, 27*, 1-22.
- Phillips, D. C. (1995). The good, the bad, and the ugly: The many faces of constructivism. *Educational Researcher, 24*(7), 5-12.

- Pillsbury, W. B. (1897). A study of apperception. *American Journal of Psychology*, 8, 315-393.
- Pinheiro, J. C., & Bates, D. M. (1995). Approximations to the log-likelihood function in nonlinear mixed-effects models. *Journal of Computational and Graphical Statistics*, 4, 12-35.
- Pinheiro, J. C. & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*, Springer.
- Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research*, 63, 167-199.
- Pintrich, P. R., & Schunk, D. H. (1996). *Motivation in education: Theories, research, and application*. Englewood Cliffs, NJ: Merrill/Prentice Hall
- Plant, E. A., Ericsson, K. A., Hill, L., & Asberg, K. (2005). Why study time does not predict grade point average across college students: Implications of deliberate practice for academic performance. *Contemporary Educational Psychology*, 30, 96-116.
- Plimpton, G. (Ed.). (1963). *Writers at work: The Paris Review interviews, second series*. New York: Penguin.
- Plimpton, G. (Ed.). (1977). *Writers at work: The Paris review. Interviews, Second Series*. New York: Penguin.
- Porter, D. (1983). The effect of quantity of context on the ability to make linguistic predictions: A flaw in a measure of general proficiency. In A. Hughes & D. Porter (Eds.), *Current developments in language testing* (pp. 63-74). London: Academic Press.
- Posner, M. I. (1988). Introduction: What is it to be an expert? In M. T. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp. xxix-xxxvi). Hillsdale, NJ: Erlbaum.
- Pressley, M., & McCormick, C. B. (1995). *Advanced educational psychology for educators, researchers, and policymakers*. New York: Harper Collins.
- Princeton Review. (2012). Retrieved from <http://www.princetonreview.com/college/sat-sections.aspx>
- Raaijmakers, J. G., & Shiffrin, R. M. (1981). A theory of probabilistic search of associative memory. *Psychological Review*, 88, 93-134.
- Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 25, 111-163.

- Raskin, E. (1936). Comparison of scientific and literary ability: A biographical study of eminent scientists and letters of the nineteenth century. *Journal of Abnormal and Social Psychology*, 31, 20-35.
- Rathbun, A., & West, J. (2004). *From kindergarten through third grade: Children's beginning school experiences*. NCES (2004-007). National Center for Education Statistics. Washington, D.C.: U.S. Department of Education.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods* (2nd ed.). Thousand Oaks, CA: Sage.
- Raufaste, E., Eyrolle, H., & Marine, C. (1998). Pertinence generation in radiological diagnosis: Spreading activation and the nature of expertise. *Cognitive Science*, 22(4), 517-546.
- Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 7, 65-81.
- Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice-Hall.
- Rayner, K., Foorman, B., Perfetti, C., Pesetsky, D., & Seidenberg, M. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, 2(2), 31-74.
- Raz, I. S., & Bryant, P. (1990). Social background, phonological awareness and children's reading. *British Journal of Developmental Psychology*, 8(3), 209-225.
- Reder, L., & Shunn, C. (1996). Metacognition does not imply awareness: Strategy choice is governed by implicit learning and memory. In L. Reder (Ed.), *Implicit memory and metacognition*. Mahwah, NJ: Erlbaum.
- Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 81, 274-280.
- Reitman, J. (1976). Skilled perception in GO: Deducing memory structure from inter-response times. *Cognitive Psychology*, 8, 336-356.
- Renninger, K. A. (1992). Individual interest and development: Implications for theory and practice. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 361-395). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Reznick, R. K., & MacRae, H. (2006). Teaching surgical skills – changes in the wind. *New England Journal of Medicine*, 355, 2664-2669.

- Rieger, M. (2004). Automatic keypress activation in skilled typists. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 555-565.
- Roeser, R. W., Midgley, C., & URdan, T. C. (1996). Perceptions of the school psychological environment and early adolescents' psychological and behavioral functioning in school: The mediating role of goals and belonging. *Journal of Educational Psychology*, 88, 408-422.
- Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin*, 90, 726-748.
- Rogosa, D. R., & Willett, J. B. (1985). Understanding correlates of change by modeling individual differences in growth. *Psychometrika*, 50, 203-228.
- Rosenblatt, L. M. (1978). *The reader, the text, the poem: The transactional theory of literacy work*. Carbondale, IL: Southern Illinois University Press.
- Rosenblatt, L. M. (1983). *Literature as exploration* (4th ed.). New York: Modern Language Association. (Original work published 1938).
- Rosenblatt, L. M. (2004). The transactional theory of reading and writing. In R. B. Ruddell, & N. J. Unrau (Eds.), *Theoretical Models and Processes of Reading* (5th ed., pp. 1363-1399). Newark, DE: International Reading Association.
- Ross, K. G., Shafer, J. L., & Klein, G. (2006). Professional judgments and “naturalistic decision making”. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 403-420). Cambridge, UK: Cambridge University Press.
- Ruddell, R. B., & Ruddell, M. R. (1995). *Teaching children to read and write: Becoming an influential teacher*. Needham Heights, MA: Allyn & Bacon.
- Ruddell, R. B., Ruddell, M. R., & Singer, H. (Eds.). (1994). *Theoretical models and processes of reading*. Newark, DE: International Reading Association.
- Ruddell, R. B., & Unrau, N. J. (2004). Section three introduction. In R. B. Ruddell, & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1116-1126). Newark, DE: International Reading Association.
- Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance VI* (Vol. 6, pp. 573-603). Hillsdale, NJ: Erlbaum.
- Rumelhart, D. E. (1994). Toward an interactive model of reading. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed., pp. 864-894). Newark, DE: International Reading Association.

- Rumelhart, D. E. (2004). Toward an interactive model of reading. In R. B. Ruddell, & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1149-1179). Newark, DE: International Reading Association.
- Rumelhart, D. E., & McClelland, J. L. (1986a). On learning the past tenses of English verbs. In J. L. McClelland, & D. E. Rumelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2). Cambridge, MA: MIT Press/Bradford Books.
- Rumelhart, D. E., & McClelland, J. L. (1986b). *Parallel distributed processing. Explorations in the microstructure of cognition. Vol. 1: Foundations*. Cambridge: MIT Press.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533-536.
- Russell, M., Bebell, I. D., & O'Dwyer, L. M. (2005). Tracking the arc of new teachers' technology use. In C. Vrasidas & G. V. Glass (Eds.), *Current perspectives on applied information technologies: Preparing teachers to teach with technology*. Greenwich, CT: Information Age Publishing, Inc.
- Russell, M., Bebell, I. D., O'Dwyer, L. M., & O'Connor, K. M. (2003). Examining teacher technology use: Implications for pre-service and in-service teacher preparation. *Journal of Teacher Education*, 54(4), 297-310.
- Russell, M., O'Dwyer, L. M., Bebell, D., & Tao, W. (2007). How teachers' uses of technology vary by tenure and longevity. *Journal of Educational Computing Research*, 37(4), 393-417.
- Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, 2, 437-442.
- Salas, E., Rosen, M. A., & Diaz-Granados, D. (2010). Expertise-based intuition and decision making in organizations. *Journal of Management*, 36(4), 941-973.
- Samuels, S. J. (2002). Reading fluency: Its development and assessment. In A. Farstrup, & S. J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 166-183). Newark, DE: International Reading Association.
- Samuels, S. J. (2004). Toward a theory of automatic information processing in reading, revisited. In R. B. Ruddell, & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1127-1148). Newark, DE: International Reading Association.
- Samuels, S. J., & Kamil, M. L. (1984). Models of the reading process. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 185-224). New York: Longman.

- Samuels, S. J., & Turnure, J. (1974). Attention and reading achievement in first grade boys and girls. *Journal of Educational Psychology*, 66, 29-32.
- Scalise, K. & Gifford, B. (2006). Computer-based assessment in e-learning: A framework for constructing “intermediate constraint” questions and tasks for technology platforms. *Journal of Technology, Learning, and Assessment*, 4(6).
- Schirmer, A., & Lockman, B. R. (2001). How do I find a book to read? Middle and high school students use a rubric for self-selecting material for independent reading. *Teaching Exceptional Children*, 34(1), 36-42.
- Schmidt, H. G., & Boshuizen, H. P. (1993). On acquiring expertise in medicine. *Educational Psychology Review*, 5(3), 205-221.
- Schnabel, K. U., Little, T. D., & Baumert, J. (2008). Modeling longitudinal and multilevel data. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling Longitudinal and Multilevel Data* (pp. 9-13). New York: Psychology Press.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: 1. Detection, search, and attention. *Psychological Review*, 84, 1-66.
- Schooler, C. (1990). Psychosocial factors and effective cognitive functioning in adulthood. In E. Birren, & W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 347-358). San Diego, CA: Academic Press.
- Schrader, P. G., Lawless, K., & Hayley, M. (2008). The model of domain learning as a framework for understanding internet navigation. *Journal of Educational Multimedia and Hypermedia*, 17(2), 235-258.
- Schulz, R., & Curnow, C. (1998). Peak performance and age among superathletes: Track and field, swimming, baseball, tennis, and golf. *Journal of Gerontology: Psychological Sciences*, 43, 113-120.
- Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychological Science*, 12, 101-108.
- Schvaneveldt, R. W., & Meyer, D. E. (1973). Retrieval and comparison processes in semantic memory. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic.
- Schwartz, B., & Reisberg, D. (1991). *Learning and memory*. New York: Norton.
- Schwartz, B., & Robbins, S. J. (1995). *Psychology of learning and behavior* (4th ed.). New York: Norton & Company.

- Schwarz, G. (1978). Estimating the dimensions of a model. *Annals of Statistics*, 6, 461-464.
- Seidenberg, M. S. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 245-263). Hillsdale, NJ: Erlbaum.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523-568.
- Seidenberg, M. S., Tannenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, 14, 489-537.
- Shaffer, L. H. (1975). Multiple attention in continuous verbal tasks. In P. M. Rabbitt, & S. Dornic (Eds.), *Attention and performance V*. New York: Academic Press.
- Shaffer, L. H. (1981). Performance of Chopin, Bach, and Bartok: Studies in motor programming. *Cognitive Psychology*, 13, 326-376.
- Shanahan, T. (1983). The informal reading inventory and the instructional level: The study that never took place. In L. Gentile, M. L. Kamil, & J. Blanchard (Eds.), *Reading research revisited* (pp. 577-580). Columbus, OH: Merrill.
- Shanahan, T. (2004). Critiques of the National Reading Panel report: Their implications for research, policy, and practice. In P. McCardle, & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 235-265). Baltimore, MD: Brookes.
- Shanahan, T. (2011). *Rejecting instructional level theory*. Retrieved from Shanahan on Literacy: <http://www.shanahanonliteracy.com/2011/08/rejecting-instructional-level-theory.html>
- Shankweiler, D., & Crain, S. (1986). Language mechanisms and reading disorder: A modular approach. *Cognition*, 14, 139-168.
- Shannon, P. (1990). *The struggle to continue: Progressive reading instruction in the United States*. Portsmouth, NH: Heinemann.
- Sharon, A. T. (1973). What do adults read? *Reading Research Quarterly*, 9, 148-169.
- Shavelson, R. J., & Towne, L. (2002). *Scientific research in education*. Washington, DC: National Academy Press.
- Shipp, S. (2004). The brain circuitry of attention. *Trends in Cognitive Sciences*, 8, 223-230.
- Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist*, 61, 394-403.

- Simonton, D. K. (1997). Creative productivity: A predictive and explanatory model of career trajectories and landmarks. *Psychological Review*, *104*, 66-89.
- Silva, E. (2007). *On the clock: Rethinking the way schools use time*. Washington, DC: Education Sector
- Singer, C. (1958). *From magic to science*. New York: Dover.
- Singer, J. D., & Willett, J. B. (2003). *Applied Longitudinal Data Analysis*. Oxford: Oxford University Press.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Free Press.
- Skinner, B. F. (1969). *Contingencies of reinforcement: A theoretical analysis*. Englewood Cliffs, NJ: Prentice Hall.
- Slavin, R. E. (1997). *Educational psychology: Theory and practice* (5th ed.). Needham Heights, MA: Allyn & Bacon.
- Sloboda, J. A. (1996). The acquisition of musical performance expertise: Deconstructing the "talent" account of individual differences in musical expressivity. In K. A. Ericsson (Ed.), *The road to excellence: the acquisition of expert performance in the arts and sciences, sports, and games* (pp. 81-106). Mahwah, NJ: Lawrence Erlbaum.
- Sloboda, J. A., Davidson, J. W., Howe, M. J., & Moore, D. G. (1996). The role of practice in the development of performing musicians. *British Journal of Psychology*, *87*, 287-309.
- Smarter Balanced Assessment Consortium (2013). About SBAC. Retrieved March 26, 2013 from: <http://www.smarterbalanced.org/about/>
- Smith, F. (1971). *Understanding reading: A psycholinguistic analysis of reading and learning to read*. New York: Holt, Rinehart, & Winston.
- Smith, F. (1973). *Psycholinguistics and reading*. New York: Holt, Rinehart & Winston.
- Snijders, T. A., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.
- Snow, C. E., Burns, W. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Cambridge, MA: Harvard University Press.
- Sonnentag, S., Niessen, C., & Volmer, J. (2006). Expertise in software design. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 373-388). Cambridge: Cambridge University Press.

- South Carolina State Department of Education. (2012). Retrieved from <http://ed.sc.gov/agency/programs-services/45/PASSELATests.cfm>
- Spilich, G. J., Vesonder, G. T., Chiesi, H. L., & Voss, J. F. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior*, *14*, 506-522.
- Spoehr, K. T. (1981). Word recognition in speech and reading: Toward a theory of language processing. In P. D. Eimas, & J. L. Miller (Eds.), *Perspectives on the study of speech*. Hillsdale, NJ: Erlbaum.
- Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, *56*(1), 72-110.
- Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, *16*, 32-71.
- Stanovich, K. E. (1984). The interactive-compensatory model of reading: A confluence of developmental, experimental, and educational psychology. *Remedial and Special Education*, *5*, 11-19.
- Stanovich, K. E. (1986). Mathew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21* (4), 360-407.
- Stanovich, K. E. (1992). *How to think straight about psychology* (3rd ed.). New York: Guilford Press.
- Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford Press.
- Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, *24*, 402-433.
- Stanovich, K. E., West, R. F., Cunningham, A. E., Cipelewski, J., & Siddiqui, S. (1996). The role of inadequate print exposure as a determinant of reading comprehension problems. In C. Cornoldi, & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 15-32). Mahwah, NJ: Erlbaum.
- Starkes, J. L., & Allard, F. (Eds.). (1993). *Cognitive issues in motor expertise*. Amsterdam: North Holland.
- Starkes, J. L., & Ericsson, K. A. (Eds.). (2003). *Expert performance in sport: Recent advances in research on sport expertise*. Champaign, IL, Human Kinetics.

- Starkes, J. L., Deakins, J., Allard, F., Hodges, N. J., & Hayes, A. (1996). Deliberate practice in sports: What is it anyway? In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 81-106). Mahwah, NJ: Erlbaum.
- Steier, D., & Mitchell, T. (Eds.). (1996). *Knowledge matters: A tribute to Allen Newell*. Mahwah, NJ: Erlbaum.
- Stenner, A. J. (2003). *Matching students to text: The targeted reader (Scholastic Professional Paper)*. New York: Scholastic. Retrieved from http://www.tomsnyder.com/reports/Target_Reader_Web2.pdf
- Stenner, A. J., & Stone, M. H. (2004). Does the reader comprehend the text because the reader is able or because the text is easy? *Paper presented at the meeting of the International Reading Association*. Reno-Tahoe, NV.
- Stenner, A. J., Burdick, H., Sanford, E. E., & Burdick, D. S. (2006). How accurate are Lexile text measures? *Journal of Applied Measurement*, 7(3), 307-322.
- Stenner, A. J., Burdick, H., Sanford, E. E., & Burdick, D. S. (2007). *The Lexile Framework for Reading Technical Report*. Durham, NC: MetaMetrics, Inc.
- Sternberg, R. J. (1996). *Cognitive psychology*. Forth Worth, TX: Holt, Rinehart, & Winston.
- Stevens, A. L., & Rumelhart, D. E. (1975). Errors in reading: Analysis using an augmented transition network model of grammar. In D. A. Norman, & D. E. Rumelhart (Eds.), *Explorations in cognition* (pp. 136-156). San Francisco: Freeman.
- Stevenson, H. W., Chen, C., Uttal, D. H. (1990). Beliefs and achievement: A study of black, white, and Hispanic children. *Child Development*, 61, 508-523.
- Stillwell, R. (2010). *Public school graduates and dropouts from the common core of data: School year 2007-08 (NCES 2010-341)*. Retrieved from National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education: <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2010341>
- Sum, A. (1999). *Literacy in the labor force: Results from the National Adult Literacy Survey*. Washington, DC: Office of the Assistance Secretary of Defense, U.S. Department of Defense.
- Swan, K., van't Hooft, M., Kratcoski, A., & Unger, D. (2005). Uses and effects of mobile computing devices in K-8 classrooms. *Journal of Research on Technology in Education*, 38(1), 99-113.

- Swartz, C. W., Burdick, D. S., Hanlon, S. T., Stenner, A. J., Kyngdon, A., Burdick, H., & Smith, M. (2011). Toward a theory relating text complexity, reading ability, and reading comprehension. *American Education Research Association*.
- Tannenhaus, M. K., Flanigan, H., & Seidenberg, M. S. (1980). Orthographic and phonological code activation in auditory and visual word recognition. *Memory and Cognition*, 8, 513-520.
- Taylor, S. E. (1965). Eye movements in reading: Facts and fallacies. *American Educational Research Journal*, 2, 187-202.
- Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, 30, 415-433.
- Texas Education Agency. (2012). Retrieved from <http://www.tea.state.tx.us/student.assessment/taks/blueprints/>
- Thomas, R. M. (1996). *Comparing theories of child development* (4th ed.). Pacific Grove, CA: Brooks/Cole.
- Thorndike, E. L. (1903). *Educational psychology*. New York: Lemcke & Buechner.
- Thorndike, E. L. (1917). Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology*, 8, 323-332.
- Thorndike, E. L. (1932). *The fundamental of learning*. New York: Teachers College Press.
- Thurlow, M., Gaden, J., Ysseldyke, J., & Algozzine, R. (1984). Student reading during reading class: The lost activity in reading instruction. *Journal of Educational Research*, 77(5), 267-272.
- Tierney, R. J. (1994). Dissensions, tensions, and the models of literacy. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed., pp. 1162-1182). Newark, DE: International Reading Association.
- Timmermans, A. A., Seelen, H. A., Willmann, R. D., & Kingma, H. (2009). Technology-assisted training of arm-hand skills in stroke: Concepts on reacquisition of motor control and therapist guidelines for rehabilitation technology design. *Journal of Neuroengineering and Rehabilitation*, 6, 1-18.
- Toppino, T. C., & Schneider, M. A. (1999). The mix-up regarding mixed unmixed lists in the spacing-effect research. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 1071-1076.
- Tracey, D. H., & Morrow, L. M. (2006). *Lenses on Reading*. New York: Guilford Press.

- Tulving, E., & Gold, C. (1963). Stimulus information and contextual information as determinants of tachistoscopic recognition of words. *Journal of Experimental Psychology*, 66, 319-327.
- Tulving, E., Mandler, G., & Bauml, R. (1964). Interaction of two sources of information in tachistoscopic word recognition. *Canadian Journal of Psychology*, 18, 62-71.
- U.S. Census Bureau (2013). State and county QuickFacts. Retrieved from: <http://quickfacts.census.gov/qfd/states/00000.html>
- U.S. Department of Agriculture (2007). National School Lunch Program. Retrieved from: http://www.fns.usda.gov/cga/factsheets/school_meals.htm
- U.S. Department of Education (2013). *Expanding evidence approaches for learning in a digital world*. Office of Educational Technology, Washington, DC.
- U.S. Department of Labor. (1991). *What work requires of schools: A SCANS report for America 2000*. Retrieved from The Secretary's Commission on Achieving Necessary Skills (SCANS): <http://wdr.doleta.gov/SCANS/whatwork/whatworks.pdf>
- Ulijaszek, S. J., Johnston, F. E., & Preece, M. A. (Eds.). (1998). *The Cambridge encyclopedia of human growth and development*. Cambridge, UK: Cambridge University Press.
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Van Orden, G. C. (1991). Phonological mediation is fundamental to reading. In D. Besner, & G. W. Humphreys (Eds.), *Basic processes in reading* (pp. 77-103). Hillsdale, NJ: Erlbaum.
- Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review*, 97, 488-522.
- VanSledright, B. (2002). *In search of America's past: Learning to read history in elementary school*. New York: Teachers College Press.
- VanSledright, B., & Alexander, P. A. (2002). *Historical knowledge, thinking, and beliefs: Evaluation component of the Corps of Historical Discovery Project (#S215X010242)*. Washington, DC: United States Department of Education.
- Vaughn, S., Moody, S. W., & Schumm, J. S. (1998). Broken promises: Reading instruction in the resource room. *Exceptional Children*, 64, 211-225.

- Vendlinski, T., & Stevens, R. (2002). Assessing student problem-solving skills with complex computer-based tasks. *Journal of Technology, Learning, and Assessment*, 1(3).
<http://www.jtla.org>.
- Venezky, R. L. (1984). The history of reading research. In P. D. Pearson, R. Barr, M. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (Vol. 2, pp. 46-47). New York: Longman.
- Vidulich, M. A., Wickens, C. D., Tsang, P. S., & Flach, J. M. (2010). Information processing in aviation. In E. Salas, & D. E. Maurino (Eds.), *Human factors in aviation* (2nd ed., pp. 175-210). Burlington, MA: Academic Press.
- Virginia Department of Education. (2012). Retrieved from
http://www.doe.virginia.gov/testing/sol/released_tests/2010/released_tests2010.shtml
- Voss, J. F., Greene, T. R., Post, T. A., & Penner, B. C. (1983). Problem solving skill in the social sciences. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research theory* (Vol. 17, pp. 165-213). New York: Academic Press.
- Voss, J. F., Tyler, S., & Yengo, L. (1983). Individual differences in the solving of social science problems. In R. Dillon, & R. Schmeck (Eds.), *Individual differences in cognition* (pp. 205-232). New York: Academic Press.
- Wagner, R. K., & Stanovich, K. E. (1996). Expertise in reading. In K. A. Ericsson (Ed.), *The road to excellence* (pp. 189-225). Mahwah, NJ: Lawrence Erlbaum Associates.
- Walberg, H. J., & Tsai, S. (1984). Reading achievement and diminishing returns to time. *Journal of Educational Psychology*, 76, 442-451.
- Wandel, J. C. (2003). Reflections on the impact of Patricia Benner's Work. *Medscape Nurses*, 5(2).
- Ware, J. H. (1985). Linear models for the analysis of longitudinal studies. *American Statistician*, 39, 95-101.
- Waters, G., Caplan, D., & Hildebrandt, N. (1987). Working memory and written sentence comprehension. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 531-555). Hillsdale, NJ: Erlbaum.
- Weber, R. M. (1970). First graders' use of grammatical context in reading. In H. Levin, & J. T. Williams (Eds.), *Basic studies in reading*. New York: Basic.
- West, L. H., & Pines, A. L. (1985). *Cognitive structures and conceptual change*. New York: Academic Press.

- White, K. R. (1982). The relation between socio-economic status and academic achievement. *Psychological Bulletin, 91*, 461-481.
- White, S., & Clement, J. (2001). *Assessing the Lexile Framework: Results of a panel meeting (NCES 2001-08)*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- Whitehurst, G. J. (1997). Language processes in context: Language learning in children reared in poverty. In L.B. Adamson & M. A. Ronski (Eds.), *Research on communication and language disorders: Contribution to theories of language development* (pp. 233-266). Baltimore: Brookes.
- Wiederholt, J. L., & Bryant, B. R. (1986). *Gray Oral Reading Tests-Revised*. Los Angeles: Western Psychological Services.
- Wigfield, A., Eccles, J. S., Rodriguez, D. (1998). The development of children's motivation in school contexts. In P. D. Pearson & A. Iran-Nejad (Eds.), *Review of research in education* (Vol. 23, pp. 73-118). Washington, DC: American Educational Research Association.
- Wigfield, A., & Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology, 89*, 420-432.
- Willett, J. B. (1988). Questions and answers in the measurement of change. In E. Rothkopf (Ed.), *Review of research in education (1988-89)* (pp. 345-422). Washington, DC: American Educational Research Association.
- Willett, J. B. (1989). Some results on reliability for the longitudinal measurement of change: Implications for the design of studies of individual growth. *Educational and Psychological Measurement, 49*, 587-602.
- Williamson, G. L. (2008). A text readability continuum for postsecondary readiness. *Journal of Advanced Academics, 19*(4), 602-632.
- Williamson, G. L., Appelbaum, M., & Epanchin, A. (1991). Longitudinal analysis of academic achievement. *Journal of Educational Measurement, 28*, 61-76.
- Williamson, G. L., Fitzgerald, J., Stenner, A. J. (2013). The Common Core State Standards' quantitative text complexity trajectory: Figuring out how much complexity is enough. *Educational Researcher, 42*(2), 59-69.
- Winne, P. H. (1982). Minimizing the black box problem to enhance the validity of theories about instructional effects. *Instructional Science, 11*, 13-28.

- Winner, E. (1996). The rage to master: The decisive role of talent in the visual arts. In K. A. Ericsson (Ed.), *The road to excellence* (pp. 271-302). Mahwah, NJ: Erlbaum.
- Wirt, J., Choy, S., Rooney, P., Provasnik, S., Sen, A., & Tobin, R. (2004). *The condition of education 2004 (NCES 2004-077)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics. Retrieved from <http://nces.ed.gov/pubs2004/2004077.pdf>
- Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading*, 5(3), 211-238.
- Wood, E., Mueller, J., Willoughby, T., Specht, J., & Deyoung, T. (2005). Teachers' perceptions: barriers and supports to using technology in the classroom. *Education, Communication, & Information*, 5(2), 183-206.
- Woodcock, R. W. (1987). *Woodcock Reading Mastery Tests-Revised*. Circle Pines, MN: American Guidance Service.
- Woolfolk, A. E. (1998). *Educational psychology* (7th ed.). Boston: Allyn & Bacon.
- Worthy, J., & McKool, S. (1996). Students who say they hate to read: The importance of opportunity, choice, and access. In D. J. Leu, C. K. Kinzer, & K. A. Hinchman (Eds.), *Literacies for the 21st century: Research and practice. 45th yearbook of the National Reading Conference* (pp. 245-256). Chicago: National Reading Conference.
- Wright, B. D., & Stenner, A. J. (1998). Readability and reading ability. *Paper presented at the meeting of the Australian Council on Educational Research*. Melbourne, Australia.
- Wright, B. D., & Stone, M. H. (2004). *Making measures*. Chicago, IL: Phaneron Press.
- 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.
- Zahorik, J. A. (1987). Reacting. In M. J. Dunkin (Ed.), *International encyclopedia of teaching and teacher education* (pp. 416-423). Oxford: Pergamon Press.
- Zarrillo, J., & Cox, C. (1992). Efferent and aesthetic teaching. In J. Many & C. Cox(Eds.), *Stance and literary understanding: Exploring the theories, research, and practice* (pp. 239-245). Norwood, NJ: Ablex.
- Zeitz, C. (1997). Some concrete advantages of abstraction: How experts' representations facilitate reasoning. In P. J. Feltovich, K. M. Ford, & R. R. Hoffman (Eds.), *Expertise in context: Human and machine* (pp. 43-65). Menlo Park, CA: AAAI/MIT Press.

- Zimmerman, B.J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*, 3-17.
- Zorzi, M., Houghton, G., & Butterworth, B. (1998a). Two routes or one in reading aloud? *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1131-1161.
- Zorzi, M., Houghton, G., & Butterworth, B. (1998b). The development of spelling-sound relationships in a model of phonological reading. *Language and Cognitive Processes, 13*, 337-371.
- Zuber, B. L., & Wetzell, P. A. (1981). Eye movement determinants of reading rate. In B. L. Zuber (Ed.), *Models of oculomotor behavior and control* (pp. 193-208). Boca Raton, FL: CRC Press.