DIFFICULTY OF CURRICULUM-BASED MEASURES IN READING (CBM-R) PASSAGES ON ORAL AND SILENT READING ASSESSMENTS

by

LESLIE ADLINE BLEVINS

(Under the Direction of Scott P. Ardoin)

ABSTRACT

Curriculum based measures in reading (CBM-R) are used in schools to progress monitor elementary students' growth in oral reading fluency. CBM-R uses alternate forms of passages to produce trend lines, which are subject to error associated with nonequivalence in passages. This study uses eye-tracking technology to analyze students' oral and silent readings of the same passage to evaluate textual characteristics that influence students' reading so that equivalence across passages might be improved. Participants were second grade average to superior-skill readers (n = 56). The passage included high, medium, and low frequency content words and function words as indicators of textual difficulty. Dependent variables included eye movements, oral word reading times, pre- and post-word pauses, and oral word errors. Results support prior research with word frequency and suggest using frequency to group words.

INDEX WORDS: elementary students, eye-tracking, eye movements, silent reading, oral reading, Curriculum-based Measures in Reading (CBM-R), spectrographic, oral reading fluency

DIFFICULTY OF CURRICULUM-BASED MEASURES IN READING (CBM-R) PASSAGES

ON ORAL AND SILENT READING ASSESSMENTS

by

LESLIE ADLINE BLEVINS

B.S., Northeastern University, 2006

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

2015

DIFFICULTY OF CURRICULUM-BASED MEASURES IN READING (CBM-R) PASSAGES

ON ORAL AND SILENT READING ASSESSMENTS

by

LESLIE ADLINE BLEVINS

B.S., Northeastern University, 2006

Major Professor: Scott P. Ardoin

Committee: Stacey Neuharth-Pritchett
Amy L. Reschly

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
May 2015

ACKNOWLEDGEMENTS

This thesis would not have been possible without the encouragement and thoughtful guidance of my advisor, Dr. Scott Ardoin. I would like to extend my gratitude to the other members of my committee, Dr. Stacey Neuharth-Pritchett for her detailed guidance during the writing process and Dr. Amy Reschly for her valuable advice and support. A special thanks to the schools, teachers, and students who allowed us to work with them in the collection of this data is extended.

TABLE OF CONTENTS

Page

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Curriculum-based measures in reading (CBM-R) are frequently used in elementary schools to monitor students' skills throughout the academic year (Reschly, Busch, Betts, Deno, & Long, 2009). This practice involves the individual student reading a grade-leveled passage while the teacher identifies errors and notes the number of words read correctly in one minute (WRCM). Teachers' frequent administration of CBM-R probes and plotting of the data (WRCM) in a time series fashion provides a vital process for teachers to graphically examine trends in students' reading fluency and to monitor individual student progress (Deno, 1985).

Although CBM-R is commonly used in schools, there are some concerns with such assessment measures. For example, there are concerns associated with measurement error associated with estimate of student growth using CBM-R progress monitoring data due to the variability in difficulty of passages with CBM-R passage sets. Research indicates error estimates associated with progress monitoring data are much greater than the established guideline for weekly gains in reading. Current passage-leveling methods, namely readability formulas, often do not result in sufficiently-equivalent passages across testing occasions, resulting in unacceptably high error in estimates of student growth. Such formulas generally rely on counts of passage variables (i.e., number of syllables in words, number of words in sentences) for estimating passage difficulty. Research suggests measuring these characteristics alone might not be sufficient to effectively equate passage difficulty (Ardoin & Christ, 2009; Ardoin, Williams, Christ, Klubnik, Wellborn, 2010). Specifically, empirical work indicates such formulas fail to explain variance in students' WRCM (Ardoin et al., 2010). In other words, the formulas fail to consider variables that impact students' reading behavior, such as speed or accuracy. Consequently, nonequivalent passages are being used to monitor student growth resulting in substantial error in progress monitoring data and thus unreliable estimates of student growth.

Recent methodological advances to equate passages have employed field testing passages to select sets of passages of similar difficulty. This approach requires the administration of numerous CBM-R passages to a large sample of students to identify which passages result in similar levels of WRCM performance across students. Although selection of CBM-R passages using such methodology is promising, there remains considerable variability in the difficulty of passages within passage sets (Ardoin & Christ, 2009). One limitation associated with the data collected through this process is that, by assessing students WRCM, the resultant data does not provide information regarding the textual aspects of passages that might hinder or enhance reading rate (Rayner, Ardoin, & Binder, 2013). Textual characteristics, such as word length, number of syllables in a word, word location in a passage, word use (content versus function), familiarity of words based on the frequency with which words appear in texts, and delimiters (e.g., commas, periods), have been associated with students' reading behavior (e.g., Miller & Schwanenflugel, 2006; Valle, Binder, Walsh, Nemier, & Bangs, 2013). Additional concerns in progress monitoring data in reading have centered on the lack of use of textual characteristics to create passages of sufficient equivalence to be used during progress monitoring.

Improving the Accuracy of Reading Growth Assessment with Eye Tracking Technology

A special issue of *School Psychology Review* in 2013 focused on the rise in research using eye movements to study children's reading skill. Joseph, Nation, and Liversedge (2013) describe eye movements as an online measure of reading in which researchers see reading behavior, or eye movements, in real time rather than using offline measures such as WRCM, which provide only an after-the-fact snapshot of students' reading. Eye-tracking studies might offer precision in measurement not feasible with standard oral reading fluency measures (Rayner, Chace, Slattery, & Ashby, 2006). Eye movements during reading are characterized by several distinct behaviors. In eye-tracking, reading behavior is captured with cameras detecting fixations, or brief pauses, and saccades, or quick jumps forward or backward in the text. Fixations on a word suggest the reader is actively processing information, with longer fixations suggesting longer processing time (Rayner et al., 2006). Forward saccades are too brief to provide information for linguistic processing (Rayner & Slattery, 2009). Backward saccades, or regressions, might indicate a reader is correcting for overshoots during saccades or is returning to prior text to allow for additional linguistic processing (Rayner et al., 2006).

Rayner and Liversedge (2011) provide context for interpreting eye movements in relation to underlying cognitive processes. Specifically, they postulate readers must first recognize a word (initial processing) before processing that word into sentence structure (syntactic processing) and giving it textual meaning (semantic processing). Gaze Duration is the sum of fixations on the first encounter of a word and generally reflects initial lexical processing. Total Time sums fixations from all encounters with a word and generally reflects later, or higher level, processing. Total Fixations might reflect the degree to which a reader struggled with initial or later processing. Regressions within a passage might reflect difficulty with later processing. Additionally, readers gain information about words contextually as they read making their predictions of each subsequent word easier (Rayner & Liversedge, 2011). Consequently, it is common for high-skill students and adult readers to skip words without losing meaning of the text (Rayner, 2009; Valle et al., 2013). The interconnectivity of reading behaviors within a passage makes accurately attributing cognitive reading processes to specific eye movements complex across situations.

Rigorous research studies employing eye-tracking methodology have provided insight into eye movements associated with reading behavior that allow inferences about cognitive processes (e.g., decoding) that are a function of textual characteristics (e.g., word frequency; Joseph, Nation, & Liversedge, 2013; Rayner & Slattery, 2009; Valle et al., 2013). The ability to measure students' reading behavior on individual words as opposed to measuring fluency across all words within a passage (i.e., WRCM) allows for the evaluation of the effects of specific word characteristics and other textual variables on students' reading.

Studying Textual Characteristics with Eye Tracking Technology

Textual characteristics within passages, such as the frequency with which a word is typically encountered, have been associated with students' eye movements (e.g., Raney & Rayner, 1995; Rayner & Well, 1996; Valle et al., 2013). High frequency (HF) words, which are words that frequently appear in text (e.g., at least 40 times out of a million words), generally take less time for students to read when compared to low frequency (LF) words (e.g., words appearing less than 10 times per million words; Ardoin, Binder, Zawoyski, Foster & Blevins, 2013; Just & Carpenter, 1980). HF and LF words have been found to significantly impact the duration of students' first pass processing time (Gaze Duration) with LF words requiring more time (Blythe, Liversedge, Joseph, White, & Rayner, 2009; Rayner & Liversedge, 2011). Although information regarding the impact of such textual characteristics may be difficult to obtain through typical CBM-R administration procedures, it is possible that the impact of textual characteristics on passage difficulty could be ascertained through measuring students' eye-movements while reading (Rayner, Ardoin, & Binder, 2013).

Although eye-tracking is used to measure silent reading skills, it is limited in assessing actual reading errors during silent reading because there is no confirmation of a student’s word reading accuracy during such processes. During silent reading, eye-tracking is not yet sensitive enough to differentiate between actual errors and eye movements that suggest other reading behaviors, such as decoding (Rogers & Ardoin, 2014).

Rogers and Ardoin (2014) utilized eye-tracking technology to observe second grade students’ eye movements while they read the same passage both orally and silently. Results indicated students spent more time reading orally than silently and were more likely to fixate on words, especially low frequency words, during oral reading. The authors, however, noted concerns regarding the internal validity of the study resulting from difficulties with the reliability of their eye-tracking data collected while students read passages orally. Specifically, the authors noted the eye-tracking equipment was prone to “losing” the location of the eye gaze when a student’s head was not stabilized, such as during oral reading. Thus, eye-tracking associated with silent reading cannot yet be easily or reliably used with oral reading. The current study bypasses this limitation with eye tracking equipment by limiting eye movement measures to only silent readings.

Groundwork for the Current Study

The current study extends the work of Valle et al. (2013), which utilized eye-tracking and spectrographic technologies to describe and analyze second grade students’ silent and oral reading behavior as those behaviors relate to textual characteristics. Valle et al. (2013) had students read two different passages, one silently and one orally. Students’ silent reading behavior was analyzed using common eye-tracking measures, and their oral reading was analyzed using spectrographic analyses.

When reading the silent passage, Valle et al. (2013) measured students' reading of six HF and six LF target words matched by word length. Eye movement measures included number of regressions, total time, gaze duration, number of skips, number of fixations, and first fixation duration. Students' oral reading was analyzed using spectrographic analyses. For the oral reading condition, spectrographic analyses were used to provide visual depictions of oral reading created by placing recordings on a time scale (plotted on the horizontal axis) with the frequency of sound waves on the vertical axis. Using spectrographic analyses, students' behavior related to five HF and five LF words were analyzed. Specifically, the researchers assessed the duration students paused before reading each HF and LF word, the duration required for them to read each word, and the duration of their pauses following the reading of HF and LF words. The authors also measured the length of students' reading pauses at commas, periods, and the end of paragraphs and evaluated differences in students' reading behavior related to content versus function words.

Results of the eye-tracking data collected by Valle et al. (2013) suggested that students read HF words more quickly than LF words both orally and silently suggesting that word frequency influences students' reading behaviors similarly between reading modalities. On measures of gaze duration, total time, and number of fixations students silently read HF words more quickly than LF words. Word frequency was found to interact with student oral reading skill. Average-skill students took longer to orally read LF than HF words, whereas high-skill students paused longer before LF words than HF words but maintained their overall times spent on HF and LF words. Although average-skill students had longer durations reading LF words, their approach to tackling reading complexity reflected less differentiation in their reading times between HF and LF words (Haenggi & Perfetti, 1994; Valle et al., 2013) suggesting words frequently encountered for these students are approached similarly as words infrequently encountered.

Valle et al. (2013) found that high-skill second grade readers made nuanced pausing between delimiter types when measures of eye-tracking behavior were used; however, average-skill readers were more likely to experience extended delays at paragraph periods, likely related to students' difficulty tracking to the beginning of the next paragraph. Similar results were found by Schwanenflugel, Hamilton, Kuhn, Wisenbaker, and Stahl (2004) suggesting skill and more nuanced pausing are positively correlated. Valle et al.'s results suggest that consideration of syntactic structure, or the use of commas and the placement of sentence periods and paragraph endings, might affect how beginning readers interact with passages.

Valle et al. (2013) suggested that content and function words are related to students' reading approaches based on students' reading skill, with better readers skipping fewer content words and more function words than poorer readers. Better readers' pattern of skipping corresponds to retaining more meaning from text while maintaining reading efficiency, an interpretation similarly provided by Rayner (1998). Valle et al. (2013) suggested maintenance of content words during reading, while skipping function words, was an approach seen in adult readers as well.

Interestingly, little information exists in the literature relating oral reading errors to the textual variables of passages, although there are both quantitative and qualitative studies to reference. Quantitative studies have compared students with dyslexia or with traumatic brain injuries to those without impairments producing inconclusive results. Qualitative research by Yetta Goodman from the 1960-70s suggests instructional decisions might be made by analyzing the pattern of errors made across a student's oral reading. Goodman (1965) found that more skilled children made semantic errors that better maintained the meaning of passages when compared to less skilled readers who made more orthographic errors. Goodman maintained that more nuanced sub-skill information (e.g., decoding, reading comprehension strategies) might be gained by qualitatively analyzing the degree to which the student's errors "disrupt the meaning" of the text and how the student responds to those errors (p.534, Goodman 1972, version 1997).

Purpose of the Current Study

The purpose of the current study was to replicate and extend the study methodology conducted by Valle et al. (2013). The current study examined second grade students' eye movements while reading a passage silently and their vocal durations and pauses associated with oral readings of words on the same passage. Additionally, oral errors vocalized during students' oral readings were analyzed. The passage was constructed to contain words of varying frequency, or familiarity, so that students' oral and silent reading behavior could be analyzed in relation to textual difficulty. Although research suggests differences exist in students' total time and gaze duration across HF and LF words, researchers have yet to examine differences using more distinctive categories of word types that include function words and HF, medium frequency (MF), and LF content words. The study also extends Valle et al.'s (2013) research on skips made by students by examining skips on words preceding HF and LF target words within the passage. This study analyzed oral reading error types and their relation with word frequencies. Spectrographic data was compared between high and low frequency target words to determine if it is consistent with prior research.

An additional purpose of this study was to extend comparisons of oral and silent reading behavior so that oral reading could be observed at the word level, a method only recently accomplished with Valle et al. (2013) and Rogers and Ardoin (2014). Students within the current study read the same passage both silently and orally, allowing for an evaluation of the relationship between their eye movements while reading silently and their speed and error types while reading orally. This study further extends the work of Valle et al. (2013) by examining readers' silent and oral reading performance on the same passage; thus, allowing a more direct evaluation of textual variables between oral and silent reading. The methods of this study also extend the work of Rogers and Ardoin (2014) where oral and silent reading were evaluated with the same passage; however the current study uses different measures on oral and silent readings. Analyses in the current study will also provide information regarding the implications of using textual variables in the evaluation of passage difficulty. The overall purpose of the current study was to provide insight into how textual characteristics might be used in the creation of equivalent passages.

CHAPTER 2

METHODS

Participants and Setting

Participants included 56 second-graders across 12 classrooms evenly distributed across three suburban schools outside a midsize city in the Southeastern United States. Criteria for inclusion included English as the child's primary language and broad reading assessment scores between the Below Average and Superior range, as measured by the Woodcock-Johnson, Third Edition, Achievement (WJ-III Ach), Broad Reading composite (Woodcock, McGrew, & Mather, 2001). Special education and students with Section 504 plans were excluded from participation. From the original 56 participants, two students' eye-tracking data were lost because of eye-tracking equipment calibration loss (i.e., securing connection between video camera and students' eye gaze). The remaining 54 participants were majority Caucasian (86.0%) with smaller proportions of the sample representing Black (1.8%), Asian (3.5%), Hispanic (3.5%), or multiracial (3.5%) ethnicities. The sample consisted of 50.9% females with an average age of 7 years, 11 months (range = 7 years, 7 months – 8 years, 7 months). Free or reduced-price lunch rates across the three schools ranged from 18-29%. Although all students in the sample received a regular education curriculum, 22.8% participated in gifted-talented programming and 8.8% participated in additional support services offered through Title 1 programming. Achievement rates on state benchmarking tests indicated that 97.9% to 99.1% of the schools' student populations met or exceeded state standards in English Language Arts/Reading in the spring of 2011.

Parental consent was obtained as part of a larger study from which the data set employed in this study was obtained. Participants were administered the WJ-III ACH Broad Reading subtests (Woodcock et al., 2001). Total Broad Reading Achievement scores for this sample ranged from 91 to 132 with average and median standard scores of 111 (*SD* = 8.88). Students between the first and third quartile within this sample scored between 106 and 117, indicating most students in the sample were performing in the Average to High Average ranges.

Materials

Analyses were based on students' silent and oral reading of a narrative passage consisting of 167 words, with a Spache (1953) readability index of 3.23. For global analyses, all words were first divided into content and function word groups, then the content words were divided into high, medium, and low frequency groups by evenly dividing the words by number, which resulted in four "word type" groups of all words in the passage. Word frequency estimates as reported in *The American Heritage Word Frequency Book* (Carroll, Davies, & Richman, 1971) ranged for each word type as follows: high frequency (HF; 211.37 to 997.29 *U*), MF (42.29 to 190.35 *U*), low frequency (LF; 0.86 to 33.97 *U*), and function words (49.88 to 4678.30 *U*).

For target word analyses, 5 HF and 5 LF target words were chosen from the passage post hoc using a systematic method. First, consideration was given to words' frequency of occurrence in children's literature as reported in *The American Heritage Word Frequency Book* (Carroll et al., 1971). Specifically, content words with the highest and lowest frequencies were considered. Next, error rate by participants on these words was considered; namely, by identifying from the HF word group those words with the fewest errors and from the LF word group those words with the highest number of errors. Finally, textual characteristics that lent to creating paired matches of high and low frequency target words were considered. Words were matched on the most dimensions of word characteristics possible (i.e., part of speech, word length, number of syllables, and sentence and paragraph placement, See Table 1), and at times this required choosing words that were less extreme in frequency or error rate. LF target words with high percentages of errors and HF target words with low percentages of errors were chosen post hoc to insure that target words represented a range of difficulty within the passage; further, it insured that the target words of this study extended the results of prior studies that suggest that frequency is an indicator of difficulty. By matching pairs of HF and LF words on textual characteristics, analyses involving target words would have less variance unrelated to word difficulty. The estimated appearance per million words (average *U* count) was 172.51 for HF target words, with a range of 53.45 to 417.67 and 7.33 for LF target words, with a range of 1.40 to 10.48 (Carroll et al., 1971).

During the silent reading condition, the passage was displayed on a computer monitor in black, 20-point, 1.5-spaced, Times New Roman font against a white background. During the oral condition, the test passage was displayed on an 8.5 × 11 inch white paper with black 14-point, 1.5-spaced Century Gothic font.

Apparatus

To track eye movements and compute gaze position on the display, an SR Research EyeLink 1000 system was used, which has a sampling rate of 1000 Hz, resolution of 0.01° of visual angle, and a range of 32° horizontally and 25° vertically. The SR Research Experiment Builder software package was connected to the EyeLink system to control eye calibration, passage presentation, and data collection. Passages were presented on either a 19-inch or a 22-inch LCD monitor. Participants indicated reading completion via a Microsoft Sidewinder Plug and Play game pad.

During the oral reading session, the test passage reading was recorded digitally using an Audio Technica AT2020USB Condenser USB Microphone connected to a laptop PC. Students' oral reading was then analyzed for durations of pauses before and after target words as well as durations for reading target words using PRAAT (version 4.2.1; Boersna & Weenink, 2003). PRAAT is a computer program that allows for the estimation of the boundaries of words to the millisecond using digital audio recordings. PRAAT software allows for spectrographic analyses, or visual analyses, of oral reading by placing recordings on a time scale to the thousandth of a second (on the horizontal axis) with the frequency of sound waves on the vertical axis. Target word duration was measured from the start of the target word, including all decoding sounds, until the end of the target word. The silences preceding and following target words were measured using the end boundary of the word preceding the target and beginning boundary of the word following the target.

Procedures

Data for the study were collected for each student as part of a larger study. Students were given the broadband reading measure on one day and within a few days they completed the second day, which contained both the silent and oral reading sessions. A total of six passages were read during silent reading sessions with the stimulus passage read second and four other passages read silently afterward. After the silent reading, students were provided with a brief break (5-10 min), and then were asked to read the stimulus passage orally. The combined silent and oral readings required approximately 30 to 45 min in total of out-of-class time to complete.

Silent reading. During silent reading sessions, individual students were escorted to a dimly lit room and told that a camera would watch their eyes as they read passages. Two trained experimenters worked with each student. One experimenter positioned the student at the eye tracker and sat next to him or her to provide instructions. The second experimenter controlled the eye-tracking computers, allowing for probe administration and data collection. Students were seated with their chins resting on a chin rest approximately 55 cm from a camera that was positioned directly in front of and below the center of a monitor. The eye tracker was calibrated to students' eye positions as they gazed onto the screen and completed a "game" that required them to follow a dot with their eyes over a 9-point grid. After students completed a familiarization task of a passage and comprehension question, a researcher then recalibrated the eye tracker to the eye gaze, reminded the child that there would be a comprehension question after each passage, and presented the first of the six silent reading session passages. Students read one passage associated with the larger study, read a second passage which was the reading analyzed for the purposes of this study, and finally read a third passage with the third passage being read four consecutive times. Following the silent readings, students took a 5-10 min break before again reading the second passage, this time orally and with no other passages.

Oral reading. This session occurred with one experimenter in a quiet room after the silent reading session. Only one passage, the stimulus passage for this study, was read aloud by each student. The experimenter provided instructions that requested students to do their best reading, informed students that they would not receive help on words, and encouraged students to keep reading after any 10 second (s) pause or 10 s stumble on a word. Error correction was not provided on any occasion; rather, the experimenter instructed students to move to the next word in the passage after 10 s.

Coding of Silent and Oral Readings

Silent reading coding. Students' silent reading of the passage was evaluated using four measures commonly assessed within the eye-tracking literature: gaze duration, total time reading, number of fixations, and number of regressions (Rayner et al., 2006; see Table 2). Gaze Duration is the total amount of time a participant fixated on a target word the first time it is encountered, including all fixations prior to regressions or forward saccades to a new word. Total Time is the total amount of time spent reading a target word including fixations resulting from a regression from subsequent text. Number of Fixations is the total number of times a target word is fixated upon during the reading. Number of Regressions indicates the number of regressions to a target word from text subsequent to the target word.

Oral reading coding. Students' oral reading of all words within the passage was coded using scoring guidelines similar to typical CBM scoring procedures (Deno, 1985), except that students were not given a word on which they hesitated for 3 s. Rather, students were given 10 s to decode each word and, if unsuccessful, were prompted to continue reading the next word of the passage.

All reading errors were coded as being one of six error types (See Table 3) from four basic reading behaviors. Miscues were mispronunciations of a word that changed the meaning of the immediate context of the word. Insertion of words into the passage was ignored, but insertion of sounds to words was counted as a miscue. For example, *the big dog* where *big* was added to the text by a student was ignored (i.e., not counted as an error) whereas *the dogs* where *s* was added to *dog* was scored as a miscue. Omissions were students' exclusion of words in the text. Reversals were coded when students transposed two consecutive words. Hesitations occurred when students stalled on a word while either vocally attempting to decode or remaining silent for at least 3 s. Notably, a Hesitation alone means the student read the word correctly after a 3 s delay—they were given 10 s total to decode each word. The remaining two types of errors combined Hesitations with Miscues and Omissions such that an error type of Hesitation-Miscue represents the student hesitating on a word before mispronouncing it, and Hesitation-Omission represents the student hesitating before skipping a word. Self-corrections were not counted as errors except where Hesitations occurred.

In addition to scoring students' reading errors during oral reading, their oral reading was analyzed using PRAAT software (Boersma & Weenink, 2011). The researcher used the software to set boundary marks so that for each of the 10 target words (5 HF and 5 LF) the durations of Pre-word Pauses, Time-on-Word, and Post-word Pauses could be measured (See Table 4.) Specifically, the Pre-word Pause included the duration of silence between the pre-target word until the student began the target word. Time on Word was measured from the first utterance of the target word through the completion of the word, including decoding efforts. Post-word Pauses were measured silences from the end of the target word until the start of the post-target word.

Data Analyses

Global analyses using all words of the passage were performed to evaluate silent reading data and oral error data. A repeated measures ANOVA was performed with each eye-tracking measure (four total) across word types (i.e., high, medium, and low frequency content words and all function words). Oral reading error data were analyzed globally with repeated measure ANOVAs for each word type across error types.

PRAAT data were gathered for only the ten target words. Oral reading spectrographic (PRAAT) data were analyzed with the Wilcoxon signed-rank test to examine differences for each of the three measures between high and low frequency target words. Spearman's correlation coefficient was used to identify correlations between spectrographic and eye-tracking measures when looking specifically at high and low frequency target words.

An additional analysis of words immediately preceding HF and LF target words was performed via a t-test to identify differences in the number of skips (via the Total Time eye-tracking measure) of these words was performed after review of the data. Specifically, after identifying a pattern of skips on these words during the oral readings, the authors added this analysis to determine significance between skip rates for words preceding HF and LF words.

Table 1

Textual Characteristics of High and Low Frequency Target Word Pairings.

Stimulus Word Pairings	Stimulus Word	Word f U	Error Rate n	Rate of Errors M	End of Sentence	End of Paragraph	Word Length n	Syllables n
LF 1	necklace	7.56	8	12.9	no	no	8	2
HF 1	opened	127.43	1	1.6	no	no	6	2
LF 2	Purse	8.79	8	12.9	no	no	5	1
HF 2	Lunch	53.45	1	1.6	no	no	5	1
LF 3	Wallet	1.40	11	17.7	no	no	6	2
HF 3	mother	417.67	1	1.6	no	no	6	2
LF 4	Limbs	10.47	23	37.1	no	no	5	1
HF 4	Truck	73.64	1	1.6	no	no	5	1
LF 5	quarrel	8.43	29	46.8	no	no	7	2
HF 5	Friends	190.35	2	3.2	no	no	7	1

Note. U is the frequency of a word's appearance per 1 million words of children's literature. HF identifies a High Frequency target word and LF identifies a Low Frequency target word.

Table 2

Eye Tracking Measure Labels and Definitions.

Measure Label	Definitions
Gaze Duration	Total amount of time spent on a target word the first time it is encountered
Total Time	The total amount of time spent on a target word including fixations resulting from a regression from subsequent text
Number of Fixations	The total number of times a target word is fixated upon during the reading
Number of Regressions	The number of regressions to a target word from subsequent text

Table 3

Error Labels And Definitions Applied Globally To Oral Readings.

Measure Labels	Definitions
Miscue	Without hesitation, misread word so that meaning changed. For instance, “dog” for “mother.”
Omission	Skipped words without hesitation.
Reversal	Transposition of two consecutive words.
Hesitation	Correct with hesitation of 3 s or more.
Hesitation-Miscue	Misread word with hesitation of 3 s or more.
Hesitation-Omission	Skipped word with at least 3 s hesitation.

Table 4

Spectrographic Measure (PRAAT) Labels and Definitions.

Measure Labels	Definitions
Time on Word	Duration of word excluding Pre- and Post-word Pauses
Pre-word Pause	Pause between target word and the immediately preceding word
Post-word Pause	Pause between target word and the immediately subsequent word

CHAPTER 3

RESULTS

Descriptive statistics for eye-tracking measures are presented in Table 5 and descriptive statistics for oral spectrographic and oral error data are presented in Table 6. Across students, errors on target words ranged from 0 to 4 errors out of 10 possible target words. The average number of LF errors for the 5 LF words across all participants was 0.06 with a range from 0 to 4 errors. HF errors averaged 0.01 (range = 0-2) for the 5 HF words across all participants.

Analyses to examine normality for the four eye-tracking measures indicated a positive skew and positive kurtosis, both overall for each measure and for each measure by word type. Evaluation of normality of eye-tracking data for only HF and LF words revealed HF data met normality standards while LF data consistently had a positive skew and positive kurtosis. Log(x) transformations were conducted to correct extreme outliers (i.e., z-score > 3.29). Assumptions were not met for overall PRAAT data except that HF time on word met the assumptions of normality. All other PRAAT data revealed a negative skew and a positive kurtosis. Transformations did not improve the normality of PRAAT data. Although parametric tests were used with transformed eye-tracking data, tests found to be robust to non-normality in data (i.e., Wilcoxon signed-rank test and Spearman's correlation coefficient, Field, 2009) were used to evaluate spectrographic data for target word differences and for analyses of PRAAT and eye-tracking data correlations.

Silent Reading Analyses

Four repeated-measure ANOVAs were performed to identify differences within eye-tracking behavior across word types. Significant differences between frequency levels were found for gaze duration, $F(2.4, 93.6) = 31.591$, $p < .001$, total time, $F(3, 159) = 111.389$, $p < .001$, regressions, $F(3, 159) = 46.035$, $p < .001$ and fixations, $F(3, 159) = 128.319$, $p < .001$. Pairwise comparisons for each eye movement measure are presented in Table 7. Pairwise comparisons revealed that for all measures except gaze duration, students' behavior on HF and MF words was similar in level to each other and were significantly ($p < .001$) shorter than LF total time durations and significantly ($p < .001$) lower than LF number of fixations. Similarly, for all measures except gaze duration, students' behavior on HF and MF words was significantly ($p < .001$) higher than LF words for number of regressions. Students' behavior for HF and MF words was significantly greater ($p < .001$) than function words in total time and number of fixations. Students' behavior for HF and MF words was significantly lower ($p < .001$) than function words for number of regressions. Students' behavior for gaze duration was similar between HF and function words, and both measures were significantly ($p < .001$) shorter in duration than MF and LF words.

An analysis of the skips made by students on the words immediately preceding HF and LF target words was completed due to observations of oral reading data revealing frequent skips. Words immediately preceding target words were all short (three or fewer letters) and served the purpose of a preposition or an article within the passage (i.e., function words). During silent readings, function words preceding HF target words had an average skip rate of 1.2 ($SD = 1.17$) words per participant, and words preceding LF target words had an average skip rate of 2.3 ($SD = 1.23$) per participant. A t-test of skip counts using Total Time as the measure of interest revealed significantly more skips occurred on words preceding LF target words than words preceding HF target words, $t(53) = -5.964$, $p < .001$.

Oral Reading Analyses

An ANOVA was completed for oral error data for each of the four word types (i.e., level of frequency [high, medium, and low] across all content words within the passage and separately for all function words). Significant differences between error types were found for the HF, $F(5, 170) = 15.442$, $p < .001$ and LF words, $F(1.599, 84.732) = 12.625$, $p < .001$, but not for MF words, $F(1.369, 72.574) = 4.682$, $p < .05$. Across function words no differences were found, $F(2.51, 133.01) = 1.218$, $p = .304$. Specific differences for word types are presented in Table 8. Pairwise comparisons within the HF level revealed misreads and skips without hesitations to be similar in number and significantly ($p < .001$) more frequent than all other error types. At the LF level, misreads with and without hesitations were similar in number and occurred significantly ($p < .001$) more often than all other error types.

Oral reading spectrographic data were analyzed with Wilcoxon signed-rank tests to estimate significant differences between HF and LF target words. Across all three PRAAT measures, significant differences ($p < .001$) were found. For each measure (i.e., time on word, $Z = -6.37$, pre-word pause, $Z = -5.75$, and post-word pause, $Z = -5.69$), LF words had significantly longer durations than HF words.

Oral and Silent Reading Analyses

Spearman's correlation coefficients were examined for PRAAT and eye-tracking measures by HF and LF target words (see Table 9). Time on word was significantly correlated ($p < .01$) with number of fixations and number of regressions for both HF and LF target words. Time on word was also significantly correlated with LF target word gaze duration ($p < .01$), HF target word total time ($p < .01$), and HF target word gaze duration ($p < .05$). Pre-word pause had two significant correlations (p < .01) with LF target word number of fixations and total time. Post-word pause was significantly correlated ($p < .01$) with HF and LF target word total time and number of regressions, LF target word gaze duration, and HF target word number of fixations.

Table 5

Descriptive Statistics of Silent Reading (Eye Tracking) Measures.

Measure	Mean (SD)	Measures	Mean (SD)
Measures across all words		Measures by Word Type	
Total Time	511.04 (402.71)	HF Total Time	2.65 (0.02)
Gaze Duration	365.97 (254.12)	HF Gaze Duration	2.5 (0.02)
Number of Fixations	1.60 (1.39)	HF Number of Fixations	0.42 (0.01)
Number of Regressions	0.32 (0.58)	HF Number of Regressions	0.84 (0.45)
		MF Total Time	2.64 (0.02)
Measures for Target Words		MF Gaze Duration	2.64 (0.02)
HF Target Total Time	2.70 (0.02)	MF Number of Fixations	0.43 (0.01)
HF Target Gaze Duration	2.54 (0.02)	MF Number of Regressions	0.84 (0.38)
HF Target Number of Fixations	0.46 (0.01)	LF Total Time	2.75 (0.02)
HF Target Number of Regressions	0.13 (0.01)	LF Gaze Duration	2.61 (0.02)
LF Target Total Time	2.93 (0.03)	LF Number of Fixations	0.49 (0.01)
LF Target Gaze Duration	2.74 (0.03)	LF Number of Regressions	0.69 (0.43)
LF Target Number of Fixations	0.59 (0.02)	Function Total Time	2.55 (0.02)
LF Target Number of Regressions	0.15 (0.01)	Function Gaze Duration	2.55 (0.02)
		Function Number of Fixations	0.36 (0.01)
Counts for missing words preceding target words		Function Number of Regressions	1.22 (.23)
HF Target Words	1.20 (1.17)		
LF Target Words	2.31 (1.23)		

Note. Eye-tracking measure means and standard deviations by word type and target words are given log(x) transformation.

Table 6

Descriptive Statistics of Oral Reading (PRAAT, Oral Errors) Measures.

Measure	Mean (SD)	Measures	Mean (SD)
Measures across all words		Measures for Word Type	
Pre-silence	0.45 (1.26)	HF Target Frequency of Error 1	0.00 (0.00)
Word Duration	0.64 (0.44)	HF Target Frequency of Error 2	0.28 (0.81)
Post Silence	0.26 (0.40)	HF Target Frequency of Error 3	0.31 (0.54)
Frequency of Error 1	0.002 (0.04)	HF Target Frequency of Error 4	0.06 (0.30)
Frequency of Error 2	0.01 (0.10)	HF Target Frequency of Error 5	0.00 (0.00)
Frequency of Error 3	0.003 (0.05)	HF Target Frequency of Error 6	0.00 (0.00)
Frequency of Error 4	0.002 (0.04)	MF Target Frequency of Error 1	0.06 (0.23)
Frequency of Error 5	0.0003 (0.02)	MF Target Frequency of Error 2	0.35 (1.05)
Frequency of Error 6	0.0009 (0.03)	MF Target Frequency of Error 3	0.04 (0.19)
		MF Target Frequency of Error 4	0.07 (0.33)
Measures for Target Words		MF Target Frequency of Error 5	0.04 (0.27)
HF Target Pre-silence	0.03 (0.06)	MF Target Frequency of Error 6	0.00 (0.00)
HF Target Word Duration	0.46 (0.13)	LF Target Frequency of Error 1	0.28 (0.81)
HF Target Post Silence	0.14 (0.25)	LF Target Frequency of Error 2	0.96 (1.64)
LF Target Pre-silence	0.87 (1.68)	LF Target Frequency of Error 3	0.02 (0.14)
LF Target Word Duration	0.83 (0.55)	LF Target Frequency of Error 4	0.07 (0.33)
LF Target Post Silence	0.37 (0.48)	LF Target Frequency of Error 5	0.00 (0.00)
		LF Target Frequency of Error 6	0.13 (0.34)
		Function Target Frequency of Error 1	0.00 (0.00)
		Function Target Frequency of Error 2	0.02 (0.14)
		Function Target Frequency of Error 3	0.04 (0.19)
		Function Target Frequency of Error 4	0.07 (0.33)
		Function Target Frequency of Error 5	0.02 (0.14)
		Function Target Frequency of Error 6	0.02 (0.14)

Table 7

Eye Tracking Measure Averages Presented by Word Type.

Word Type	Eye Tracking Measures				
		Average Time		Average Frequency	
	U range	Gaze Duration	Total Time	Regressions	Fixations
High (e.g., father, read)	211.37 to 2468.90	2.50 (.02)[a]	2.62 (.02)[a]	0.84 (0.45)[a]	0.42 (0.01)[a]
Medium (e.g., friends, candy)	42.289 to 193.67	2.64 (.02)[b]	2.64 (.02)[a]	0.84 (0.38)[a]	0.43 (0.01)[a]
Low (e.g., backpack, necklace)	0.0243 to 36.942	2.61 (.02)[b]	2.75 (.02)[b]	0.69 (0.43)[b]	0.49 (0.01)[b]
Function (e.g., to, an, it, the)	0.114 to 73,122.00	2.55 (.02)[a]	2.53 (.02)[c]	1.22 (0.23)[c]	0.36 (0.01)[c]

Note. U gives the frequency of a word per million words while adjusting for the word's presence in different subjects and grade levels. High frequency words are commonly encountered by students while low frequency words are seldom encountered by students. Values with same superscripts did not differ significantly from each other (within each eye-tracking measure, or column) whereas different superscripts denote significant differences ($p < .001$) across values within the same column.

Table 8

Oral Reading Error Type Distribution for Each Word Type.

Word Type	n	Error 1	Error 2	Error 3	Error 4	Error 5	Error 6
		Percentage of Error Occurrence within Word Type					
Total Percentage of Errors Averaged across Word Type	151	38	32	17	6	2	5
Content Words by Frequency							
High	35	0^{a}	43^{b}	49^{b}	9^{a}	0^{a}	0^{a}
Medium	30	3^{a}	70^{a}	7^{a}	13^{a}	7^{a}	0^{a}
Low	79	19^{ab}	66^{a}	1^{b}	5^{b}	0^{b}	9^{b}
All Function Words	9	0^{a}	11^{a}	22^{a}	44^{a}	11^{a}	11^{a}

Note. Errors: 1) Misread with hesitation, 2) Misread without hesitation, 3) Skip without hesitation, 4) Skip with hesitation, 5) Transposed words, 6) Correct with hesitation. Frequency ranges for error words by word type: High Frequency (211.37 to 997.29 U), Medium Frequency (42.29 to 190.35 U), Low Frequency (0.86 to 33.97 U), and Function Words (49.88 to 4678.3 U). Values with the same alphabetic superscript do not differ significantly from each other (within each word type) whereas different superscripts denote significant differences ($p < .001$ at the high and low frequency level and $p < .05$ at the MF level) across values within the same row.

Table 9

Spearman Correlation Coefficients between Silent Reading Measures and Oral Reading Spectrographic Measures by Target Word Frequency.

Eye Tracking Measures	Spectrographic (PRAAT) Measures		
	Time on Word	Pre-word Pause	Post-word Pause
Gaze Duration			
High Frequency Target Words	.339*	.027	.163
Low Frequency Target Words	.434**	.216	.363**
Total Time			
High Frequency Target Words	.433**	-.043	.219**
Low Frequency Target Words	.414	.224**	.358**
No. of Regressions			
High Frequency Target Words	.023**	-.157	.021**
Low Frequency Target Words	-.020**	-.109	-.138**
No. of Fixations			
High Frequency Target Words	.363**	-.045	.191**
Low Frequency Target Words	.342**	.181**	.312

*$p < .05$, **$p < .01$

CHAPTER 4

DISCUSSION

CBM-R passages are used within schools as a means of assessing students' reading achievement and monitoring their response to intervention (Reschly et al., 2009). Unfortunately, research suggests that the passages used to monitor students' response to intervention generally vary in level of difficulty resulting in inaccurate estimates of student growth (Ardoin & Christ, 2009; Rayner, Ardoin, & Binder, 2013). The use of eye-tracking technology, however, might facilitate researchers' evaluation of text difficulty by allowing for a more detailed measurement of variables that might influence passage difficulty. By employing both silent and oral reading tasks in the current study, an evaluation of the relationship between students' eye movements and oral reading rate and errors on individual words was possible.

Results Consistent with Past Findings

Data from the eye-tracking procedure used in this study revealed silent reading behavior differed across word type. Results were consistent with past research indicating HF words are more quickly and easily read than LF words. Specifically when reading LF words, longer total times, more fixations, and fewer regressions were observed as compared to HF words. Consistent with the empirical literature (Rayner & Liversedge, 2011; Valle et al., 2013), results from this study suggested words encountered infrequently are given more attention during the initial reading pass (i.e., more fixations), suggesting decoding, while higher frequency words are given less initial pass attention but are more frequently returned to later in the passage, suggesting later processing.

Function word results were also consistent with previous research where students attended to function words by minimally fixating on them and more frequently returning to them from subsequent text (Rayner & Liversedge, 2011). These results might indicate an array of complex reading processes, such as decoding at the word level, rereading the text to improve decoding, or using regressions to infer meaning. Eye-tracking technology does not yet allow accurate conclusions about the relationships between the measures obtained on one word to another word in a passage. Future research might support identifying patterns of eye movements within clauses, sentences, or phrases that include HF and LF words to better understand how students move between words to engage in initial and later processing of text. Given the range of frequencies of function words within the stimulus passage of this study, future research might delineate more stringent categories of function words in order to distinguish varied textual effects on reading behavior.

Spectrographic data analyses of students' oral reading of the passage suggested that they made significantly longer pauses before and after LF target words and read LF target words for longer durations when compared to HF target words. These data are similar to prior research by Valle et al. (2013) suggesting LF words consume more of students' reading time than HF words. These data inform CBM-R passage development by suggesting that individual words can be divided into informative categories that provide meaningful information about the difficulty of passages. Further, students' behavior on the individual words of passages is observable on oral readings, which might help initial development of more equivalent passages using technology more readily available and usable than eye-tracking equipment.

Results Extending Past Research

Eye-tracking measures. The unique contribution of this study was that word type might be usefully delineated by more than just high and low frequency of content words. It might be beneficial to categorize words into more than two word types (i.e., HF, MF, LF, and function) to increase the extent to which categories differentiate students' familiarity with the words. MF words corresponded with both HF and LF words on different measures of eye movement. HF and MF words were similar on measures of total time, number of fixations, and number of regressions. Students' readings of HF and MF words, when compared to LF words, had shorter total times and fewer fixations and regressions. MF words were more similar to LF words when considering Gaze duration. This suggests that students' fixations during their initial encounter with a word were similar on both MF and LF words indicating similar initial processing of these word types. These data indicate that, on these measures, HF and MF words function similarly within a passage when considering difficulty, or when students engage in higher linguistic processing of text (Joseph, Nation, & Liversedge, 2013). These results might not hold when different words or even the same words are used in a different passage. Further research elucidating how frequencies of words impact student reading behavior might provide insight into how CBM-R passages can be developed to accommodate the incorporation of MF words.

Results suggested that function words might also have a unique contribution to passage difficulty in that they function distinctly from HF, MF, or LF words. Function words had shorter total times and number of fixations and regressions than all other word types. Gaze duration data provided evidence indicating students spent similarly shorter amounts of time reading function and HF words the first time such words are encountered than MF or LF words; however, total time data indicated students spent more time on HF words than function words after considering cumulative reading passes. Although students regressed to function words most often, their total time on HF words was more than that on function words, which suggests students might have fixated on HF words longer to decode. Alternatively, students might have regressed to function words when comprehending at the sentence level, which would support interpretations by Rayner and Liversedge (2011). These data need baring out by further research that examines the relationship between eye movements of separate words within a passage. CBM-R passage development might be improved with consideration of function words in relation to student behavior.

Based upon observations made when analyzing students' oral reading errors, which suggested that students made more oral omissions of function words preceding LF target words than HF target words, an analysis was conducted to examine silent reading skips of function words for both HF and LF target words. An analysis of total time revealed that LF target words were associated with more frequent skips of the function word that immediately preceded LF target words than function words preceding HF target words. These results might suggest students changed their reading behavior based on the difficulty of the words within the passage. These results also suggest students' behavior might be affected by the relationship of words within a passage, namely that relative frequency of words might have a degree of impact on the reading behavior of students. Valle et al. (2013) suggested that high-skill students who made more frequent skips of function words while meticulously reading content words did so to improve reading efficiency. Given that the current study's sample made more skips on function words preceding LF content words, results from the current study support Valle et al.'s interpretation.

Oral reading errors. The oral reading error type analyses provided information regarding the relation between error types and word frequency. Oral reading error types were evaluated by word type, revealing that HF and LF word types, but not MF or function words, differed in the proportions of error types associated with each word type. HF words corresponded with more misreads and omissions without hesitations than other error types. HF words also generally resulted in fewer errors than MF or LF words. LF words corresponded with more misreads with and without hesitations. By recognizing that words of varying frequency contribute to student oral reading performance differently, CBM-R passages can be equated better to prevent situations as the example described above. Although the current study does not distinguish between substitutions of incorrect words for correct words and phonological errors on LF words, this distinction would be an important area for future study given the general focus of qualitative research on miscue analysis (Goodman, 1965). Given the lack of quantitative background analyzing reading errors (Goodman, 1965), future research might address which errors are most related to textual characteristics and eye movements.

Spectrographic measures related to eye movements. The spectrographic results of this study allowed for analyses which moved beyond prior research on the association between oral reading durations and silent reading eye movements. Oral and silent reading measure correlations revealed the time spent reading a word aloud and the pauses before and after words read aloud were related to silent reading eye movements in a variety of ways. HF target word data suggested second grade students demonstrated proportionally similar times on easier words when reading either silently or orally. With LF target words, students also spent proportional amounts of time reading orally and silently during their first-pass, which extends findings by Rogers and Ardoin (2014). Students had fewer instances of regressing to LF words later in the passage suggesting students had less difficulty with incorporating LF words into the context of the sentence (i.e., comprehending). Conversely, students most frequently regressed in oral and silent reading to function and HF words. All of these results compliment and extend the findings of Valle et al. (2013) by establishing that elementary readers demonstrate comparable behaviors during oral and silent readings on the same passage, which is suggestive of the influence of textual characteristics.

Students' pausing after HF and LF target words was related to eye movements suggesting later processing (Total Time, Number of Regressions), findings that were not indicated by Valle et al. (2013). LF target words had pre-word pauses that corresponded with eye movements indicative of initial processing (Number of Fixations, Total Time). This last finding supports work by Valle et al. (2013) as students in that study paused just before relatively difficult words. The sample of students from the current study might reflect the highly skilled readers similar to those who participated in the Valle et al. (2013) study.

In summary, the correlations between spectrographic measures and eye movement measures in general provided evidence that students' oral reading behaviors are proportionally related to their eye movements when reading silently. Specifically, longer oral reading durations in time-on-word, pre-word pause, and post-word pause all correlated significantly with eye movements during silent reading suggesting increased time spent reading words orally translate into longer time spend reading these words silently. Additionally, more difficult words (i.e., LF target words) appear to require more time during first pass reading and are not often reread, whereas easier words (i.e., HF target words, function words) seem to be read quickly during students' first pass but will be reread more frequently. Importantly, these results demonstrate that passage difficulty during oral readings corresponds with the difficulty of specific words within the passage during silent readings. Implications of spectrographic and eye movement analyses suggest that HF and LF words can be used as proxies within a passage to better estimate the difficulty of a passage than perhaps other textual characteristics, or in combination with other textual characteristics.

Limitations

Limitations to the present study should be considered in the interpretation of findings. As a function of the larger study for which these data were collected, students always read the passage first silently and then aloud. The results for the oral session thus might reflect practice effects because the same passage was used for both conditions. Practice effects might have been mitigated by four readings of other passages between the silent and oral sessions of the assessment passage. These four passages were read after the stimulus passage for this study during the silent session. The students also received a short break when they moved to another room to begin the oral reading session.

A second limitation is that the word types within this study were chosen based on the range of frequencies of all content words within the stimulus passage. Although this provided a range of word frequencies, the delineations for each level of frequency (i.e., high, medium, low) might not be consistent with other research using categorical groupings of word frequency. No guidelines for choosing a high frequency versus a low frequency word were found within the literature, suggesting inconsistency between studies' definitions of high and low frequency words. A rough guideline used by the current study was to model identification of HF and LF words after Ardoin et al., (2013), which identified HF words as those appearing at least 40 times out of a million words in children's literature and LF words as those appearing less than 10 times per million words in children's literature.

Similarly, global analyses for this study used a MF word type and a function word type in addition to commonly used HF and LF word types. Unfortunately for target word analyses three LF target words had frequencies that placed them in the MF group for global analyses. These overlaps suggest that the LF target words used in target analyses were of a moderate frequency when compared to all content words within the passage. The significance of target word analyses might have been more apparent if target words with even lower frequencies (that met the textual requirements for target word pairings) were available within the stimulus passage. Of note, one-third of content words within the passage were technically considered LF for global analyses; however, only LF words with specific textual characteristics led to appropriate matching with HF target words and, thus, were chosen as LF *target* words. Given the lack of established guidelines for HF and LF words, the results of this study might vary with different target words, different passages, and samples of students with varying compositions of reading skill.

A third limitation is the relatively highly skilled reading ability of the sample. This resulted in relatively few oral errors and might have affected the overall pattern of eye movements observed on target words. Hyönä and Olsen (1995), found differences in oral and silent reading measures between high and low skilled reading groups such that oral errors and silent total time measures occurred more often and lasted longer, respectively, in lower skilled readers than in higher skilled readers. Valle et al. (2013) found that intra-word regressions accounted for up to 75% of the variance for average-skill readers, as compared to high-skill readers, in total reading time on silent passages and oral passages, an indirect measure of decoding. Additionally, Valle et al. (2013) found other indicators of skill related differences in reading, such as the frequency of skipping words, number of fixations, pausing before words, and oral word duration. The sample of the current study might have led to decreased variation in student behavior, thus limiting the degree to which generalization might be made to other student populations. This limitation might be diminished when considering past findings that skill level with elementary students does not interact with other reading behaviors. Namely, gaze duration and number of regressions, did not seem to be affected by skill level in Hyönä and Olsen (1995). Additionally, Valle et al. (2013) and Huestegge, Radach, Corbic, and Heustegge (2009) found no evidence for an interaction between word frequency and elementary students' reading skills.

Application and Importance of Findings

Overall, findings from this study support the idea that word frequency is useful in predicting the difficulty of a passage at the word level and that difficulty might be determined by more than just the lowest and highest frequency words. Function words and high, medium, and low frequency content words are useful in determining the overall difficulty of a passage.

Given the extent to which CBM-R passages are used within schools to gauge student progress in reading, it is especially important to use passages that are tightly controlled for passage difficulty as represented by students WRCM. In practice, passages are often equated using methods that create standard errors of measurement much larger than expected reading growth (Ardoin & Christ, 2009). By using eye movement data to better identify characteristics of passages that lead to difficulty, it might be possible to create more equivalent passages resulting in reduced SEMs for WRCM.

Future research should establish guidelines for the categorization of high, medium, and low frequency words so that consistency across studies might occur. Research to determine if relative frequency or absolute frequency of the words within a passage contributes to the overall difficulty of a passage is needed. Additionally, equations for establishing passage difficulty might be created if the absolute frequency of words is found to be key to overall passage difficulty.

REFERENCES

Ardoin, S. P., Binder, K. S., Zawoyski, A. M., Foster, T. E., & Blevins, L. A. (2013). Using eye-tracking procedures to evaluate generalization effects: Practicing target words during repeated reading within versus across texts. *School Psychology Review, 42*, 477-495.

Ardoin, S. P., Williams, J. C., Christ, T. J., Klubnik, C., & Wellborn, C. (2010). Examining readability estimates' predictions of students' oral reading rate: Spache, Lexile, and Forcast. *School Psychology Review, 39*, 277-285.

Ardoin, S. P., & Christ, T. J. (2009). Curriculum-based measurement of oral reading: Standard errors associated with progress monitoring outcomes from DIBELS, AIMSweb, and an experimental passage set. *School Psychology Review, 38*, 266-283.

Blythe, H.I., Liversedge, S.P., Joseph, H.S.S.L., White, S.J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vision Research, 49,* 1583-1591.

Boersma, P., & Weenink, D. (2011). PRAAT: Doing phonetics by computer (Version 5.2.26).

Carroll, J.B., Davies, P., & Richman, B. (1971). *The American Heritage Word Frequency Book*. Boston: Houghton Mifflin Company.

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

Field, A. (2009). *Discovering statistics using SPSS* (3rd Ed.). Thousand Oaks, CA: SAGE Publications, Inc.

Goodman, Y. M. (1997). Reading diagnosis—qualitative or quantitative?. *Reading Teacher, 50,* 534.

Haenggi, D., & Perfetti, C. A. (1994). Processing components of college-level reading comprehension. *Discourse Processes*, 17, 83–104.

Huestegge, L., Radach, R., Corbic, D., Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision Research*, *49*(24), 2948-2959. doi:10.1016/j.visres.2009.09.012

Hyönä, J., & Olson, R. K. (1995). Eye movement patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1430–1440.

Joseph, H. L., Nation, K., & Liversedge, S. P. (2013). Using Eye Movements to Investigate Word Frequency Effects in Children's Sentence Reading. *School Psychology Review*, *42*, 207-222.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*, 329-354. doi:10.1037/0033-295X.87.4.329

Miller, J., & Schwanenflugel, P. J. (2006). Prosody of syntactically complex sentences in the oral reading of young children. *Journal of Educational Psychology*, *98*, 839-853. doi:10.1037/0022-0663.98.4.839

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. doi: 10.1037/1196-1961.49.2.151

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

Rayner, K. (2009). The 35th Sir Frederick Bartlett lecture: Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*, 1457-1506.

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.

Rayner, K., & Liversedge, S. P. (2011). Linguistic and cognitive influences on eye movements during reading. In S. P. Liversedge, I. D. Gilchrist, S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 751-766). New York, NY, US: Oxford University Press.

Rayner, K., & Slattery, T. J. (2009). Eye movements and moment-to-moment comprehension processes in reading. In R. K. Wagner, C. Schatschneider, C. Phythian-Sence (Eds.), *Beyond decoding: The behavioral and biological foundations of reading comprehension* (pp. 27-45). New York, NY US: Guilford Press.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. Psychonomic Bulletin & Review, *3*, 504-509

Reschly, A., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology*. doi: 10.1016/j.jsp.2009.07.001

Rogers, L. S. and Ardoin, S. P. (2014). Investigating eye-movement behavior during reading utilizing eye-tracking technology. (Doctoral Dissertation). University of Georgia, Athens, Georgia.

Schwanenflugel, P. J., Hamilton, A., Kuhn, M. R., Wisenbaker, J. M., & Stahl, S. A. (2004). Becoming a Fluent Reader: Reading Skill and Prosodic Features in the Oral Reading of Young Readers. *Journal of Educational Psychology*, *96*, 119-129.

Spache, G. (1953). A new readability formula for primary grade material. *Elementary English*, 53, 410–413.

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.