USING EYE TRACKING TO OBSERVE DIFFERENTIAL EFFECTS OF REPEATED READINGS FOR LOW- AND HIGH-PERFORMING READERS

by

ANDREA M. ZAWOYSKI

(Under the Direction of Scott P. Ardoin)

ABSTRACT

Repeated Readings (RR) is an evidence-based fluency intervention in which students read the same text multiple times. Currently, it remains unknown whether readers of different skill levels benefit in different ways from RR. Eye tracking provides a means to examine intervention effects more closely because it permits measurement of subtle changes that occur during RR. The current study measured changes in underlying reading behavior of low-performing (n=22) and high-performing 2nd graders (n=22). Participants read a grade-level passage 4 times in a single session while their eye movements were recorded. Findings replicated previous research, suggesting that both groups benefited from RR. Additionally, results implied that effects were greater for low-performing readers, although they were typically unable to match levels of eye movement efficiency exhibited in the high-performing readers' first reading. Findings have implications for improving future eye tracking reading research with children and the efficiency of RR in the classroom.

INDEX WORDS: Elementary students, Eye tracking, Eye movements, Low-performing readers, High-performing readers, Reading, Repeated reading, Word frequency

USING EYE TRACKING TO OBSERVE DIFFERENTIAL EFFECTS OF REPEATED READINGS FOR LOW- AND HIGH-PERFORMING READERS

by

ANDREA M. ZAWOYSKI

B.S., University of Florida, 2011

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF ARTS

ATHENS, GEORGIA

2013

USING EYE TRACKING TO OBSERVE DIFFERENTIAL EFFECTS OF REPEATED READINGS FOR LOW- AND HIGH-PERFORMING READERS

by

ANDREA M. ZAWOYSKI

Major Professor: Scott P. Ardoin
Committee: A. Michele Lease
Kevin M. Ayres

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2013

ACKNOWLEDGEMENTS

This research was funded by the Institute of Education Sciences, U.S. Department of Education, via Grant R305A100496 to the University of Georgia. The opinions expressed are those of the author and do not represent views of the Institute or of the U.S. Department of Education.

I would like to acknowledge all the individuals who supported me throughout this project. First, I must express my sincerest gratitude to my advisor, Dr. Scott Ardoin. I appreciate how he made each revision a learning opportunity and always found the time to answer my questions. I could not have completed this thesis without his patience, support, and feedback. I am also grateful for the assistance of Dr. Kathy Binder, Carolyn Nemier, and the students in Dr. Binder's lab for their valuable insight and guidance. Furthermore, I would like to acknowledge my committee members, Drs. Michele Lease and Kevin Ayres, and thank them for their time and helpful feedback. Finally, I must express gratitude to my family, my friends, and my boyfriend, for their unwavering encouragement and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Reading fluency is a key component of proficient reading and an essential area of instruction for early elementary readers (NICHD, 2000). Despite efforts to foster a nation of proficient readers, data from the National Center for Education Statistics show that within the last decade, the percentage of fourth grade students reading at or above the proficient level increased by merely 5%, with 33% still reading below the basic level (Aud et al., 2012). These disappointing findings underscore the need for a better understanding of the specific effects that intervention has on students' reading. Knowing how interventions work will help teachers make more appropriate instructional modifications so that a greater proportion of students can develop adequate reading skills.

Repeated Readings

A well-known reading intervention with a substantial research base is repeated readings (RR), which fundamentally requires students to reread a text until they reach a criterion (Chard, Vaughn, & Tyler, 2002; Therrein, 2004). The foundational principles of RR are derived from LaBerge and Samuels' (1974) theory of automaticity. This theory suggests that readers must achieve automaticity in lower-level processes (e.g., decoding and word recognition) before they can allocate enough of their attention to higher-level processes (e.g., comprehension), to permit sufficient understanding of the text. When lower-level skills are not automatic, readers must expend a majority of their cognitive resources on deciphering individual words, which reduces their attention to the meaning of the text. It is thought that repeated practice of a text improves readers' automaticity with the decoding and recognition of words within the text, thus permitting them to focus their attention on the meaning of the text (Huey, 1908/1968; Samuels, 1979; Samuels, 2006).

Findings from applied research on RR demonstrate that repeated practice can improve reading for myriad populations, including skilled readers, (Ardoin, Morena, Binder, & Foster, submitted; Levy, Nicholls, & Kohen, 1993; Sindelar, Monda, & O'Shea, 1990), readers with learning disabilities (Chard et al., 2002), and all students through fourth grade (NICHD, 2000). In a meta-analysis of RR, Therrein (2004) analyzed mean improvements on reread passages and reported that the increase in effect size was .83 for fluency and .67 for comprehension. Despite strong support for the effectiveness of RR, it remains unknown whether the procedure improves fluency and comprehension for all readers equally, or if the effects are divided differentially, such that lower-performing readers may benefit more in word recognition, whereas higher-performing readers benefit more in comprehension (NICHD, 2000). For example, Levy et al. (1993) found evidence that overall, RR resulted in improved reading rates, comprehension, and error detection (i.e., detecting both nonwords and real words that did not make sense in context), for high- *and* low- performing third, fourth, and fifth graders on passages that were below, at, and above their grade level. However, differences between low- and high-performing readers were observed in measures of lower- and high-level processing. Specifically, on measures of lower-level processing (e.g., detecting nonword errors), high-performing readers reached a ceiling after the initial readings, whereas low-performing readers continued to improve on lower-level processes throughout all four practice readings. In contrast, on measures of higher-level processing (e.g., comprehension questions, words out of context) readers in both groups improved across all readings. Overall, low- and high-performing readers exhibited similar patterns of improvement throughout repetitions, but the magnitude of improvement was greater for the low-performing readers. According to Levy et al., this finding suggests that with sufficient practice, low-performing readers could achieve acceptable levels of fluency.

Similarly, research examining generalization passages implies that the effects of RR may vary based on skill level. Using students' achievement scores, Faulkner and Levy (1994) identified groups of low- and high-performing readers and asked them to read pairs of easy and difficult passages. Intervention effects were then examined by assessing students' reading of a second passage, which was either: (a) the same passage, (b) a passage with high word overlap, (c) a passage with high content overlap, or (d) a passage which lacked word and content similarity. Not surprisingly, students' reading fluency improved the most when they reread the same passage and improved the least when they read the passage lacking word and content similarity to the first passage. Interesting findings were observed in the word and content overlap passages, which revealed that relative text difficulty impacted RR effects. When passages were presumably easy for students to read (e.g., when high-performing students read grade-level passages), their reading times did not decrease significantly on word overlap passages. In contrast, their reading times did significantly decrease when they read high content overlap passages. Faulkner and Levy suggested that when students read passages that were easy for them, they were able to focus attention on comprehension as opposed to word recognition. Therefore, repeated practice with the same words in a different context (i.e., word overlap condition) was not likely to facilitate transfer. Rather, repeated practice with similar context (i.e., high content overlap condition) was more beneficial. In contrast, when passages were difficult for students to read (e.g., when low-performing students read grade-level passages), benefits were observed on both high word and high content overlap passages. Furthermore, when high-performing students read upper grade-level passages which were presumably difficult for them to read, they made significant improvements on high word and high content overlap passages, just as their low-performing peers did when they read grade-level (i.e., relatively challenging) passages. These results were largely replicated in a follow-up study involving fourth grade students and undergraduates (Faulkner & Levy, 1999).

Together, findings from Faulkner and Levy (1994, 1999) imply that RR generally improves reading, but difficulty level of the practice passage, relative to the student, may affect the magnitude or type of the benefits. A notable limitation to Faulkner and Levy (1994, 1999) is that conclusions about how skill and difficulty level impacted text processing were based on data collected from transfer passages, which may have added variability as a result of different words and content. Furthermore, the RR condition involved only one repetition, which does not reflect recommended RR procedures (Samuels, 1979; Therrein, 2004).

A limitation present in the aforementioned studies is that researchers examined outcome measures (e.g., word reading rate and accuracy), as opposed to the changes in reading behavior that occur *during* each reading to produce differential outcomes. In fact, despite substantial evidence supporting RR's effectiveness in improving reading fluency (Therrein, 2004), researchers have yet to determine what specifically causes readers to benefit from RR, and whether the benefits are differentiated on the basis of skill level. Fortunately, recent improvements in eye tracking technology permit data collection on students' eye movements during RR (Rayner, Ardoin, & Binder, 2013), thus allowing for direct observation of minute changes in reading behavior that are otherwise impracticable to measure reliably.

Eye Movements and Eye Tracking Dependent Measures

Although readers may feel as if their eyes move fluidly across text, eye tracking research reveals that during reading, the eyes actually make a series of quick jumping movements, called *saccades*, as well as pauses, called *fixations* (Rayner, 1998; Rayner, Chace, Slattery, & Ashby, 2006). The eyes also make backward saccades, or *regressions*. Regressions can be *inter-word* (i.e., regression to previously read text) or *intra-word* (i.e., regression within a single word). Eye trackers can record count and duration of these eye movements during reading. Each measurement is thought to represent different aspects of reading behavior (e.g., decoding, comprehension). (See Table 1 for detailed information). Using these measures, Ashby, Rayner, and Clifton (2005) found that arduous reading is characterized by longer fixations on words, shorter saccades, and more frequent regressions. In contrast, efficient reading is characterized by shorter fixations, longer saccades, and less frequent regressions.

Eye movements of children. Research on developmental patterns in children's eye movements during reading began in the early 20th century (e.g., Buswell, 1922; Huey, 1908/1968). Changes across age groups are thought to reflect improved efficiency in reading. Specifically, saccade length increases, fixations become shorter and less frequent, and word skipping increases (Buswell, 1922; McConkie et al., 1991; Rayner, 1986; Taylor, 1965). Also, findings reveal that regardless of grade, elementary students make approximately the same number of regressions, but with age, regression type gradually shifts from intra-word to inter-word (McConkie et al., 1991), perhaps indicating a shift from word-level to passage-level reading. Despite nearly a century of research documenting children's developmental changes in eye movements during reading, little was known about how eye movements in children change as a function of RR until very recently.

Eye tracking studies on rereading with adults. The majority of eye tracking research examining the effects of rereading a passage has measured the eye movements of skilled adult readers. Findings from this literature indicate that rereading facilitates processing and results in faster reading of a text and more efficient eye movements (Hyona & Niemi, 1990; Levy & Burns, 1990; Shebilske & Fisher, 1980). Interestingly, adults remain sensitive to word frequency, meaning that even with repeated practice, low-frequency words require more processing time than high-frequency words (Raney & Rayner, 1995).

Despite evidence that adults benefit from rereading, it is inappropriate to generalize findings from adult eye tracking research to the understanding of reading behavior in early elementary readers during RR. This is primarily because adults and children process text in different ways (Faulkner & Levy, 1999; Rayner, 1986), which likely reflects differences in their purposes for reading. For example, adults may reread for the purposes of retaining information or redirecting attention to the text, whereas beginning readers may reread for the purposes of practicing and achieving fluency in lower-level processes such as word recognition or decoding. Thus, in order to gain a better understanding of how RR improves fluency in early elementary readers, it is essential to directly examine reading behavior in this population.

Eye tracking studies on RR with early elementary students. Foster, Ardoin, and Binder (2013) examined changes in the eye movements of second grade students during RR. Reading behavior was analyzed globally (i.e., across the passage) and on low- and high-frequency target words. Consistent with the adult research on rereading (Hyönä & Niemi 1990; Raney & Rayner, 1995; Shebilske & Fisher, 1980), data gathered from students' eye movements (i.e., passage reading time, first fixation duration, gaze duration, total fixation time, average number of fixations per word, and number of regressions) revealed that RR resulted in overall more efficient reading. Overall, findings from Foster et al. indicated that RR reduced the amount of time required for students to process both individual words and meaning of the text. Furthermore, RR resulted in immediate effects on measures indicative of lower-level processing (i.e., first fixation duration, gaze duration) and continued effects for measures indicative of higher-level processing (i.e., total fixation time, number of inter-word regressions, average number of fixations per word). These findings provide further support for LaBerge and Samuels' (1974) theory of automaticity, as RR seemingly helped second graders develop enough automaticity in word-level reading to permit them to focus more attention on passage-level processing.

Foster et al. (2013) identified several similarities in eye movements between adults during rereading and children during RR, such as significant effects of word frequency and reduced overall reading time. However, it was the differences between adults' and children's eye movements which permitted a better understanding of how RR benefits early elementary readers. Specifically, children's gaze duration and total fixation time decreased across readings for low-frequency target words only, whereas findings from adult literature indicate decreases across these dependent variables for low- *and* high-frequency target words (Raney & Rayner, 1995). These differences suggest that RR primarily improves children's reading efficiency on *low-frequency words* and reduces the amount of additional processing time required for them to read text (Foster et al., 2013). Findings from Foster et al. were largely replicated in a follow-up study examining the impact of RR on target words presented in generalization passages, with effects observed only on low-frequency, not high-frequency words (Ardoin, Binder, Zawoyski, Foster, & Blevins, in press). Together, these studies provide a preliminary understanding of the behaviors that change while students engage in RR, which ultimately lead to improvements in outcome measures (e.g., WRCM). Unfortunately, the implications are limited because analyses grouped all students together regardless of skill level, therefore making it difficult to ascertain whether RR improved oral reading rate with accuracy for high- and low-performing readers in the same manner. However, findings from eye tracking research conducted with average-performing readers and high-performing readers suggest that significant differences between readers of different skill levels are likely. For example, Valle, Binder, Walsh, Nemier, and Bangs (2013) measured eye movements of average- and high-performing early elementary readers during one reading of a text. Findings indicated that high-performing readers made fewer and shorter fixations than the average-performing readers. Average-performing readers made almost twice as many intra-word regressions as high-performing readers, signifying that they had more difficulty with decoding and word recognition.

Given that the National Reading Panel (NRP) suggested high- and low-performing readers may benefit differently from RR (NICHD, 2000, p. 3-3—3-4), failure to address this question prohibits a comprehensive understanding of how RR improves students' oral reading rate with accuracy. Furthermore, eye tracking research on rereading with adults (Rayner et al., 2006; Raney & Rayner, 1995) and applied RR research with children (Faulkner & Levy, 1994, 1999; Levy et al., 1993) suggest that intervention effects may differ depending upon a reader's skill level. Despite preliminary evidence (e.g., Ardoin et al., in press, Foster et al., 2013), eye tracking research has not yet examined potential differential effects of RR for low- and high-performing early elementary-aged students.

Purpose and Hypotheses

The purpose of the current study was to extend the RR and eye tracking literature by evaluating reading behavior of low- and high-performing readers during RR on a grade-level passage. Changes in eye tracking measures reflective of reading processes were examined across four readings of a text. As in Foster et al. (2013) and Ardoin et al. (in press), this study examined the effects of RR on reading behavior in typically-developing second grade readers, with the added benefit of separating a more academically diverse sample into groups of low- and high-performing readers.

Due to differences in the low- and high-performing readers' skill levels, it was hypothesized that findings would reflect between groups differences in overall benefits of RR and general characteristics of eye movements during RR. First, regardless of skill level, both groups were expected to benefit in some manner from RR. However, high-performing readers were not expected to require all four readings to attain the maximum benefits of RR, whereas low-performing readers were expected to significantly improve between all readings. Additionally, throughout RR, high-performing readers were expected to exhibit more efficient eye movements than low-performing readers (e.g., they would make fewer and shorter fixations, less intra- and inter-word regressions).

On global analyses, it was hypothesized that high-performing readers would improve on measures associated with higher-level processing (i.e., average fixation count, number of regressions, total fixation time) in fewer readings than low-performing readers. Also, RR was expected to help low-performing readers make enough progress throughout readings that their eye movement efficiency in the final reading would be comparable with the high-performing readers' eye movement efficiency in their initial reading. Given this expectation, it was hypothesized that the magnitude of improvement would be greater for low-performing readers.

In target word analyses, low- and high-performing readers were expected to remain sensitive to word frequency effects, meaning that low-frequency target words would require greater amounts of processing time, relative to high-frequency target words. Findings were expected to replicate Foster et al. (2013), in that RR would reduce additional text processing requirements most significantly for low-frequency target words, although the magnitude of effect would be greater for low-performing readers. Similar to expectations for global analyses, high-performing readers were expected to make significant improvement in measures associated with higher-level processing (i.e., average fixation count, number of regressions, total fixation time) in fewer readings than low-performing readers.

Table 1

Eye Movement Terminology and Dependent Measures Defined (McConkie et al., 1991; Rayner et al., 2006)

Terminology

Fixation: pause between saccades, 300 ms or longer for early elementary readers; new information is encoded during this time

Saccade: rapid, forward eye movements during which vision is suppressed and new information cannot be encoded

Regression: a backward saccade

Inter-word regression: regression between words, from the current word to a previously read word

Intra-word regression: regression within a single word

Dependent Measures

<u>Measures Expected to Represent Lower-Level Processing</u>

First fixation duration: length of time required for the first fixation on a word before a saccade is made to another word or part of that same word

Gaze duration: duration of all fixations made on a word *before* a saccade is made to another word

Number of Intra-word regressions: total count of regressions made *within* words

<u>Measures Expected to Represent Higher-Level Processing</u>

Total fixation time: duration of *all* fixations made on a word, including fixations that occur after a regression is made to that word

Average fixation count: average number of fixations made on each word (for global analyses) or on each target word (for target word analyses)

Number of Inter-word regressions: total count of regressions made *between* words

CHAPTER 2

METHOD

Participants and Settings

Participants were 44 second grade students (20 males, 24 females). Mean age was 7 years, 11 months (range = 7 years, 4 months to 8 years, 11 months). The majority of students in the sample were Caucasian (81%). The remaining students were identified as multiracial (9%), Asian (5%), or Black (5%). Participants attended one of three public suburban schools in the southeastern part of the US. Free and reduced lunch rates across the schools ranged from 18%-29%. Most second grade students at these schools met state standards for Reading (91-100%) and English/Language Arts (87%-95%).

Participants were selected from a group of students who were part of a larger study which involved pretesting, 10 weeks of intervention or assignment to a control group, and post testing. Data for the current study were collected in the post testing period. Due to requirements for the larger study, no participants in the current study received special education, gifted, or English as second language services. In terms of reading skill level, participants in the current study represented the bottom 25% (low-performing readers, n = 22) and top 25% (high-performing readers, n = 22) of the sample for the larger study. Rank order was determined by each student's median oral reading fluency scores (ORF) on Formative Assessment Instrumentation and Procedures for Reading (FAIP-R) curriculum-based measurement probes (Christ, Ardoin, Monaghen, Van Norman, & White, 2013). On average, the low-performing readers attained FAIP-R median ORF scores of 74 words read correctly in a minute (WRCM), (range = 43

WRCM- 92 WRCM) whereas the high-performing readers attained FAIP-R median ORF scores of 144 WRCM, (range = 127 WRCM -224 WRCM).

All students completed four subtests of the Woodcock-Johnson Tests of Achievement – Third Edition, Form A (WJ-III ACH; Woodcock, McGrew, & Mather, 2001). Broad Reading composite standard scores (SS) fell within the average range for low-performing readers (mean SS=106, range SS = 95-111) and the average to high average range for high-performing readers (mean SS=117, range SS= 106-126).

As FAIP-R and WJ-III ACH scores indicate, low-performing readers were not normatively low-performing. Rather, they were low-performing in relation to the high-performing students in the sample. Participants' FAIP-R scores suggest that low-performing readers were generally reading grade-level passages at the instructional level, whereas high-performing readers were reading grade-level passages at the mastery level. ,

Apparatus

Eye movement data were collected with an SR Research EyeLink 1000 system, which has a sampling rate of 1000 Hz, a resolution of 0.01 degrees of visual angle, and a range of 32 degrees horizontally and 25 degrees vertically. In this system, eye movements are recorded by a desktop-mounted camera positioned directly in front of a computer screen, which was either a 19” (48.26 cm) ViewSonic VG930m or a 22” (55.88cm) ViewSonic VX2268wm LCD display monitor. The camera recorded movements from only one eye, yet participants' viewing was binocular throughout the assessment. Typically, the system records movements from the right eye, but data were collected from the left eye in the event of right eye tracking difficulties.

Experimenters verified via tape measure that the camera was positioned at the recommended tracking distance, approximately 50-55 cm from the chin rest where participants placed their chins during tracking. The brightness and height of the monitor were attuned prior to assessments. Participants indicated that they finished reading each passage and responded to comprehension questions by pressing the left and right analog trigger keys on a Microsoft Sidewinder Plug and Play game pad.

Materials

The stimulus was an experimenter-created passage modeled after stories from second and third grade basal readers. It consisted of four paragraphs with 16 sentences and 162 words. The Spache (1953) readability estimate was 2.88. Black text appeared on the computer monitor screen against a white background in 20-point Times New Roman font with 1.5 line spacing. The 16 sentences of the passage spanned 13 lines, with length ranging from 18-87 characters. Distributed across the text were 5 low-frequency target words and 5 high-frequency target words. Word frequency of the target words was determined using *The American Heritage Word Frequency Book* (Carroll, Davies, & Richman, 1971), which defines words of low frequency as appearing 10 or fewer times per million words of text and words of high frequency as appearing 40 or more times per million words of text.

Procedure

Two examiners trained in eye tracking procedures conducted the eye tracking assessments, which were held in a quiet classroom. Participants completed the assessment individually, which consisted of a single eye tracking session. Upon entering the room, they sat facing the monitor and placed their heads on a chin rest intended to minimize head movements during tracking. One examiner positioned participants and explained testing procedures while the other examiner adjusted the eye tracking camera and ensured that the computer settings met threshold requirements. Then, calibration and validation commenced using a nine point grid presented in the form of a child-friendly "follow the dot game," in which participants were instructed to look at fixation dots that appeared in various locations on the screen. The calibration process informs the eye tracking system where participants' eyes are located when they are looking at specific points on the screen. Calibration is followed by validation, in which the fixation dots are presented again in the same locations to ensure that the eye tracking system can reliably identify the participants' eye positions.

Successful calibration and validation were followed by a practice trial intended to familiarize participants with assessment procedures (i.e., reading text off a screen and answering questions with the game pad). Next, an examiner informed participants that they would need to read several passages silently and urged them to do their best reading. Participants were also told that the experimenters could not provide assistance to them when reading and that they would be required to answer a comprehension question following each passage. Calibration and validation procedures were then repeated and upon success, the experimenter presented a fixation dot in the upper left hand corner of the screen. The first passage was presented after participants fixated on the fixation dot. Participants read the passage, pressed a button on the game pad to indicate they had finished reading, and then responded to a comprehension question with the game pad. This completed the first trial.

Passages and comprehension questions were presented in the same fashion for a total of six trials. Data for the current study were gathered from the RR passage that was presented in the third, fourth, fifth, and sixth trials. Between the second and third trials, examiners presented additional instructions about RR procedures prior to participants' first reading of the RR passage. After reading each RR passage, participants responded to a comprehension question. Examiners then informed participants of how long they had taken to read the passage.

Each participant's session lasted approximately 15-20 min, unless technical difficulties and/or a participant's excessive body movement necessitated recalibration. Participants were allowed short breaks between readings if they seemed fatigued. After post testing was complete, participants selected a small toy as a reward for participation in the study.

Data Analyses

The effects of RR on eye movement parameters were examined via repeated measures mixed analyses of variance (ANOVAs). Analyses were conducted at the global level (i.e., across all words in the passage) and at the target word level (i.e., on low- and high- frequency target words) across groups of low- and high-performing readers. Multivariate statistics (i.e., Wilks' Lambda) were reported when Mauchly's test indicated a violation of sphericity. Log(x) transformations were conducted to correct extreme outliers (i.e., z-score > 3.29) or violations of the assumption for homogeneity of variance. Bonferroni-corrected follow-up analyses were conducted to examine significant interactions between skill level and/or readings and target word frequency. Overall effect sizes for intervention are reported as partial eta squared, which correspond to Cohen's (1988) benchmark *f* values indicating small ($\eta_p^2 = .0099$), medium ($\eta_p^2 = .0588$), and large effects ($\eta_p^2 =.1379$). Cohen's *d* was calculated to represent effect size between two readings in pairwise comparisons for global analyses. For all global and target word analyses, significant interactions that pertained to differences between low- and high-performing readers were followed up with pairwise comparisons and *t*-tests.

CHAPTER 3

RESULTS

Global Analyses

Global measures included first fixation duration, gaze duration, total fixation time, number of inter-word regressions, number of intra-word regressions, and average fixation count (per word). A 4 (readings) x 2 (skill level) mixed ANOVA was conducted for each variable. Due to extreme outliers and violations of Levene's test for homogeneity of variance, log (x) transformations were conducted for gaze duration, total fixation time, number of intra-word regressions, and average fixation count per word. Assumptions for normality were met following log(x) transformations, with the exception of one data point in average fixation count. All variables met assumptions for the homogeneity of variance after transformation, except for number of intra-word regressions. Further attempts to achieve homogeneity of variance through transformations (i.e., square root and reciprocal methods) also failed. Although findings for this measure must be interpreted with caution, changes between readings occurred in the expected direction. Of note, there were no differences in statistically significant findings between non-transformed and transformed data. Means and test statistics for global measures are presented in Table 2.

As expected due to the design specification of grouping participants by achievement scores, analyses across all measures confirmed significant between-groups effects. Significant main effects for readings were also observed for all measures, indicating improvement in the expected direction across all four readings of the passage, with medium effects for first fixation duration (η_p^2=.11) and large effects for all other measures (gaze duration, η_p^2=.60; total fixation time, η_p^2=.78; number of inter-word regressions, η_p^2=.41; number of intra-word regressions, η_p^2=.68; average fixation count, η_p^2=.81). Pairwise comparisons between the first and fourth readings indicated significant aggregate effects across all measures in the expected direction (first fixation duration, p = .014; total fixation time, p = .01; gaze duration, number of intra-word regressions, average fixation count, and number of inter-word regressions, p < .001). Furthermore, pairwise comparisons between the first and second readings revealed immediate, significant effects for measures indicative of both lower-level processing (i.e., gaze duration and number of intra-word regressions, p < .001) and higher-level processing (i.e., number of inter-word regressions, p = .001; total fixation time and average fixation count, p < .001). Significant effects of readings persisted between the second and third readings for all measures indicative of higher-level processing (i.e., number of inter-word regressions, p = .04; total fixation time, p = .001; average fixation count, p < .001), but only one measure thought to represent lower-level processing (i.e., number of intra-word regressions, p < .001).

Average fixation count was the only dependent measure to reveal a significant readings x skill level interaction (η_p^2=.21). Follow-up analyses suggested that average fixation count differed significantly across readings for both low-performing readers (η_p^2=.60) and high-performing readers (η_p^2=.74) (See Figure 1). Analyses of significant aggregate effects indicated large effect sizes for both high-performing readers (d = 1.58) and low-performing readers (d = 1.08). Pairwise comparisons revealed significant decreases in average fixation count between the first and second readings for both groups, although effect size was large for high-performing readers (d = .86) and medium for low-performing readers (d = .60). Pairwise comparisons were also significant for both groups between the second and third readings, with small effects indicated for high-performing readers (d = .42), in contrast to large effects for low-performing readers (*d* = .83). Collectively, findings suggest that RR facilitated processing sooner for high-performing readers than low-performing readers, and that of the two groups, high-performing readers made the greatest gains between the first and fourth readings. Yet, these findings should be interpreted with caution given the presence of a significant outlier that could not be normalized with transformations.

Target Word Analyses

Target word analyses were conducted on first fixation duration, gaze duration, total fixation time, and average fixation count (per target word). A 4 (readings) x 2 (skill level) x 2 (word frequency) mixed ANOVA, with readings and word frequency as within factors, was conducted for each measure. Due to extreme outliers and violations of Levene's test for homogeneity of variance, log (x) transformations were conducted for all variables. Following transformation, data were successfully normalized, with the exception of one data point in gaze duration. Although data from this measure must be approached with caution, interpretations are most likely valid given that results aligned with other target word measures and findings from previous studies (e.g., Foster et al., 2013). After transformation, all variables met assumptions for homogeneity of variance.

For average fixation count, transformation revealed a significant readings x skill level interaction that was not identified in the non-transformed data. Large variability in non-transformed data led to the identification of a significant skill level x word frequency interaction in total fixation time, which was not significant after data were normalized. All other significant findings remained consistent between transformed and non-transformed measures. Detailed test statistics are presented below; see Figures 2-5 for visual comparison of means for target word measures separated by skill level.

First Fixation Duration. Analyses revealed no significant between-groups differences, $F(1,42) = 3.66$, $p = .062$ for target word analyses of first fixation duration. There was also no significant main effect of readings, $F(3,126) = 2.19$, $p = .092$, $\eta_p^2 = .05$, although a significant main effect for word frequency was observed, $F(1,42) = 18.39$, $p < .001$, $\eta_p^2 = .31$. Interaction effects were also not observed for readings x word frequency, $F(3,126) = .68$, $p = .567$, $\eta_p^2 = .02$.

Analyses indicated that the three-way interaction of readings x skill level x word frequency was not significant, $F(3,126) = 1.40$, $p = .248$, $\eta_p^2 = .03$. Two-way interactions were not significant for readings x skill level, $F(3,126) = .73$, $p = .537$, $\eta_p^2 = .02$ or skill level x word frequency, $F(1,42) = .84$, $p = .366$, $\eta_p^2 = .02$. Lack of interaction effects taken together with lack of between-groups effects suggests that with respect to first fixation duration, reading behavior is similar for low- and high-performing readers.

Gaze Duration. As expected, significant between-groups differences were observed for target word analyses on gaze duration, $F(1,42) = 47.90$, $p < .001$, a measure thought to represent lower-level processing. Also significant were main effects for readings, $F(3, 126) = 15.15$, $p < .001$, $\eta_p^2 = .27$, and word frequency, $F(1, 42) = 158.46$, $p < .001$ $\eta_p^2 = .79$, which were qualified by a significant readings x word frequency interaction, $F(3,126) = 7.92$, $p < .001$, $\eta_p^2 = .16$. Results indicated that collectively, gaze duration on target words declined across readings and readers required more time to engage in lower-level processing of low-frequency target words than high-frequency target words. Furthermore, presence of an interaction between these main effects indicates that facilitative effects of RR were greatest on low-frequency target words.

The three-way interaction of readings x skill level x word frequency was not significant for gaze duration, $F(3,126) = .68$, $p = .565$, $\eta_p^2 = .02$. Examination of two-way interactions indicated that effects were not significant for readings x skill level, $F(3, 126) = 1.43$, $p = .237$,

$\eta_p^2 = .03$, but were significant for skill level x word frequency, F (1,42) = 8.98, $p = .005$, $\eta_p^2 = .18$. Pairwise comparisons of the skill level x word frequency interaction indicated significant effects of word frequency for both groups (Low-Performing, $F(1,21) = 113.50$, $p < .001$, $\eta_p^2 = .84$; High-Performing $F(1,21) = 49.46$, $p < .001$, $\eta_p^2 = .70$), although the magnitude of the word frequency effect was greater for low-performing readers.

Total Fixation Time. Consistent with expectations, significant between-groups differences were observed for target word analyses on total fixation time, $F(1,42) = 51.64$, $p < .001$, a measure thought to represent higher-level processing. Also significant were main effects of readings, Wilks' Lambda = .338, $F(3, 40) = 26.17$, $p < .001$, $\eta_p^2 = .66$, and word frequency, $F(1, 42) = 142.15$, $p < .001$, $\eta_p^2 = .77$. The significant main effects were qualified by a significant readings x word frequency interaction, $F(3,126) = 10.25$, $p < .001$, $\eta_p^2 = .20$. As with gaze duration, findings suggested that for participants in both groups, total fixation time on target words decreased across readings, and that low-frequency words required longer processing time. The significant interaction indicates that facilitative effects of RR on this measure of higher-level processing were greatest on low-frequency target words.

Target word analyses of total fixation time did not indicate a significant readings x skill level x word frequency interaction, $F(3,126) = 2.00$, $p = .118$, $\eta_p^2 = .05$. Furthermore, interaction effects were not significant for readings x skill level, Wilks' Lambda = .838, $F(3,40) = 2.59$, $p = .066$, $\eta_p^2 = .16$, nor for skill level x word frequency $F(1,42) = 3.09$, $p = .086$, $\eta_p^2 = .07$. Lack of interactions for target word analyses of total fixation time involving skill level as a factor suggest that RR facilitates higher-level processing similarly for low- and high-performing readers.

Average Fixation Count per Target Word. Significant between-groups differences were observed for target word analyses of average fixation count per target word, $F(1,42) =$

47.38, $p < .001$. As with total fixation time, this measure is thought to represent higher-level processing. Consistent with findings from gaze duration and total fixation time, main effects were significant for readings, $F(3,126) = 30.47$, $p < .001$, $\eta_p^2 = .42$ as well as word frequency, $F(1,42) = 157.67$, $p < .001$, $\eta_p^2 = .79$, and were qualified by a significant readings x word frequency interaction, $F(3,126) = 8.88$, $p < .001$, $\eta_p^2 = .18$. Together, findings indicate that the average number of fixations made on target words decreased across readings, although participants made more fixations on low-frequency target words than on high-frequency target words. The significant interaction indicates that facilitative effects of RR on this measure of higher-level processing were greatest for low-frequency target words.

Although the three-way interaction for readings x skill level x word frequency was not significant, $F(3,126) = 2.28$, $p = .083$, $\eta_p^2 = .05$, all two-way interactions involving skill level as a factor were significant. Further analysis of the readings x skill level interaction, $F(3,126) = 3.10$, $p = .029$, $\eta_p^2 = .07$, indicated that effects of readings were significant for high-performing readers between the first and second readings ($p < .001$), whereas significant effects for low-performing readers were not observed until between the second and third readings ($p = .025$). The significant skill level x word frequency interaction, $F(1,42) = 6.26$, $p = .016$, $\eta_p^2 = .13$, was further explored to reveal that the magnitude of the word frequency effect on average fixation count per target word was greater for low-performing readers, $F(1,21) = 103.76$, $p < .001$, $\eta_p^2 = .83$, than for high-performing readers, $F(1,21) = 55.71$, $p < .001$, $\eta_p^2 = .73$.

Taken together, findings reveal that RR permits low-performing readers to make greater reductions in average fixation count across readings, particularly on low-frequency target words. Furthermore, high-performing readers make greater gains sooner (i.e., between the first and second readings) than the low-performing readers (i.e., between the second and third readings).

Unlike target word analyses for total fixation time, another measure of higher-level processing, all two-way interactions for target word analyses of average fixation count involving skill level as a factor were significant.

Table 2

Summary of Global Eye Movement Parameters across Readings by Skill Level

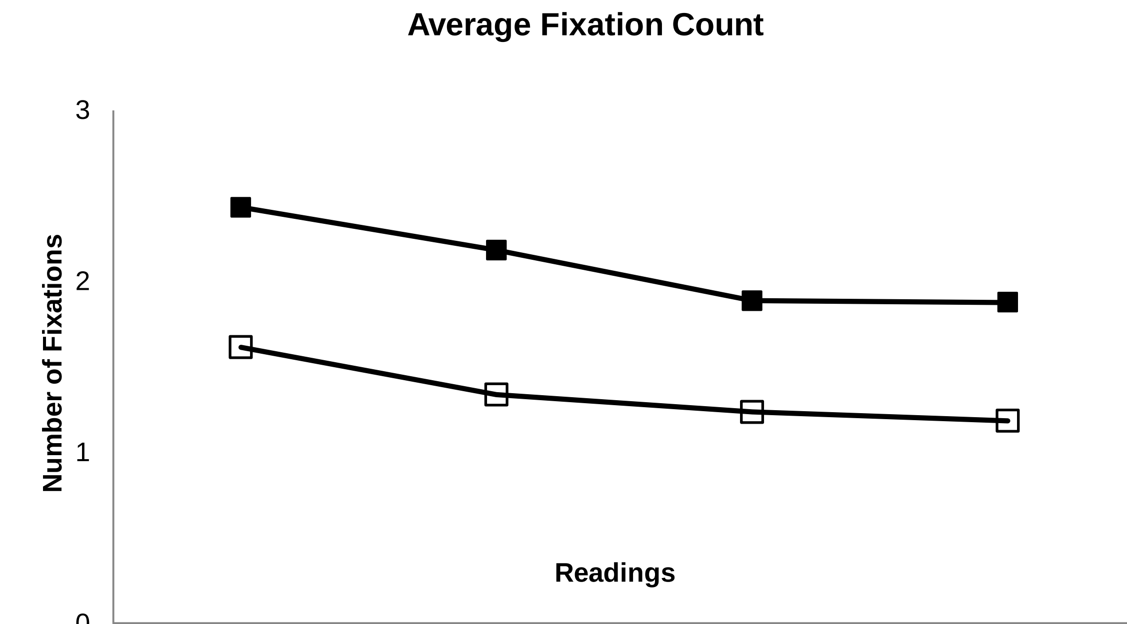
Measure	Reading 1 M (SD)	Reading 2 M (SD)	Reading 3 M (SD)	Reading 4 M (SD)
First fixation duration (ms) (*N*=44)	280(31)	274(32)	274(31)	270(31)[+]
Between Groups: $F(1, 42) = 8.916$, $p = .005$				
Main effect (Readings): $F(3, 126) = 4.91$, $p = .003$				
Interaction (Readings x Skill Level): $F(3, 126) = 1.014$, $p = .389$				
Low-Performing Readers (*n*=22)	291(26)	287(29)	288(26)	281(26)
High-Performing Readers (*n*=22)	269(32)	262(30)	259(30)	259(32)
Gaze duration (ms) (*N*=44)	430(106)	397(97)*	391(90)	377(86)[+]
Between Groups: $F(1, 42) = 56.563$, $p < .001$				
Main effect (Readings): $F(3, 40) = 19.820$, $p < .001$[a]				
Interaction (Readings x Skill Level): $F(3, 40) = 1.515$ $p = .225$[b]				
Low-Performing Readers (*n*=22)	510(76)	464(77)	443(77)	434(78)
High-Performing Readers (*n*=22)	350(61)	329(61)	320(52)	320(47)
Total fixation time (ms) (*N*=44)	634(194)	543(172)*	504(156)*	490(161)[+]
Between Groups: $F(1, 42) = 70.531$, $p < .001$				
Main effect (Readings): $F(3, 40) = 48.337$, $p < .001$ [c]				
Interaction (Readings x Skill Level): $F(3, 40) = .998$, $p = .404$[d]				
Low-Performing Readers (*n*=22)	782(147)	674(134)	614(143)	601(155)
High-Performing Readers (*n*=22)	486(99)	412(83)	393(64)	380(59)
Number of inter-word regressions (#) (*N*=44)	.317(.112)	.273(.105)*	.243(.095)*	.229(.099)[+]
Between Groups: $F(1, 42) = 19.830$, $p < .001$				
Main effect (Readings): $F(3, 126) = 28.898$, $p < .001$				
Interaction (Readings x Skill Level): $F(3, 126) = 1.108$, $p = .387$				
Low-Performing Readers (*n*=22)	.362(.092)	.335(.089)	.292(.081)	.284(.084)
High-Performing Readers (*n*=22)	.271(.113)	.211(.080)	.195(.084)	.175(.082)

Table 2 Continued

Number of intra-word regressions (#) (*N*=44)	.251(.154)	.187(.122)*	.160(.010)*	.150(.108)$^{+}$
Between Groups: $F(1,42) = 51.161$, $p < .001$				
Main effect (Readings): $F(3, 40) = 28.205$, $p < .001$[e]				
Interaction (Readings x Skill Level): $F(3, 40) = 2.558$, $p = .069$[f]				
Low-Performing Readers (*n*=22)	.357(.140)	.269(.111)	.233(.091)	.219(.112)
High-Performing Readers (*n*=22)	.146(.074)	.106(.065)	.088(.033)	.080(.038)
Average fixation count (#) (*N*=44)	2.02(.568)	1.76(.544)*	1.56(.428)*	1.53(.487)$^{+}$
Between Groups: $F(1,42) = 57.924$, $p < .001$				
Main effect (Readings): $F(3, 40) = 58.395$, $p < .001$[g]				
Interaction (Readings x Skill Level): $F(3, 40) = 3.474$, $p = .025$[h]				
Low-Performing Readers (*n*=22)	2.43(.447)	2.18(.381)*	1.89(.315)*	1.89(.437)$^{+}$
Follow up: $F(3,63) = 31.119$, $p < .001$				
High-Performing Readers (*n*=22)	1.61(.336)	1.34(.295)*	1.23(.232)*	1.18(.207)$^{+}$
Follow up: $F(3, 63) = 58.682$, $p < .001$				

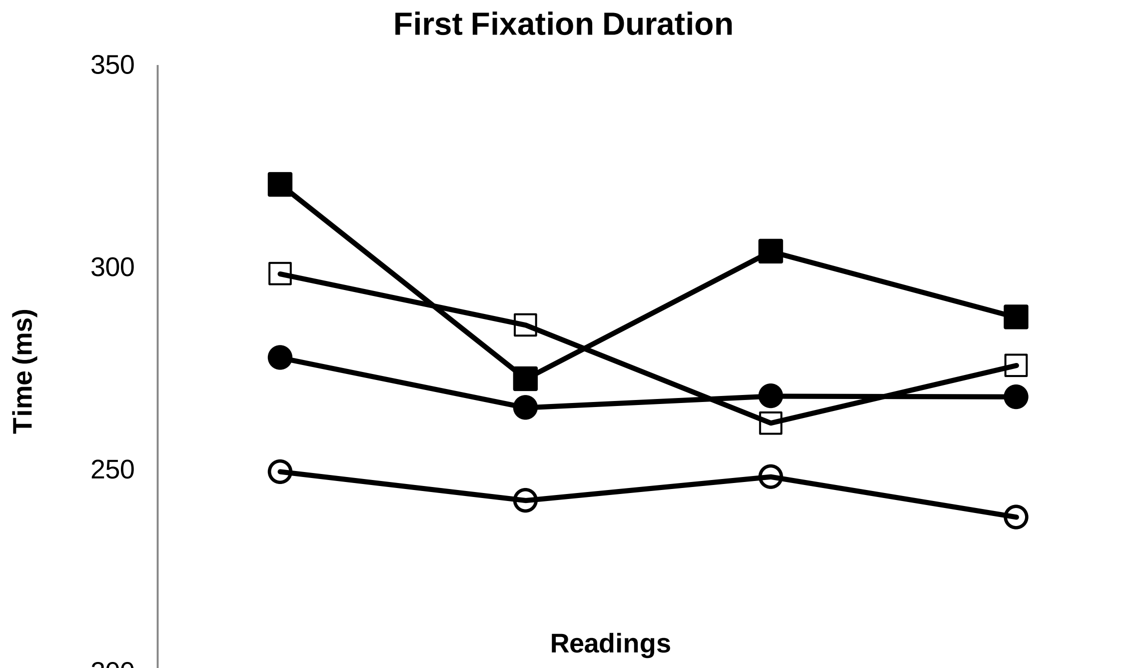
Gaze Duration: [a]Wilks' Lambda = .40. [b]Wilks' Lambda = .90
Total Fixation Time: . [c]Wilks' Lambda = .22 [d]Wilks' Lambda = .93
Number of Intra-word Regressions: . [e]Wilks' Lambda = .32 [f]Wilks' Lambda = .84
Average Fixation Count: [g]Wilks' Lambda = .19 [h]Wilks' Lambda = .79

*Significant pairwise differences between the denoted reading and its previous reading, p <.05
$^{+}$Significant pairwise differences between the first and fourth reading, p < .05



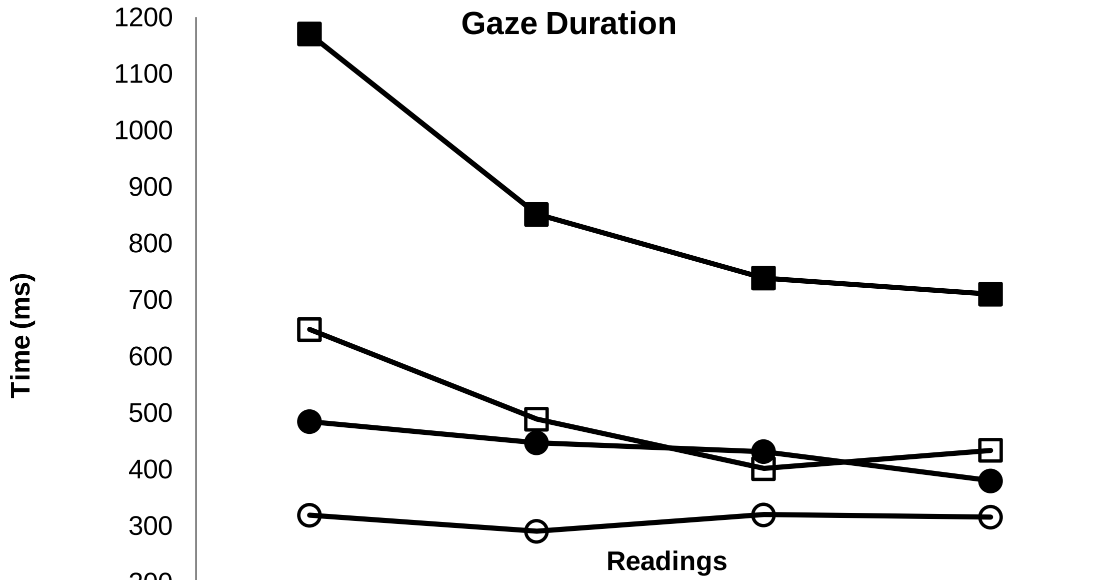
	1	2	3	4
LP	2.431	2.181	1.885	1.877
HP	1.614	1.337	1.235	1.184

Figure 1. Average fixation count across readings, separated by skill level (LP= Low-Performing; HP= High-Performing).



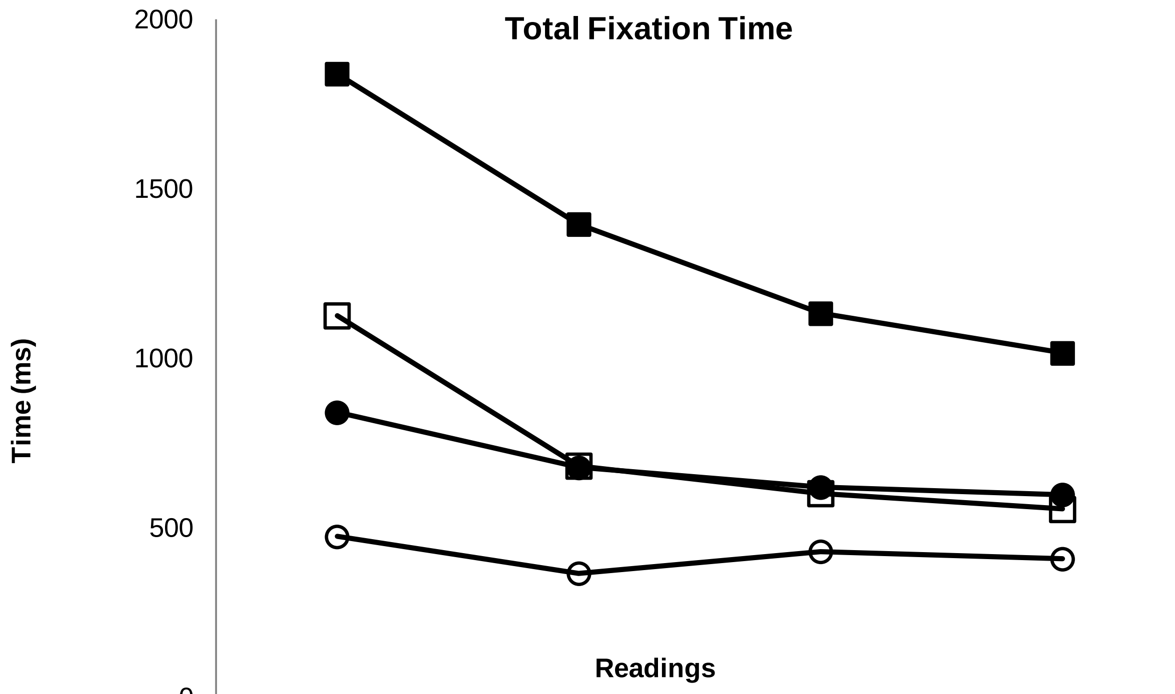
	1	2	3	4
LP-High Freq	277.657	265.314	268.177	267.932
LP-Low Freq	320.432	272.4	303.929	287.675
HP-High Freq	249.42	242.35	248.201	238.213
HP-Low Freq	298.389	285.692	261.427	275.698

Figure 2. Average first fixation duration on low- and high-frequency target words across readings, separated by skill level. (LP= Low-Performing; HP= High-Performing; Low-Freq= Low-Frequency; High-Freq= High-Frequency)



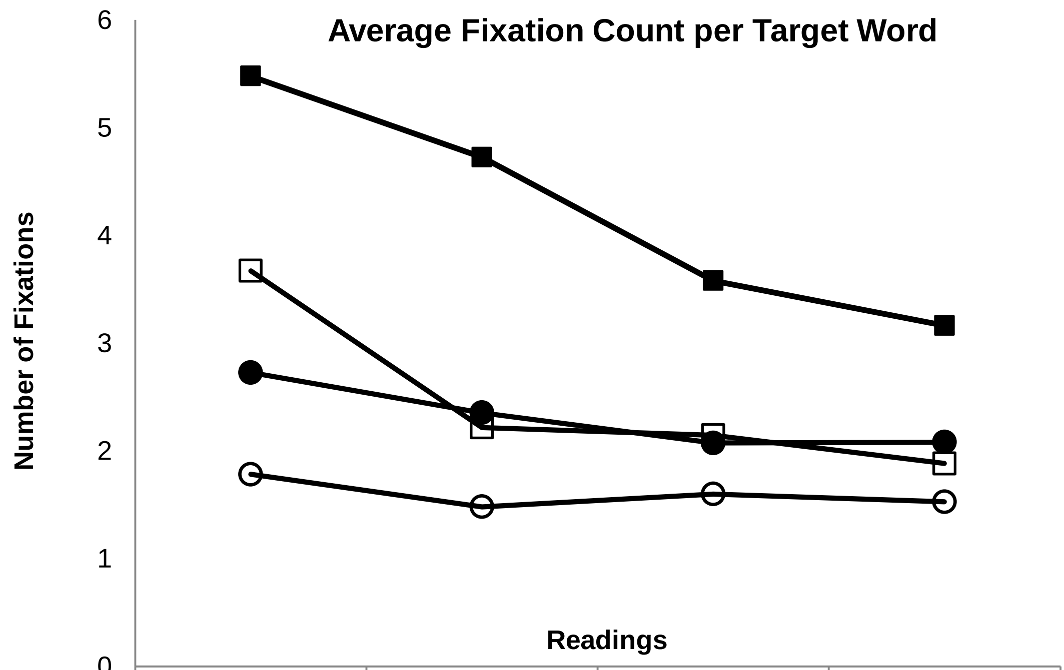
	1	2	3	4
LP- High Freq	484.284	446.718	431.191	379.2
LP- Low Freq	1170.439	850.9	738.361	709.784
HP- High Freq	318.55	290.095	319.071	315.125
HP Low Freq	647.186	488.469	400.923	433.511

Figure 3. Average gaze duration on low- and high-frequency target words across readings, separated by skill level. (LP= Low-Performing; HP= High-Performing; Low-Freq= Low-Frequency; High-Freq= High-Frequency)



	1	2	3	4
LP-High Freq	839.125	676.283	618.468	596.548
LP-Low Freq	1838.316	1394.645	1131.42	1014.384
HP-High Freq	472.761	364.264	428.851	406.979
HP-Low Freq	1124.986	681.642	600.457	553.652

Figure 4. Average total fixation time on low- and high-frequency target words across readings, separated by skill level. (LP= Low-Performing; HP= High-Performing; Low-Freq= Low-Frequency; High-Freq= High-Frequency)



	1	2	3	4
LP-High Freq	2.727	2.355	2.073	2.082
LP-Low Freq	5.482	4.727	3.582	3.164
HP-High Freq	1.782	1.482	1.6	1.527
HP-Low Freq	3.673	2.218	2.145	1.882

Figure 5. Average fixation count on low- and high-frequency target words across readings, separated by skill level. (LP= Low-Performing; HP= High-Performing; Low-Freq= Low-Frequency; High-Freq= High-Frequency)

CHAPTER 4

DISCUSSION

Findings from numerous applied research studies suggest that RR is an effective reading intervention (Chard et al., 2002; Therrein, 2004) for readers through fourth grade (NICHD, 2000, pp. 3-3-3-4). Despite evidence of effectiveness, methodological limitations leave unanswered questions regarding how and why readers of different skill levels benefit from the same intervention. Fortunately, eye tracking technology provides a means for researchers to observe how intervention impacts the behaviors that result in differential reading outcomes. Several eye tracking studies conducted with early elementary readers (Ardoin et al., in press; Foster et al., 2013) support findings from applied research by providing more specific information about children's eye movements during rereading, but conclusions cannot be extended to low-performing readers due to inadequate representation within the samples. In order to better understand how second grade readers of different skill levels benefit from rereading, the current study examined underlying changes in reading behavior of low- and high-performing students during RR via eye tracking technology.

In general, findings were consistent with expectations for between-groups differences in overall benefits of RR and general characteristics of eye movements. First, results suggested that students benefited from RR regardless of their skill level, which was evidenced by the increased efficiency in their eye movements across the four reading trials. Particularly, there were significant decreases in fixation frequency, fixation duration, and number of regressions made within words or between portions of text. It was also hypothesized that high-performing readers would not require all four readings to attain full benefits of RR, whereas low-performing readers would improve significantly throughout all readings. Surprisingly, findings indicated that neither group improved significantly between the third and fourth readings, suggesting that readers in both groups required only three readings to attain full facilitative effects of RR. Foster et al. (2013) also observed that students made optimal gains within three readings, thus findings from the current study support and extend these implications to low- and high-performing readers. Finally, findings were consistent with the expectation that high-performing readers would exhibit more efficient eye movements than low-performing readers. Specifically, high-performing readers made fewer and shorter fixations than low-performing readers. They also made fewer inter- and intra-word regressions, indicating that high-performing readers did not need to reread previous text or parts of individual words as often as low-performing readers.

Global Analyses

Global analyses of eye movements during RR revealed improvement for low- and high-performing readers on all measures between the first and fourth readings, with evidence of a greater magnitude of improvement for the low-performing readers. However, in contrast to expectations, the low-performing readers never matched the levels of reading efficiency observed during the high-performing readers' first reading. Levy et al. (1993) reported similar findings with low- and high-performing third grade readers on reading outcome measures such as overall reading time, performance on comprehension questions, and detection of word and nonword errors during RR. Thus, results from the current study measuring differences in underlying reading behavior of low- and high-performing readers are consistent with findings from studies measuring outcomes of reading.

Despite differences in processing efficiency, the pattern of improvement across readings was similar between groups. This finding aligns with LaBerge and Samuels' (1974) theory of automaticity. RR initially helped readers attain automaticity in lower-level processing, which allowed them to continue making significant gains in higher-level processing throughout later readings. Specifically, low- and high-performing readers made the greatest gains on measures of lower-level processing (e.g., gaze duration) between the first and second readings, and continued to make significant gains on measures considered to reflect higher-level processing through the second and third readings (e.g., total fixation time, number of inter-word regressions, average fixation count).

Interestingly, the only measure associated with lower-level processing that students continued to significantly improve on between the second and third readings was number of intra-word regressions, which is a measure of the number of regressions a student makes within a word *throughout the entire reading*. Given that other measures of lower-level processing (i.e., first fixation duration and gaze duration) are recorded only on the reader's *first attempt* at reading a word, the continued improvement for intra-word regressions between the second and third reading might suggest that when second grade readers make regressions to previous text, they still require time to decode the words they are rereading. Yet, other factors, such as differences in measurement dimensionality (i.e., count vs. duration) or functions of these reading behaviors may also explain why findings for number of intra-word regressions were different. Further research on intra-word regressions, a measure rarely examined within the adult literature, is necessary to support this interpretation.

The only significant interaction between readings and skill level observed was for average fixation count, suggesting that this measure may be sensitive enough to detect subtle differences in how RR differentially benefits low- and high-performing readers. Specifically, the greatest decrease in average fixation count for high-performing readers occurred between the first and second readings, whereas the greatest decrease in average fixation count for low-performing readers occurred between the second and third readings. Given that decrease in fixation count during RR is thought to represent higher-level processing (Hyona & Niemi, 1990), results imply that high-performing readers, who were rereading the passage at a mastery level, comprehended information to a higher level sooner than low-performing readers, and that low-performing readers, who were rereading the passage at an instructional level, required another reading to facilitate their word recognition and decoding before they could make significant gains in comprehension.

Target Word Analyses

Target word analyses supported conclusions from previous research suggesting that RR primarily improves fluency in second grade readers by reducing the amount of additional processing required to read low-frequency target words (Foster et al., 2013). Results extend these conclusions by illuminating that the effects were more pronounced for low-performing readers. In fact, it was only on *low-frequency target words* that the low-performing readers' performance on the fourth reading surpassed that of the high-performing readers' first reading (See Figures 2-5). The only exception to this finding was for low-frequency target word analyses of gaze duration (See Figure 3). A potential explanation for the inconsistency is that high-performing readers may have comparatively greater skill in swiftly recognizing and decoding low-frequency target words than low-performing readers. Furthermore, despite lack of group differences in analyses, high-performing readers evidenced particular strength in initial processing of high-frequency target words. Specifically, the high-performing readers' first fixation durations on high-frequency target words (first reading: 249ms, final reading: 238ms) were strikingly similar to those made by adult readers in Raney and Rayner (1995) (first reading:

250ms, final reading: 236 ms). Results imply that first fixation duration did not improve after RR due to floor effects (i.e., readers require a certain amount of time to recognize words). That is, readers had attained automaticity in word recognition for these high-frequency target words and could no longer improve beyond this time, which reflects the amount of time required for initial word recognition.

Findings from target word analyses pertaining to the impact of word frequency replicate previous research and extend implications to low- and high-performing readers. As observed by Foster et al. (2013) and Raney and Rayner (1995), word frequency effects were significant across target word analyses, revealing that low-frequency target words required more processing time than high-frequency target words. Interactions were, however, observed between skill level and word frequency for the measures of gaze duration and average fixation count. Consistent with expectations, results suggest that the magnitude of the word frequency effect was greater for low-performing readers. This finding implies that low-frequency target words had a greater negative impact on the reading of low-performing than on high-performing readers. Conversely, the low-frequency target words were potentially not as challenging for high-performing readers. As an example, averaged across readings, the difference between gaze duration on low- and high-frequency target words was 432ms for low-performing readers, but only 182ms for high-performing readers. Similar findings were also observed for average fixation count. Although the skill level and word frequency interaction was not significant for total fixation time, lack of significance may be due to large amounts of variability in the data for this measure.

Consistent with global findings, RR impacted the low- and high-performing readers' fixation frequency and duration on low-and high-frequency target words in a similar manner. For example, in total fixation time on low-frequency words, low- and high-performing readers improved by exactly the same amount (443ms) between their first and second readings. Also consistent with global analyses, the only measure to reveal a significant interaction between skill level and readings was average fixation count, suggesting that low-performing readers required an additional reading to make gains in comprehending target words. This consistency suggests that average fixation count may be the only measure sensitive enough to detect subtle differences among early elementary readers. Further research on this measure is clearly warranted, given the potential for advancing knowledge about underlying behaviors involved in reading comprehension.

Limitations and Future Directions

Conclusions from the current study should be interpreted with consideration of multiple limitations. First, given the variability in data and a relatively small sample size (N = 44, n = 22 per group), replication with a larger sample is necessary to make interpretations about the subtle group differences on average fixation count measures more conclusive. A related limitation is that 17 of the eligible low-performing readers and 14 of the eligible high-performing readers were excluded from the sample due to observed mindless reading. Mindless reading is characterized by a sudden decrease in fixations or regressions and a sudden increase in erratic eye movements (Nguyen, Binder, Nemier, & Ardoin, submitted). Given that readers can only extract information from the text when making fixations (Rayner, 1998), mindless reading behavior is thought to indicate that a reader has "zoned out" and is no longer reading (Reichle, Reineberg, & Schooler, 2010; Schad, Nuthmann, & Engbert, 2012). The excluded participants' reading behavior was not appropriate for analyses in the current study because their data did not represent a complete session of RR. If the students were not reading the words presented within the text across trials, RR did not occur. Identification and awareness of mindless reading behavior creates multiple avenues for future research on mindless reading in children, particularly because this behavior was observed in both low- and high-performing readers. A final limitation pertaining to the sample is that the participants of Foster et al., (2013) made up a small portion (20%) of the sample from which the participants in this study were selected. However, given that 80% of participants in the current sample were unique to this study, it is unlikely that results are merely a function of shared participants.

An important perennial limitation to eye tracking research on RR with children is that findings reflect changes in students' reading behavior across silent rereading of a passage as opposed to intervention. Therrien (2004) suggested that RR is most effective when students read aloud to an adult who provides error correction and performance feedback regarding improvements in reading time. Unfortunately, implementing such an intervention while monitoring children's eye movements presents numerous obstacles (e.g., significant head and mouth movement) that are unfavorable to quality data collection using current eye tracking technology. Once technology allows for valid oral reading data collection, future research should examine skill level differences in eye movements during RR conducted with typical classroom implementation procedures.

Summary and Implications

Overall, findings from the current study imply that repeated practice will help second grade readers regardless of skill level improve their word- and passage-level processing of text. Furthermore, differences between low- and high-performing readers' eye movements reveal differences in their text processing efficiency. Together, findings have important implications for improving eye tracking research and measurement, tailoring classroom instruction by skill level, and enhancing the effectiveness of RR.

First, researchers interesting in studying eye movements of children during reading are encouraged to consider how skill level may impact their findings. Although many researchers are aware that inferences from adult research should not be extended to children due to differences in their eye movements (Blythe & Joseph, 2011), findings from the current study extend this implication to early elementary readers at different skill levels. Given the observed within grade level differences in eye movements, researchers are cautioned against generalizing findings from one group of second grade readers to all second grade readers. Instead, findings should be considered in light of the participants' skill levels, particularly when participants span across grade levels.

Findings from target word analyses of first fixation duration have important measurement implications for future eye tracking research. Despite expectations for differences in group performance and evidence for between groups differences on all measures, between group differences were not present for target word analyses of first fixation duration. This finding is consistent with Valle et al. (2013), who did not observe significant differences between average- and high-performing readers on first fixation duration. Similarly, Foster et al. (2013) did not report significant improvement in first fixation duration throughout RR in global or target word analyses. Given the consistent differences between first fixation duration and all other measures, researchers may consider interpreting findings from children's first fixation duration with caution until further research can sufficiently explain why this measure differs from all other measures.

In addition to research implications, results have applied implications for modifying classroom instruction to benefit low- and high-performing readers. Importantly, most measures indicated that reading a passage four times was not enough to help the low-performing readers match the high-performing readers' first attempt, in terms of reading efficiency. However, the magnitude of improvement was greater for the low-performing readers, suggesting that in order for them to improve to the level of their high-performing peers, they may benefit from the addition of evidence-based enhancements to RR. Findings also imply that in relation to high-performing readers, the low-performing readers had poorer word recognition and decoding skills.

Finally, findings have implications for improving the effectiveness and efficiency of RR. First, because neither the low- nor the high-performing readers improved significantly between the third and fourth readings, all students may experience diminishing returns from reading a text four or more times in one session. This has particular significance for students who receive RR as a pull-out service. Programming for three readings of a text would seem to maximize the benefits of RR and minimize the amount of time students spend outside of general classroom instruction. A second implication for improving RR pertains to the importance of considering intervention goals when selecting intervention passages. Past research recommends conducting reading intervention with challenging text that matches a student's instructional level (Daly, Marten, Kilmer, & Massie, 1996; Therrien & Kubina, 2006). However, these studies measured reading fluency only and not the impact of RR on students' comprehension. Recall that in the current study, high-performing readers, who read at a mastery level, improved faster than low-performing readers on average fixation count, a measure thought to reflect comprehension. The low-performing readers, who read at an instructional level, first improved on word recognition and decoding and exhibited significantly slower processing times on comprehension measures. Together, these findings suggest that when comprehension is the goal, students will make improvements sooner if they read a mastery level passage. On the other hand, if word recognition and faster decoding are the goals of intervention, then students will make greater gains on instructional level passages.

Ultimately, questions remain regarding subtle differences in how RR impacts low- and high-performing readers' eye movements. Yet, the general implication that RR is effective for early elementary readers of varying skill level is important for providing preliminary understanding about the development of reading. This study provides another demonstration of the successful marriage between applied and eye tracking research, with hope of continued collaboration between these fields for the greater benefit of all students.

REFERENCES

Ardoin, S. P., Binder, K. S., Zawoyski, A. M., Foster, T. E., & Blevins, L. A. (in press). Using eye-tracking procedures to evaluate generalization effects: Practicing target words during repeated reading within versus across texts. *School Psychology Review.*

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology, 58A*, 1065–1086.

Aud, S., Hussar, W., Johnson, F., Kena, G., Roth, E., Manning, E., Wang, X., & Zhang, J. (2012). *The Condition of Education 2012* (NCES 2012-045). U.S. Department of Education, National Center for Education Statistics. (pp. 62-63) Washington, DC.: U.S. Government Printing Office.

Buswell, G. T. (1922). *Fundamental reading habits: A study of their development.* Chicago: University of Chicago Press.

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American Heritage word frequency book.* Boston, MA: Houghton Mifflin Company.

Chard, D. J., Vaughn, S., & Tyler, B. J. (2002). A synthesis of research on effective interventions for building fluency with elementary students with learning disabilities. *Journal of Learning Disabilities, 35,* 386-406.

Christ, T. J., Ardoin, S. P., Monaghen, B., Van Norman, E., & White, M. J. (2013) *CBMReading: Technical Manual*. Minneapolis, MN: University of Minnesota, Department of Educational Psychology.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Daly, E. J., Martens, B. K., Kilmer, A., & Massie, D. R. (1996). The effects of instructional match and content overlap on generalized reading performance. *Journal of Applied Behavior Analysis*, 507-518.

Faulkner, H. J., & Levy, B. A. (1994). How text difficulty and reader skill interact to produce differential reliance on word and content overlap in reading transfer. *Journal of Experimental Child Psychology, 58*, 1-24.

Faulkner, H. J. & Levy, B. A. (1999). Fluent and nonfluent forms of transfer in reading: Words and their message. *Psychonomic Bulletin & Review, 6*, 111-116.

Foster, T. E., Ardoin, S. P., & Binder, K. S. (2013). Underlying changes in repeated reading: An eye movement study. *School Psychology Review, 42,* 140-156.

Huey, E. B. (1908/1968). *The psychology and pedagogy of reading.* Cambridge, MA: MIT Press. (Original work published 1908).

Hyönä, J., & Niemi, P. (1990). Eye movements during repeated reading of a text. *Acta Psychologica, 73*, 259-280.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293-323.

Levy, B. A. & Burns, K. I. (1990). Reprocessing text: Contributions from conceptually driven processes. *Canadian Journal of Experimental Psychology, 44*, 465-482.

Levy, B. A., Di Persio, R., & Hollingshead, A. (1992). Fluent rereading: Repetition, automaticity, and discrepancy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 957-971.

Levy, B. A., Nicholls, A., & Kohen, D. (1993). Repeated Readings: Process benefits for good and poor readers. *Journal of Experimental Child Psychology, 56,* 303-327.

McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W, Bryant, N. R., & Wolff, P.M. (1991). Children's eye movements during reading. In J. F. Stein (Ed.), *Vision and visual dyslexia* (pp. 251-262). London: Macmillan Press.

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: Reports of the subgroups* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.

Nguyen, V. T., Binder, K. S., Nemier, C., & Ardoin, S. P., (submitted). Gotcha! Catching kids during mindless reading. *Psychological Science.*

Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology, 49*, 151-173.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41*, 211-236.

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

Rayner, K., Ardoin, S. P., & Binder, K. S. (2013). Children’s eye movements in reading: A commentary. *School Psychology Review, 42*, 223-233.

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading, 10*, 241-255.

Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science, 21*, 1300-1310.

Samuels, S. J. (1979). The method of repeated readings. *The Reading Teacher, 32*, 403-408.

Samuels, S. J. (2006). Looking backward: Reflections on a career in reading. *Journal of Literacy Research, 38,* 327-344.

Schad, D. J., Nuthmann, A., & Engbert, R. (2012). Your mind wanders weakly, your mind wanders deeply: Objective measures reveal mindless reading at different levels. *Cognition, 125*, 179-194.

Shebilske, W. L., & Fisher, D. F. (1980, December). *Eye movements reveal components of flexible reading strategies*. Paper presented at the Annual Meeting of the National Reading Conference, San Diego, CA.

Sindelar, P. T., Monda, L. E., & O'Shea, L. J. (1990). Effects of repeated readings on instructional- and mastery-level readers. *Journal of Educational Research, 83*, 220-226.

Spache, G. (1953). A new readability formula for primary-grade reading materials. *Elementary School Journal, 55,* 410–413.

Taylor, S. E. (1965). Eye movements while reading: Facts and fallacies. *American Educational Research Journal, 2,* 187-202.

Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading. *Remedial and Special Education, 25,* 252-261.

Therrien, W. J., & Kubina, R. M. (2006). Developing reading fluency with repeated reading. *Intervention in School and Clinic, 41*, 156-160.

Valle, A., Binder, K. S., Walsh, C. B., Nemier, C., & Bangs, K. E. (2013). Eye movements, prosody, and word frequency among average- and high-skilled second grade readers. *School Psychology Review, 42, 171-190.*

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing.