Investigating the Existence of Word Callers

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Rhonda Rae DeJournett
August 2017

Copyright © 2017 by Rhonda Rae DeJournett
All rights reserved.

Dedication

To Pierre, Danae, and Tre`, for the endless love, inspiration, support, and patience you have shown me throughout this arduous journey.

Acknowledgements

As I complete this long journey toward my Ph.D., I want to express my deepest appreciation to family, friends, and professors who encouraged me along the way. I cannot begin to name every person, but the most important to mention are my husband, Pierre, and children, Tre` and Danae, whom I asked to make the greatest life-changes. Thank you for being willing to move to Tennessee, for bravely and selflessly leaving behind your home, friends, school, and community as we faced a new adventure together. You are my everything. Thanks to my best friend Heather who was always one call away, and to my friends, Drs. Martha Meyer and Marlaine Chase, who convinced me I had a place in academia. Thank you, Dr. Meyer, and to my mentors, Dr. Kristin Rearden, and Nancy Stansberry, for completing the numerous references I requested as I searched for employment. I cannot begin to express my appreciation for sharing your valuable time. Thanks to my new-found friends, from among the amazing group of Ph.D.s and peers who collaborated, supported, imparting knowledge, and inspiring passion that made me eager to get to class. Thanks to the small group of UT friends who believe in the power of unicorns, and still believe we can make important differences in K-12 public, private, and charter schools, college classrooms, and in research. Thanks to extended family for helping to make our transition from Tennessee to the next great adventure. A special thanks to my sister, Renee, and sister-in-law, Belynda, for the unexpected surprises that always seemed to come when I needed them most. Finally, thanks to my dissertation committee for sticking with me to the end. I extend a special thanks to my committee. Amy Broemmel, for your patience, guidance, and support. Dr. Dennis Ciancio, for so patiently enduring what must have seemed like endless requests for meetings as my topic and research questions seemed to elude me. Dr. Sherry Mee Bell, thank you for being the amazing role-model and example of commitment to others, as well as for allowing me an opportunity to work with you; for providing good advice, and for making time for me in your tremendously busy schedule. I extend my gratitude to Dr.s Allington and McCallum for being a part of my journey. Finally, I want to extend the deepest appreciation to the administrative assistants in the TPTE department who unselfishly provided support and assistance as I maneuvered through this life-changing, challenging endeavor. Thank you, Teresa, Karen, Patti, Pat, and Vicki, and to all of the friends, family, and staff who contributed to my successful journey. I could never have done it without an incredible network of support.

Abstract

Numerous investigations of reading ability have identified students who demonstrate adequate oral reading fluency rates but fail to reflect understanding on measures of reading comprehension (Buly & Valencia, 2002; Catts, Compton, Tomblin, & Bridges, 2012; Hamilton & Shinn, 2003; Jenkins, Hudson, & Johnson, 2007; Jorm, 1983; Meisinger, Bradley, Schwanenflugel, Kuhn, & Morris, 2009; Morris, 1998; Torppa et al., 2007). This group of struggling readers is sometimes referred to as word callers; however, there has been debate among scholars (Hamilton & Shinn, 2003; Stanovich, 1993) regarding whether such a group exists, if the term is a misnomer (Spencer, Quinn, & Wagner, 2014), or if the label is too broad (Stanovich, 1993). Word callers are typically defined by their performance on two factors, reading fluency and reading comprehension. So far, researchers have mainly focused on data from common, universal assessments to determine the existence of word callers. In this investigation data-mining techniques were used to determine if word callers exist among a sample of first through fourth grade students attending a Tennessee school district. Identification was based on criteria requiring a reading fluency score that is a minimum of 14.04 Normal Curve Equivalent points greater than a student's reading comprehension score. A small number of word callers were identified using a single assessment, the STAR-Reading Assessment (Renaissance Learning, Inc., 2015), which includes both reading fluency and reading comprehension scores. Instructional implications of this research are important, as the earlier that word callers are identified, the easier their challenges are to remediate (Catts, 1997; Johnson, Jenkins, Petscher, & Catts, 2009; Snow, Burns, & Griffin, 1998; Torgesen, 2002).

Key words; word callers, poor reading comprehension, late-emerging reading disability.

Table of Contents

Chapter 1 Introduction 1
Purpose of the Study 3
Need for the Study 4
Delimitations 5
Definitions 5
Chapter 2 Literature Review 8
Reading Theory 8
The Simple View of Reading 9
Automaticity Theory 9
Schema Theory 10
Transactional Theory 11
Theoretical Influences on Instruction 12
Good Versus Poor Readers 13
Poor Comprehenders 14
Non-Verbal Learning Disability 15
Late-Emerging Reading Disability 16
English Language Learners 17
Word Callers 18
Reading Fluency of Word Callers 18
Quantifying Reading Comprehension Deficits of Word Callers 20
Assessing WCs 21
Comprehension Measures 22
Reading Fluency Measures 24
Curriculum Based Measures of Reading 26
Cloze and Maze Formats 27
Dynamic Indicators of Basic Early Literacy Skills (DIBELS) 28
Reliability and Validity of Fluency Measures 30
Summary 31
Chapter 3 Materials and Procedures 33
Theoretical Framework 33
Research Questions 34
School and Participant Demographics 35
Assessment Description - STAR Reading 36
Data Collection Procedures 38
Data Cleaning 39
Word Caller Variables 41
Chapter 4 Results 43
Analysis 43
WC Prevalence by Grade 46
Word Caller by Gender 47
Word Callers versus Non-Word Callers 48
Summary 49

Chapter 5 Conclusions and Recommendations 50
Limitations 52
Recommendations 52
Contribution to Research 53
References 56
Vita 79

List of Tables

Table 1 District Demographics 2014-15……..... 37
Table 2 STAR-Reading and Dynamic Indicators of Basic Skills Correlations 38
Table 3 STAR Reading Reliability Estimates ……. 39
Table 4 District Demographics Gender by Grade ……... 41
Table 5 Final School Count by Grade ……..... 42
Table 6 NCE Variable Descriptive Statistics ……43
Table 7 Variable Descriptives by Grade ……. 44
Table 8 Frequency and Range of Word Caller Scores by Normal Curve Equiv…. 45
Table 9 Word Caller Prevalence by Grade ……. 47
Table 10 Descriptives by NCE Variable, Gender …….. 47
Table 11 Word Callers by Grade, Gender ……. 48
Table 12 Word Caller versus Non-Word Caller Status by Gender …… 49

Chapter 1 Introduction

Numerous investigations of reading ability have identified students who demonstrate adequate oral reading fluency rates but fail to reflect understanding on measures of reading comprehension (Buly & Valencia, 2002; Catts, Compton, Tomblin, & Bridges, 2012; Hamilton & Shinn, 2003; Jenkins, Hudson, & Johnson, 2007; Jorm, 1983; Meisinger, Bradley, Schwanenflugel, Kuhn, & Morris, 2009; Morris, 1998; Torppa et al., 2007)). This group of struggling readers is sometimes referred to as word callers (WCs). Generally speaking, WCs possess adequate phonological awareness, word-decoding skills, and oral reading fluency (ORF) that is on grade level, but they do not comprehend well what they read. There has been some debate among scholars (Hamilton & Shinn, 2003; Stanovich, 1993) regarding whether such a group exists, if the term is a misnomer (Spencer, Quinn, & Wagner, 2014), or if the label is too broad (Stanovich). It is the opinion of some that word calling may be the result of assessment-driven classroom instruction, suggesting that teachers only teach what is tested (Catts, Petscher, Schatschneider, Bridges, & Mendoza, 2009; Pressley, Hilden, & Shankland, 2005; Samuels, 2007). This makes some sense considering the current educational landscape and the pressure on teachers and students to perform. Low-level specific skills, such as list reading, non-word decoding, and ORF are easily measured and tracked in terms of baselines and improvement. In these cases, component skills may often be taught and practiced as isolated skills. Some suggest that an over-simplified interpretation of reading theory, i.e., the simple view of reading (SVR; Gough, 1996; Hoover & Gough, 1990), has led to literacy instruction, interventions, and assessments that emphasize phonics,

decoding skills, vocabulary, and ORF conceptualized as simply reading quickly and accurately (Hamilton & Shinn, 2003; Jenkins, et al., 2007; Nation & Snowling, 2000) without consideration or inclusion of the comprehension component. To illustrate, a common definition of ORF measures speed and accuracy and is typically reported as words read correctly per minute, or wcpm (Eldredge, 2005; Fuchs, Fuchs, Hosp, & Jenkins, 2001; Silverman, Speece, Harring, & Ritchey, 2013; Wolf & Katzir-Cohen, 2001). As such, common early elementary practices often involve practicing reading strategies e.g., repeated reading, choral reading, that promote ORF and improve reading comprehension (RC). These two components are both typically measured in early grades and specified in the definition of word calling. Online assessment software, such as STAR Reading (STAR-R; Renaissance Learning, Inc., 2015), do not calculate an estimated ORF for students beyond fourth grade, which would be helpful in isolating WCs in later grades, if they exist.

Research varies in regard to the types of struggling readers identified, ranging from three to 17 subgroups, from kindergarten to grade seven (Buly & Valencia, 2002; Cain & Oakhill, 2006; Leach, Scarborough, Rescorla, 2003; Meisinger, Bloom, Hynd, 2010; Morris, 1998; Torppa et al., 2007). Characteristics of each group differed on various measures, such as word-level skills, e.g., reading, spelling, decoding, vocabulary, and RC. Buly and Valencia described groups of readers fitting the WC definition as automatic WCs, struggling WCs, and word stumblers, reflecting the ORF component, while Leach and colleagues (2003) referred to similarly performing students as having a late-emerging reading disability.

The component skills used to identify WCs simply focus on a reading fluency (RF) measure and RC, information which should be available to classroom teachers for each student. Reading fluency measures may be oral or silent (SRF). The RF measures obtained from common, online, universal assessments, such as the STAR-Reading assessment, are usually an estimated oral reading fluency measure (RF-E). Routine classroom assessments and universal screeners should provide enough information to identify WCs. The identification of WCs, should they exist, is important as their skill deficits differ from other struggling readers and, thus, require instruction tailored to their needs if they are to become good readers.

Purpose of the Study

Teachers and researchers have identified groups of students who perform at various levels of competency in terms of reading ability. However, there is disagreement among scholars regarding the existence of the group known as WCs. The two skills consistently addressed in WC research are an RF measure, ORF, SRF, or RF-E, and RC. The first factor, RF, is considered adequate if students demonstrate average or above-average oral reading ability on grade level text. Typically, the RC factor is considered low, measured by standardized assessments, if scores are at least one standard deviation below the mean (Buly & Valencia, 2002; Dymock, 1993; Meisinger, et al, 2010; Morris, 1998), or if the standard score ≤ 85 (Leach, et al, 2003). The purpose of this study is to determine if WCs exist based on the two measures, RF and RC.

Need for the Study

As previously mentioned, research on WCs has included a variety of component skills. However, the different components, and the way they are measured seem to confuse the interpretation of results. Research clearly isolating RC and RF components is needed to determine if there are WCs, as defined, among student populations. The investigation regarding WCs is also important beyond simply establishing their existence. The prevalence of WCs may increase with each grade level, and some may not be identified until around the fourth grade, possibly as a result of exposure to increasingly complex vocabulary, text, and concepts (Chall & Jacobs, 2003), or because of deficits in specific cognitive ability (Catts, et al., 2012). As schools and teachers face continued pressure to improve all students' performance on high stakes testing, the instructional implications of this research are important, as the earlier that WCs are identified, the easier their challenges are to remediate (Catts, 1997; Johnson, Jenkins, Petscher, & Catts, 2009; Snow, Burns, & Griffin, 1998; Torgesen, 2002). So far, no researcher has used data from school- or classroom-based assessments, exclusively, to identify WCs. So, the question remains as to whether classroom-based universal assessments and standardized state exams can be used to identify this group of students.

My research goal is to answer the following questions:

1. What percentage of the students can be identified as WCs based on the definition of word callers where RF-E score is equal to grade level and RF-E score is greater than the RC score by at least 14.04 NCE points?
2. Does the percentage rate of WCs increase across grade levels?

3. Are there differences in number of WCs vs. non-WCs in relation to the following variables:
    a. Gender
    b. Grade level
    c. Free and reduced lunch status
    d. Identified as having a disability

Delimitations

The main delimitation of this project relates to the selection of the research population. I selected the participating district based on ease of access. The University of Tennessee, and my department specifically, has an ongoing, collaborative relationship with the school district, making access to students and data easier than trying to establish trust and collaboration with a new school or district. This choice impacts the generalizability of any results to other districts with different demographics. The selection of school district also dictated the universal assessment. The participating district used STAR Reading assessment in all elementary schools.

Definitions

Automaticity theory-when a skill or subskill can be performed while attention is directed elsewhere. Specific to reading, when decoding and word-reading are automatic, attention can be given to the process of comprehension (LaBerge & Samuels, 1974).

English Language Learners (ELL)-children who speak a language other than English at home and who are learning English as their second language (Kena et al., 2016). There

have been changes in terminology in reference to students who speak a language other than English at home, but will be referred to here as English language learners (ELL) for consistency.

Late-emerging reading disability (LERD)-students who seem to get off to strong start in terms of reading ability but experience reading difficulties in later grades (Catts, et al, 2012).

Oral reading fluency (ORF)-measure of rate and accuracy in terms of words read correctly per minute (Hasbrouck & Tindal, 1992).

Prosody-intonation, loudness, and timing used when reading aloud (Dowhower, 1991); a measure of timing, phrasing, and intonation when reading aloud (Kuhn et al., 2010)

Silent Reading Fluency (SRF)- a measure of students' reading comprehension determined by having students read a passage silently and circling the last word read when told to stop two minutes. Students answer standardized comprehension questions following silent reading (Fuchs, Fuchs, Eaton, & Hamlett, 2000, in Fuchs, et al., 2001)

Response to Intervention (RtI) - is a multi-tier approach to the early identification and support of students with learning and behavior needs. The RtI process begins with high-quality instruction and periodic screening of academic progress of all children in the general education classroom. The process allows educators to determine if interventions are effective in improving students' performance in targeted area(s) of intervention (Fuchs & Fuchs, 2006).

Word caller (WC)-children who are typically identified in late elementary school who demonstrate on grade level oral reading fluency (ORF) SS=85 or above, with RC that is 14.04NCEs below expected performance based on RC measures (Meisinger, et al, 2009) Universal screening- typically brief, conducted with all students at a grade level, and followed by additional testing or short-term progress monitoring to corroborate students' risk status ("Universal Screening | Center on Response to Intervention," n.d.).

Chapter 2 Literature Review

In order to determine the existence of the WC, it is important to understand some of the reading theories that have influenced instruction and assessment in today's classrooms. Characteristics of good readers will segue to a discussion of WC characteristics and students who demonstrate similar reading profiles. Common assessment tools for fluency and RC will be reviewed, but with particular emphasis on clarifying a definition of fluency and how it is measured. The chapter will end with a review of some assessments commonly used as they relate to investigating the possible existence of WCs.

Reading Theory

Word callers are defined in terms of adequate ORF and RC deficits. Stanovich (1993) was critical of research describing WCs, specifically as automaticity theory relates to ORF and RC in that deficits in these areas may be influenced by factors such as exposure to print, reading experience, and listening vocabulary. Hamilton and Shinn (2003) investigated the existence of WCs comparing teacher judgments of student reading ability to simple curriculum-based oral fluency measures, i.e. rate and error, and comprehension. Researchers found little support for the existence of teacher-identified WCs among participating third grade students. Specifically, teachers overestimated ORF scores for fluent readers and WCs. Furthermore, the teacher-identified WCs' ORF and RC scores were lower than their more fluently reading peers. Current reading theory emphasizing ORF and RC measures may influence the defining characteristics and, perhaps, the identification of WCs. A review of influential theory is necessary to

understand the possible framework for understanding how WCs have been defined and identified.

The Simple View of Reading

The simple view of reading, SVR, (Hoover & Gough, 1990) contends that reading ability is based on decoding and linguistic comprehension. The SVR emphasizes bottom-up phonics, decoding, and ORF instruction, and is widely supported in the literature (Gough & Tunmer, 1986; Hoover & Gough, 1990; LaBerge & Samuels, 1974; National Reading Panel, 2000; Perfetti et al., 1987; Scarborough, 1998; Snow et al., 1998; Wise, et al., 2010). Hoover and Gough did not discount the complexity of reading processes, but suggested that becoming a good reader is a cumulative process of skill acquisition and practice. The report by the National Reading Panel (2000) has been used to support the development of curriculum, instruction, and assessment focused on phonics, decoding, and reading fluency, as these skills are the easiest to improve and assess. Furthermore, they asserted decoding and linguistic comprehension contribute mutually to good reading, and that good reading will not develop in absence of one or the other skill.

Automaticity Theory

Automaticity theory (LaBerge & Samuels, 1974) is closely aligned to the SVR, and emphasizes the importance of developing readers' decoding skills for unknown words, and automatic word recognition to promote reading fluency. In theory, when readers are able to read accurately and fluently, their cognitive energy and attention can be used to understand the text. Automaticity theory (LaBerge & Samuels; Samuels,

1976) is well-grounded and supported in research of skills, such as phonological awareness, decoding, and reading fluency, as they relate to RC (Anderson, 1981; Bashir & Hook, 2009; Burns et al., 2011; Fuchs, et al, 2001; Nation, Snowling, & Clarke, 2007; Perfetti, et al., 1987; Reutzel & Hollingsworth, 1993; Therrien, 2004; Vellutino, 1991) The theory of automaticity (LaBerge & Samuels, 1974) aligns with a behaviorist approach which theorizes that frequent and repeated practice is needed to teach new skills. The theory is that with practice, when interpreting graphemes, e.g., phonemic and phonological units, is immediate or automatic, decoding will become automatic or fluent, then cognitive attention can be focused on comprehending what is read, rather than on word-level decoding processes (LaBerge & Samuels; Perfetti & Hogaboam, 1975). Automaticity theory also emphasizes repeated practice in teaching of low-level phonics skills for reading, which allow the reader to effortlessly interpret words on a page, and to simultaneously comprehend.

Schema Theory

Schema theory is widely supported and explains how a student's culture and experiences, or schemata (Anderson & Pearson, 1984), relate to functions of selective attention and inference-making, i.e., promoting attention, self-editing to determine what is or is not relevant, and inferential reconstruction. Schema theory is aligned with more constructivist views of teaching and learning in which teachers facilitate investigating and problem solving in ways that incorporate social experiences and knowledge using activities and projects to create new knowledge (Dewey, 2004). Anderson (1981) suggests teachers include instructional practices such as activating prior knowledge,

building background knowledge and incorporating activities that help children relate content material to their own experiences. Ruddell and Unrau (2004) extended these suggestions, adding value to promoting reader motivation, setting purpose for reading, helping students understand that reading and constructing meaning is purposeful and interactive, and that sharing authority and negotiating meaning allows students to validate and verify their own understanding.

Transactional Theory

Related to schema theory and constructivism, transactional theory (Rosenblatt, 1993) incorporates prior learning, background knowledge, and experiences to relate to new learning. Transactional practice incorporates instructional approaches using a variety of teaching methods that encourage students to interact with texts to comprehend what they read and write. Important components of transactional instruction include teachers modeling their thinking when they read and write, having students connect to their own background knowledge of a topic or book, making connections to other texts, and writing about what they read. Teachers who aspire to incorporate practices aligned with transactional theory include classroom and peer discussions to help students relate to a text. They guide students to activate their own background knowledge as a springboard for understanding characters and situations; they promote reader motivation by giving them choices of reading materials; they explain the purpose for reading so students understand the goal, leading students to understand that reading and constructing meaning is purposeful and interactive.

Theoretical Influences on Instruction

Understanding the main components of these influential theories is important as they play a role in textbook publications, in what teachers are expected to teach if students are to become good readers, and in the choice of skills deemed important enough to assess. Lipson and Wixson (1986) predicted that limiting measures of reading to a narrow set of skills would result in the narrowing of instruction. In the decades following Lipson and Wixson, other scholars weighed in supporting the negative impact of assessment practices on classroom instruction (Airasian & Madaus, 1983; Barone, 2013; Falk-Ross, Szabo, Sampson, & Foote, 2009; Madaus, 1988). Madaus and Russell (2010), for instance, pointed out a long history of political influence on classroom instruction via rewards or sanctions, which, still today continues to influence what children should know and understand. Furthermore, they assert that testing "shapes important educational values (p. 22)" as choice of subjects tested suggests greater value for some coursework above others. These ideas could be interpreted and compared to Campbell's Law, which argues that established norms influence the value on specific skills, and may negatively impact those who do not fit within accepted parameters (Campbell, 1976).

Common Core State Standards emphasize reading instruction that is focused on reading comprehension. Current instruction and assessment practices used in many RtI (Response to Intervention) programs, however, continue to emphasize low-level skills, like decoding, and oral reading fluency. Many assessments use multiple-choice formats selected for their efficiency in terms of time for administration, scoring and reporting. Results may overlook students' backgrounds and cognitive processes that might explain

differences in what good readers and struggling readers do. This might suggest that current testing practices overlook thinking processes of children who are not raised in the dominant culture, resulting in over-identification of struggling learners. Madaus and Russell (2010) further suggest that common high-stakes assessments narrow the content and skills taught, placing more value on instructional time spent teaching to the test and on learning how to answer the type of questions asked, rather than teaching students to be critical thinkers. What, then, do good readers do?

Good Versus Poor Readers

People read for a multitude of reasons, but whether one reads for pleasure or for information, the end goal should be to comprehend. Good readers (GR) know this, and develop strategies to use before, during, and after reading (Pressley & Gaskins, 2006). In their qualitative analysis of verbal protocols during reading, Pressley and Afflerbach (2009) reported that prior to reading, GRs may preview a book or other text, to judge whether they want to read the book. Good readers look at illustrations, read headings or chapter titles, and judge length, to determine if they are interested. While they are reading, GRs may look up new words, connect to background knowledge, take notes, visualize, slow down, ask questions of themselves, or reread confusing portions of the text (Pressley & Afflerbach,; Pressley & Gaskins; Therrien, Wickstrom, & Jones, 2006). After they read, GRs are more likely to think about what they just read, make connections to other ideas and concepts, returning to the text to insure adequate understanding (Pressley & Gaskins).

Poor readers may use some of these same strategies that good readers do; however, they do not use them effectively. Poor readers may additionally be bogged down by decoding processes, which may cause a bottleneck and prevent or inhibit comprehension (Perfetti & Lesgold, 1977). Because they are likely to be deficient in these skills, poor readers are likely to read less than GRs (Allington, 1977; Cain & Oakhill, 2011; Stanovich, 1986), resulting in obvious decreases in print exposure, reading experience, and perpetuating the continued challenges in reading, i.e., the Matthew effect (Cain & Oakhill; Stanovich, 1986). The Matthew effect is the concept that the rich get richer and the poor get poorer. In terms of reading development, this means the more a person reads, the better they get at reading, and the converse would also be the case. This does not claim to explain the existence of students who have disabilities specifically related to RC, but there may be some overlap with poor readers who have RC problems related to lack of experience or exposure with reading.

Poor Comprehenders

Students with deficits in specific reading skills may be identified by the more general term, Specific Learning Disability (SLD). For example, a student previously labeled as dyslexic, who has poor ORF but good listening comprehension, may be identified as having an SLD in RC. A student who reads fluently but without comprehension may also be identified as having an SLD in RC. A factor that may further confound the issue of WCs is the use of the other terms, "nonverbal learning disability (NVLD)," (Cornoldi, Vecchia, & Tressoldi, 1995; Harnadek & Rourke, 1994; Mammarella & Pazzaglia, 2010; Rourke, 1995) and "late-emerging reading disability

(LERD)" (Compton, Fuchs, D., Fuchs, L. S., Elleman, & Gilbert. (2008); Leach et al., 2003) from the fields of special education and neuropsychology respectively. The three groups of children exhibit very similar reading skills, and virtually identical academic performance-profiles. However, the identification of subgroups in some literature suggests differences that due to potential language, reading, and experiential deficits common among English language learners, referred to here as ELL (Buly & Valencia, 2002; Catts et al., 2012; Meisinger, et al, 2009; Torppa et al., 2007). These potential, non-academic factors are not routinely measured by standard assessments used in schools, but this information may be useful for determining students' instructional needs. Regardless, some discussion regarding the similarities and differences in RF and RC abilities of these groups is merited.

Non-Verbal Learning Disability

Rourke (1995), a neuropsychologist, described students with good phonological skills and low reasoning ability, diagnosing them as NVLD. Children with NVLD were described as having strong abilities in phonemic awareness, real-word decoding, spelling, and writing, but spelling and writing ability was best in mid elementary years after decoding and phonics became rote. In addition, they presented with perceptual and psychomotor difficulties, with clear preferences and better motor ability on the right side of the body. These children had severe deficits in visuo-spatial organization, pronounced deficits in nonverbal reasoning and processing, and highly developed rote verbal-memory abilities. The children were also characterized as very talkative, yet deficient in the understanding of psycholinguistic pragmatics, social cues and social interactions. Finally,

they also struggled with visuo-spatial organization and higher order cognitive processes, such as problem-solving. Rourke's concern was that these students might be overlooked until the demand for cognitive and reading skills increased in later elementary school. Children with NVLD had some difficulty with nonsense word decoding. Rourke's (1989) early estimations of 5-10% among children receiving special education services, was challenged by later findings of Denckla (1979), who estimated prevalence from 0.1 to 1% of the children with learning disabilities. The children he described were referred for evaluation. These are the children whose deficits were noticed. Based on the similar descriptions of the NVLD and WCs, it is possible Rourke's estimate is low and that some children are never identified. A search for other estimates provided only Cornoldi's (1999) estimated prevalence rate of 2.5%.

Late-Emerging Reading Disability

In a longitudinal investigation of reading achievement of kindergarten students, Judge and Bell (2010) identified 3.6% of students who received special education services as fitting the late-emerging learning disability criteria. Dennis (2013) identified 32% of students identified as having LERD demonstrated good low-level skills and low comprehension skills. In a large study (n=493) involving multiple grade levels (K, 2nd, 4th, 8th, 10th), Catts, et al, (2012) referred to students as having LERD, who exhibited profiles similar to WCs, i.e., strong word-level skills paired with weak comprehension. The group did not identify LERDs in kindergarten; however, students were identified beginning in second grade. Out of 13.4% of poor readers identified in this research population, more than half (52%) were identified as having only comprehension deficits.

Prevalence rates increased from second grade through eighth grade at which point rates appear to stabilize. The greatest increase was between second and fourth grade. Researchers also found that students with LERD specific to comprehension difficulties had a history of nonverbal cognitive deficits and oral language impairments. Over half of the LERDs identified in their research struggled with comprehension and not decoding, and most students were identified in fourth grade. Other research supports the notion that the prevalence of WCs increases with grade level, as the deficits become more apparent (Compton, et al., 2008; Knight-Teague, Vanderwood, & Knight, 2014; Leach et al., 2003). Leach and colleagues (2003) reported a significantly larger percent of students at a middle school who were identified as automatic WCs (n=23, 24%), mostly students who are learning English as a second language. Also, it is at 4th grade where comp becomes the primary way teachers measure reading performances.

English Language Learners

Buly and Valencia (2002) found that 60% of students identified as automatic WCs were ELL students who were not receiving services. Another study including ELL students (Stothard & Hulme, 1992) found that among their seven and eight-year-old ELL participants, 10-15% had adequate decoding skills but struggled to understand what they read. A more recent investigation (Knight-Teague et al., 2014) with a small sample (n=26) of Spanish-speaking third and fifth grade children identified as ELL identified 6% of these third grade and 8% of fifth grade students as WCs.

Word Callers

One of the challenges of investigating WCs is that they are often overlooked or unidentified until late elementary years, which may explain the use of the term late-emerging reading disability (Leach, et al., 2003). Performance on RC measures is inconsistent, depending on the methodological approach, the cognitive requirements involved, and the format of the assessment used. Researchers have used quantitative and mixed-methods to investigate the prevalence of reading subgroups (Meisinger, et al, 2010; Meisinger et al., 2009). Hamilton and Shinn (2003) used a mixed methods approach to determine if third grade teachers' ratings or CBM-R (curriculum-based ORF measures) would more accurately identify WCs in their classrooms. Their results suggested that CBM-R was more accurate and that teachers overestimated reading fluency and RC abilities of their students, similar to others (Feinberg & Shapiro, 2009; Hamilton & Shinn; Meisinger, et al., 2010).

Reading Fluency of Word Callers

As previously discussed, reading theory recognizes the relationship between reading fluency and RC. Investigations of RC often include measures of fluency and provide convincing support for what seems obvious; we cannot comprehend what we cannot read well (Alber-Morgan, Ramp, Anderson, & Martin, 2007; Begeny & Martens, 2006; Daly, Bonfiglio, Mattson, Persampieri, & Foreman-Yates, 2006; O'Connor, White, & Lee Swanson, 2007; Therrien, 2004; Therrien et al., 2006). Researchers (Allington, 1983; Holliman et al., 2014; Klauda & Guthrie, 2008; Pinnell, 1995; Schwanenflugel, Westmoreland, & Benjamin, 2015; Veenendaal, Groen, & Verhoeven, 2015) argue that

features such as prosody and intonation should be included in any measure of ORF, but the most consistent components measured are speed and accuracy, or wcpm, which is a measure of the number of words read correctly per minute. What must be acknowledged, is that fast reading does not guarantee good RC (Danks & Fears, 1976; Fleisher, Jenkins, & Pany, 1979).

Numerous studies support the correlation between RF and RC (Armbruster, 2003; Kim, Wagner, & Lopez, 2012; NRP, 2000; Price, Meisinger, Louwerse, & D'Mello, 2016; Snow, et al., 1998). In a study of sixth grade students, Fuchs, et al. (2001) reported a validity coefficient of .91 on one-minute RF to RC. However, their experiment did not include measures of expression or prosody, which may contribute to RC (Allington & Brown, 1979; Dowhower, 1991; Klauda & Guthrie, 2008). Silberglitt and colleagues (2006) noted the rate of growth in reading rate decelerates in later grades and may be less important as ORF begins to plateau from fifth-grade and beyond. Others found no correlation between reading speed and RC defined in terms of wcpm (Cramer & Rosenfield, 2008; Pressley, et al., 2005). Cramer and Rosenfeld evaluated the relationship between urban fourth graders' ORF and RC but reported no significant correlation (2008).

Assessments, such as STAR-R, use the relation between ORF and RC to estimate an SRF, despite the conflicting results of ORF-RC investigations. Improvements in students' ORF rates may not guarantee increases in SRF (Freeland, Skinner, Jackson, McDaniel, & Smith, 2000). Estimated silent reading fluency (RF-E) has replaced

individual ORF measures as a direct result of the format necessitated by current online assessment practices.

Quantifying Reading Comprehension Deficits of Word Callers

Because WCs are defined in terms of RF and RC, it is important to examine how these skills are measured. Reading fluency is not always assessed, particularly for students who are not showing signs of RC or RF difficulties. However, when students are assessed, teachers and researchers generally define adequate ORF in terms of wcpm, using grade level cut scores, i.e. Hasbrouck and Tindal's National Oral Reading Fluency Norms (1992), percentiles, and various standardized scores. Computer-based assessments, such as STAR Reading, include similar tools for norming Estimated ORF scores. Adequate ORF ranges require a minimum SS $\geq$ 85 when reading grade level text (Meisinger et al., 2009).

Researchers have used various measures to quantify or define the comprehension skills that differentiate WCs from other poor readers. Buly and Valencia (2002), for instance, found automatic WCs among fourth grade students scoring 374 – 378 out of 700 points (53 – 54%) based on the Washington Assessment of Student Learning (WASL), the state assessment which focuses on comprehension, rather than fluency. Investigators also assessed students' RC with the Qualitative Reading Inventory-Second Edition, with open-ended response design, which reflected RC levels one to two years below grade level on narrative and expository passages. In this investigation, researchers also used the single word decoding subtest from the Woodcock Johnson-Revised, and vocabulary, measured by Peabody Picture Vocabulary Test-Revised. These assessments

were inconsistently related to the WASL on the skills they measure. More than half (60%) of WCs in the sample were identified as second language learners. In other investigations, researchers used a cut score ranking and percentiles where 25% is the lowest score in the average range (Barth et al., 2008; Etmanskie, Partanen, & Siegel, 2016; Fletcher, et al., 2014). Using this approach, WCs were identified if ORF performance was ≥ 35th percentile and RC was ≤ 25th percentile. A central problem with such definitions is the standard error of measurement.

Assessing WCs

Choice of assessments is an important factor in identifying WCs. Reading comprehension and RF are the principle characteristics used to define WCs. Thus, how these components are evaluated is important. Per the National Reading Panel (2000), elementary students and readers who struggle should receive reading instruction that includes phonemic awareness, phonics, fluency, vocabulary, and RC. Researchers investigated time spent on RC instruction (Duke, 2000; Durkin, 1979; Fisher & Berliner, 1985) and conducted evaluations regarding the effectiveness of specific reading instruction and interventions (Block, Parris, Reed, Whiteley, & Cleveland, 2009; Guthrie et al., 2004; Homan, Klesius, & Hite, 1993; McBride, 2005; Popplewell & Doty, 2001; Therrien, 2004). All have suggested areas for improvement regarding what skills should be measured and how those skills should be measured. Researchers advise including multiple measures to obtain an accurate interpretation of the targeted skills measured. Computer-based RtI intervention software provides assessments and generate individual student performance reports designed to identify areas of weakness. A review of some

common assessments used in schools may be helpful, particularly as they relate to specific RC and ORF measures.

Comprehension Measures

The RAND report (Snow, 2002) defined RC in terms of sociocultural context, in that the reader interacts with the text, the activity, and their experiences to create meaning when they read. The report, similar in content to the NRP report (2000), acknowledged that fluency alone does not equate to comprehension, and further recognized the contributions of factors such as vocabulary, oral language skills, higher-order thinking skills, such as problem-solving, analyzing, and visualizing, motivation, purpose, background knowledge, discourse knowledge, self-monitoring, and self-efficacy. Accurate measurement of the component skill development is difficult as not all are externally observable, and the reliability of self-reported information is questionable.

Investigations related to teaching component skills related to RC present positive findings on the efficacy of teaching specific skills and strategies to students (Block, et al., 2009; Eason, Goldberg, Geist, & Cutting, 2012; Guszak, 1967; National Reading Panel, 2000; Pressley, Johnson, Symons, McGoldrick, & Kurita, 1989; Valencia & Buly, 2004). Teachers use instructional strategies such as questioning, pre-teaching vocabulary, building background knowledge, modeling self-questioning, and by thinking aloud, to help students to improve RC.

Researchers have posited several concerns regarding assessments used for evaluating RC (Cutting & Scarborough, 2006; Keenan & Betjemann, 2006; Keenan, Betjemann, & Olson, 2008). Different testing formats used for measuring RC may also

require different skills, e.g., silent RC, listening comprehension, and answering implicit, and explicit questions. When RC is assessed in the classroom, multiple-choice, true-false, and open-ended questions formats are often used to evaluate recall and background knowledge, rather than higher order cognitive reasoning, such as inference-making, problem-solving, and evaluating. It is important to evaluate potential assessments for research in terms of format to ensure the data gathered will provide the information needed to answer the research question. For example, a single vocabulary assessment may not be the best choice for predicting students' RC ability. Ideally, a vocabulary assessment would be used in combination with other RC assessments to evaluate potential factors that may contribute to RC ability.

Skills related to RC, such as language ability (Cain, Oakhill, & Bryant, 2004; Cain, Oakhill, Barnes, & Bryant, 2001; Nation, Cocksey, Taylor, & Bishop, 2010; Nation et al., 2007), verbal memory (Carretti, Cornoldi, De Beni, & Romanò, 2005; Palladino, Cornoldi, Beni, & Pazzaglia, 2001; Pimperton & Nation, 2010), and attention (Ghelani, Sidhu, Jain, & Tannock, 2004; McInnes, Humphries, & Hogg-Johnson, 2003), are not typically assessed in the classroom. High stakes state tests, such as the Tennessee's state exam, typically assess students' vocabulary and RC via multiple-choice formats, which are easily scored. Informal RC assessments include variations of cloze assessments which omit every nth word and give three to five choices for an appropriate word choice. Other formats include true/false tasks; sentence verification tasks, which are a variation of true/false assessments; and open-ended questions, which are more subjective in terms of scoring responses. All but the open-ended questions are good for group testing in

terms of ease administration and scoring, but there are issues of test length, text interactions, student difficulties in formulating responses, and lack of standardization that pose potential problems when using them for evaluating RC. Specific examples of these formats include CBM-R (Deno, 1985), and editions of the Qualitative Reading Inventory (Leslie & Caldwell, 1995, 2010).

Reading Fluency Measures

Aligning with automaticity theory, research confirms a correlation between RF, RC, and state assessments (Homan, et al., 1993; Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2008; Silberglitt, et al., 2006). Therefore, it is important that RF and RC be assessed together (Snow et al., 1998). As discussed previously, there is support for including prosodic measures when assessing ORF. However, reading prosody is not typically assessed in classrooms, and current online assessment practices do not account for prosody. The two measures of RF, ORF and Silent Reading Fluency (SRF), can be used as the RF metric for determining the existence of WCs. As defined in the definitions in Chapter 1, SRF is a measure of students' reading comprehension often determined by having students read a passage silently and circling the last word read when told to stop after two minutes. Students may be required to answer standardized comprehension questions following silent reading (Fuchs, et al., 2000, unpublished, as cited by Fuchs, et al., 2001). Although individually administered ORF assessments provide important information to teachers and researchers, one downfall is that they are time-consuming to administer to large groups of students. Beyond elementary school, students' reading activities are usually silent. Thus, researchers have investigated the relationship between

SRF and RC. Online assessments enable teachers to assess groups of students simultaneously, therefore making SRF assessment efficient and practical in terms of time. The main limitations of online SRF assessments is that there is no way to assure that a student has read the required text. For individually administered, timed SRF assessments, students may not accurately identify where they ended their reading. Another issue is the potential difference between a student's comprehension of texts when reading aloud versus reading silently, as ORF may not reflect SRF ability (Freeland, et al., 2000).

Another assessment available to schools is AIMSWeb, which is often a component of their Response to Intervention (RtI) program. AIMSWeb is used as a universal classroom screening and to monitor student progress during RtI; it includes CBM-Rs to assess RF, and maze assessments for RC skills. AIMSWeb includes numerous CBM-R passages (23 to 33 for each grade level from first to eighth grade) for evaluating RF performance measured by wcpm. The availability of numerous passages per grade level means it can be used frequently throughout the school year for progress-monitoring. Other standardized assessments for ORF, such as the Diagnostic Assessment Battery and the Gray Oral Reading Test, and SRF assessments, such as Woodcock–Johnson IV-Reading Fluency subtest (WJ-IV; Woodcock, Schrank, McGrew, & Mather, 2014) the Test of Silent Contextual Reading Fluency–Second Edition (TOSCRF-2; Hammill, Wiederholt, & Allen, 2012) and Test of Silent Word Reading Fluency (TOSWRF-2; Mather, Hammill, Allen & Roberts, 2014), include only two to four versions. Therefore, these assessments are more useful as part of a diagnostic battery, or as pre- and post-intervention assessments. Common online assessments, such as STAR-

R, provide multiple texts and are used for universal screening and progress monitoring which calculate RF-E for grades 1-12. Examples of other online assessments are i-Ready Diagnostic for Reading / English Language Arts for grades 1-8, and Scholastic Reading Inventory for grades K-12. Online assessments typically assess component skills i.e., phonological awareness, phonics, decoding, vocabulary, and RC for literature and informational text using a multiple-choice format. These assessments typically provide scaled scores, instructional placement levels, normed scores, lexile levels, and suggested areas of instruction.

Curriculum Based Measures of Reading

Curriculum based measures of reading, or CBM-R, were designed to evaluate general reading proficiency, measured as words read correctly per minute (wcpm), and to monitor student progress (Deno, 1985). Although the idea for CBM-R was that teachers could construct their own assessments using classroom materials, it has become somewhat synonymous with a standardized test based on national norms. The CBM-R design requires students to read aloud and to select words deleted from the text (Deno, 2003). Hasbrouck and Tindal (1992) extended Deno's research by establishing grade leveled RF norms so student scores could be compared. Silberglitt and colleagues (2006) reported the correlation between CBM-R and State assessment score for third and eighth graders was .71 and .51 respectively. The decreased correlation from third to eighth grade may be explained by differences in text difficulty, increased linguistic and cognitive demands, and the developmental stage of reading required by a student to read early elementary material compared to fourth grade material and beyond (Chall, 1983). Chall

explains this developmental change from lower to higher elementary grades a transition from learning to read, to reading to learn. As text becomes more difficult, the relation between ORF and RC may decrease as students slow down or reread passages to understand (Chall; Spear-Swerling, 2004).

There is some debate on the use of CBM-R for identifying struggling students, as it is only a measure of ORF and not a direct measure of RC (Fuchs & Fuchs, 1992). Based on ORF scores WCs would miss being identified for interventions. Measures of ORF are more sensitive for younger students and those needing support services from special education (Baker et al., 2015). Recent evaluation on the use of CBM-ORF measures to evaluate middle school students (not receiving services in special education) suggested it does not provide enough information for teachers to determine if students were improving or struggling (Baker et al.). Baker and colleagues reported that ORF is less predictive of reading achievement as students entere higher grade levels.

Cloze and Maze Formats

Cloze and maze provide measures of fluency and RC. Cloze and timed maze formats are similar in that a word is removed from a text at specific intervals. Both have the same goal, which is to determine RC by asking the student to read and construct, or reconstruct, text. The student may be required to generate an acceptable word, provide a specific word (cloze), to select the appropriate word given three choices (maze) that reflects comprehension; maze assessments are timed. Research suggests a correlation between maze and standardized tests of .82 and a test-retest reliability of .90 (Guthrie, Seifert, Burnham, & Caplan, 1974). The Woodcock-Johnson family of assessments uses

cloze format for passage RC. Although teachers may be giving up ease of scoring in using cloze or maze assessments, the results may provide timely, teacher-friendly information that is more pertinent to classroom instruction than data retrieved from standardized assessments. Klein-Braley (1997) investigated several measures of language proficiency related to reading comprehension, comparing five different versions of cloze test formats to determine which had the greatest validity and reliability. She also concluded that the C-test, which provides the first two or three letters of missing words in a passage, has the greater reliability, with rtt ($_{Alpha}$) estimated at .85 compared others, such as typical multiple-choice cloze formats, estimated at .55.

Dynamic Indicators of Basic Early Literacy Skills (DIBELS)

The development of DIBELS (Good & Kaminski, 2002) reflects aspects of automaticity theory. It was designed to measure the fluency of isolated skills emphasizing phonemic and phonological awareness, letter naming, nonsense word decoding, and ORF. The use of DIBELS is common, but some researchers (Catts et al., 2009; Pressley, et al., 2005; Samuels, 2007) argue against the use of DIBELS due to ceiling and floor effects, and the notion that DIBELS testing leads to word-calling due to its emphasis on fast reading and responses. Good, Simmons, and Kame`enui (2001) reported predictive validities of DIBELS subtests ranged from .34 to .82 for state exams, thereby concluding there was some utility of for assessing foundational skills. Good and colleagues also reported more than 90% of students who met DIBELS benchmark goals also met or exceeded expectations on the state assessment.

Abbott, Wills, Miller, and Kaufman (2012) analyzed the correlation between DIBELS correct words per minute (CWPM) and passage comprehension on the Woodcock Reading Mastery Test -Revised (WRMT; Woodcock, 1998). The group found strong correlations between WRMT and DIBELS CWPM (.68 to .73) and error rates (-.72 to -.74) for second and third grade students needing intensive reading instruction, and moderate correlations for second and third grade students who met benchmarks. Other research including DIBELS measures of CWPM as a predictor of performance on state exams yielded correlations of .70 (Roehrig, et al, 2008) and .67 (Good et al., 2001), although differences between the state tests should be acknowledged when comparing the two. The more problematic aspect of DIBELS is that grade level passages are used to measure reading fluency. Thus, passage difficult is harder for struggling readers than for achieving readers. The apparent lack of specificity and precision of DIBELS subtests for the identification of struggling students should be a concern, as interventions may not be appropriate without an understanding of the scores and what they are measuring.

The main goal here is not to criticize DIBELS or other ORF assessments, but rather to explain how automaticity theory led to a classroom emphasis on fluency instruction and assessment. The importance of automaticity and fluency cannot be diminished as potential predictive tools for reading achievement and for performance on state assessments, or for progress monitoring, and for identifying struggling students. However, individual assessments should not be used in isolation, but rather should be used in tandem with other metrics.

Reliability and Validity of Fluency Measures

There are several concerns regarding the use of ORF measures. Some scholars posit concerns regarding the reliability and validity of using it to predict RC (Chall, 1983; Neddenriep, Hale, Skinner, Hawkins, & Winn, 2007; Poncy, Skinner, & Axtell, 2005), arguing that it is not a direct measure of comprehension. The measures used and the grade level of the student may account for some of the disagreement; the relation between ORF and RC appears to plateau in middle school (Silberglitt et al., 2006). Differences may also be explained by the format of the assessment and how ORF is measured.

As examples, Eason, Sabatini, Goldberg, Bruce, and Cutting (2013) found that ORF, as measured by the Gray Oral Reading Test-Fourth Edition (GORT-4; Wiederholt & Bryant, 2001), correlated ($r = .77$) with the Word Identification and Word Attack subtests of the Woodcock Reading Mastery Test (WMRT; Woodcock, R. W., 1998), but only ($r = .53$) with the GORT-4 Comprehension test. Good and colleagues calculated a correlation coefficient of .67 (2001), while Roehrig, et al. (2008) reported .70 for third graders. More recently, Kim, Wagner, and Foster (2011) examined the relationship between ORF and SRF reporting a strong correlation between the two variables for the full sample ($\varphi = .89$, $p < .001$). Their research also suggested the correlation between ORF and SRF was greater for skilled readers ($r= .79$) than for average readers. The researchers also reported a slightly higher correlation between ORF and RC than between SRF and RC.

Despite some of the questions regarding the reliability and validity of different ORF assessments, schools and classrooms continue the practice of adopting and using

assessment tools, such as DIBELS, Curriculum-Based Measures of reading (CBM-R), Read Naturally (Read Naturally & Ihnot, 2007), Qualitative Reading Inventory-5 (Leslie & Caldwell, 2010), and lesser-known Six-Minute Solutions (Adams & Brown, 2007). Each of these assessments provide leveled materials, and sometimes subtests, for continuous progress monitoring in elementary grades, intervention groups, and special education students. Scores differ by assessment, but commonly identify struggling students based on percentiles and standard scores as follows:

- At/Above Benchmark ≥ 40th percentile or NCE ≥ 44.7
- On Watch, ranging from ≥25th percentile to < 40th percentile, or NCE < 44.7
- Intervention ranging from ≥10th percentile to <25th percentile, or NCE < 35.8
- Urgent Intervention = Below 10th percentile or SS ≤ 23 (Renaissance Learning, 2012)

Information provided by such assessments is valuable in identifying student strengths and weaknesses. More importantly, they help guide classroom and individual instruction. The question is whether they can be used exclusively to identify WCs based on specific, limited criteria, ORF and RC, without additional psychoeducational testing. Time is of the essence in terms of intervention and remediation, thus an investigation may be merited.

Summary

Reading theory influences education policy and research. Automaticity theory is supported and has been further explained to influence RC. After all, a student cannot understand what he cannot read. Word callers are defined as students who read aloud at

an acceptable rate of speed, but with limited or no comprehension of what was just read. Researchers have landed on opposite sides of the discussion regarding the very existence of WCs. Assessments used in the research are largely unavailable, and even impractical, for teachers to administered and interpret efficiently. In order to address the existence of wcs and important instructional needs of these students, it is important that teachers have the ability to identify this group. The larger question is whether teachers can use data from a universal assessment, such as star-r to answer these questions. To this date, no researcher has sought to answer these questions.

Chapter 3 Materials and Procedures

In this chapter, I will describe the theoretical framework driving my investigation, identify the dependent and independent variables. I will present the four questions that will guide my research methodology, including providing participant demographics, assessment description, and the procedures I used to collect, clean, and analyze the data for the participating school district.

Theoretical Framework

Although I do not fully ascribe to a single theoretical perspective in terms of instruction or assessment, current practices influence my approach to investigating the possible existence of WCs. The theoretical framework for my investigation is based on existing instructional and assessment practices following behaviorist theories underlying the simple view of reading and automaticity, involving low-level skills. As it relates to literacy, behaviorist theory asserts that reading is a behavior and that literacy skills are built by developing low-level skills related to reading fluency, phonemic awareness, phonological awareness, and single word decoding, leading to vocabulary development and ultimately to RF and RC. The implication is also that practicing the behavior, e.g. repeated reading, improves RC, the primary goal of reading. These various theories support the practice of assessing the low-level, component skills and lend themselves to using data mining as a methodology for investigating relationships between these skills and WC status. Using data mining procedures, I also investigated the potential influence of dichotomous, independent demographic variables, such as gender, special education status, and free or reduced lunch eligibility, and grade level, which is an ordinal variable.

In this research, the dependent variables include the STAR-R Assessment's RC equivalent, which is the independent reading level (IRL). The IRL is a combination of domains comprised of a set of skills expected at each grade level, including comprehension based on separate scores for literary and informational texts, and vocabulary. The dependent variables are dichotomous, meaning there are only two possibilities, either 0 = non-WC, or 1 = WC. The independent variables, estimated silent reading fluency (RF-E) and IRL are quantitative scores which were converted to Normal Curve Equivalents for statistical analyses, which will be explained.

Research Questions

As established in the previous chapter, there is some support for the notion of the existence of WCs who have been identified using combinations of informal, diagnostic, and other standardized assessments in fourth and fifth grades (Buly & Valencia, 2002; Cain & Oakhill, 2006; Meisinger et al., 2009). Even though classroom teachers identify WCs within their classrooms, researchers have not looked beyond teacher ratings and CBM-R (Feinberg & Shapiro, 2009; Hamilton & Shinn, 2003) to determine whether WCs exist using common, standardized classroom and state assessments. My primary goal for this investigation is to determine if the STAR-R assessment can be used to answer questions regarding the existence of WCs for students in first through fifth grade. My research question focuses on the existence of WCs defined as students whose estimated reading fluency (RF-E) score is 14.04 NCE points greater than RC score (Meisinger et al., 2009). My specific research questions are as follows:

1. What percentage of the students can be identified as WCs based on the definition of word callers where RF-E score is equal to grade level and RF-E score is greater than the RC score by at least 14.04 NCE points?

2. Does the percentage rate of WCs increase across grade levels?

3. Are there differences in number of WCs vs. non-WCs as a function of the following:

    a. Gender

    b. Grade level

    c. Free & reduced lunch status

    d. Those identified with a disability

School and Participant Demographics

In order to conduct my investigation, I contacted two school districts in Tennessee. Both were willing to participate and provided letters of support. However, as I investigated the assessment protocols in the districts, I found one unsuitable, as it did not consistently use the same assessments for all grades in all of its elementary schools. The participating school district is in East Tennessee and made up of mostly small communities. Elementary students are predominantly white, and all schools have been classified as Title I. Each school serves students from kindergarten through fifth grade and range in size, from 328 to 680 pupils. Specific demographics for the six elementary schools are provided in Table I. The number of ELL students in this case, identified as Hispanic, is negligible, with the school reporting the largest population as 1.6% of total students, much lower than the 24% reported by the National Center for Education

Statistics (NCES; 2015). Of interest, however, is the 20% of this sample identified as students with disabilities, which is much higher than estimated 12.9% of the general population (NCES). Participants are first through fifth grade students from the six elementary schools.

Table 1

District Demographics 2014-15

School	White	African-American	Hispanic	Identified Disabilities	Economic Disadvantage	Total Students
1	83.6	11.0	1.6	25.4	81.6	621
2	96.4	2.8	--	20.0	61.9	467
3	91.9	5.1	--	16.9	52.4	681
4	93.6	4.6	--	21.0	57.6	328
5	99.0	--	--	24.9	67.0	374
6	86.8	8.1	--	16.7	85.4	652
Total	91.9	**5.3	1.6	**20.1	**67.7	3123

Reported as percentages and number of students. Other scores reported at percentages.
Retrieved from Tennessee Department of Education website (2015).
** Average percent for district

Assessment Description - STAR Reading

The STAR-R assessment is trademarked, and previously known as Standardized Testing and Reporting (STAR-R; Renaissance Learning, 2005). It is a norm-referenced, computer-based measure designed for students in grades 1-12, and uses item response theory for item selection and adaptive branching. For this reason, kindergarten students were not included in this research. Adaptive branching adjusts test difficulties based on the individual student's responses. STAR measures students' reading abilities in the classroom setting providing estimated reading levels, and suggestions for instruction based on national norms. At each grade level, there are five domains assessed: word

knowledge and skills, comprehension strategies and constructing meaning, understanding author's craft, analyzing literary text, and analyzing argument and evaluating text.

STAR also provides an estimated oral reading fluency (RF-E) per minute. Test developers found a statistical link between RF-E and DIBELS oral reading fluency for students in grades 1-4 (Renaissance Learning, Inc., 2015). Median concurrent and predictive validity coefficients for STAR-R range from 0.68-0.84. Correlations between DIBELS-ORF and STAR-Reading RF-E are provided in Table 2.

Table 2

STAR-Reading and Dynamic Indicators of Basic Skills Correlations

		STAR-R Scale Score		DIBELS wcpm		
Grade	N	M	SD	M	SD	Correlation
1	205	179.31	100.79	45.61	26.75	0.86
2	438	270.04	121.67	71.18	33.02	0.83
3	362	357.95	141.28	86.26	33.44	0.78
4	190	454.04	143.26	102.37	32.74	0.74

M=mean; SD=Standard Deviation.

The STAR technical manual describes the investigation of the correlations between STAR-R and the Stanford Achievement Test, Ninth Edition (Stanford 9; Harcourt Educational Measurement, 1996), and the California Standards Tests, or CST, reporting correlation coefficients ranging from 0.78 to 0.83 for students in grades 3-5. STAR-R scores are reported in several formats e.g., scaled scores (0 – 1400), percentiles, normal curve equivalents, or NCEs, (1 – 99), and grade equivalents. The estimate for overall reliability of the scores is reported as 0.95. Reliabilities for grades 1 to 5 are provided in Table 3.

Table 3

STAR Reading Reliability Estimates

		Generic	Split-Half	Test-Retest		
Grade	N	ρxx	ρxx	N	ρxx	Ave. Days Between Testing
1	7,523	0.91	0.88	298	0.89	8
2	10,132	0.90	0.89	296	0.85	7
3	10,476	0.89	0.89	297	0.82	7
4	9,984	0.89	0.89	297	0.83	7
5	8,352	0.90	0.89	300	0.83	7

ρxx =population reliability of the sum score.

Data Collection Procedures

The director of schools provided a private room, district laptop, and access to individual student summaries and diagnostic reports by school. The data came from three reports I accessed through the Renaissance Learning website. The first report provided demographic information, i.e. grade, gender, special education status, and free/reduced lunch status; the second provided assessment summaries, including percentiles, NCEs, IRL, and RF-E scores; the final report provided diagnostic information reported to parents including domain scores for literature, informational text, and language. I used the student demographic spreadsheet as the base to build my data file. I compared the school demographic and assessment spreadsheets, adding students' individual scores in small sections to insure accuracy. I deleted students with incomplete data. The student summaries included scale scores, grade equivalents, percentiles, NCEs, instructional reading level (IRL), RF-E, ZPD (zone of proximal development) and lexiles. The STAR-R created scaled scores for reporting which are different from commonly used scale scores. Therefore, NCE scores were created and used for analysis. I knew from reviewing the STAR-R technical manuals that RF-E were only reported for first through fourth

grade students. I included students' testing times for all grade levels to analyze potential correlations between testing times and RF-E scores. I did so hoping to create an RF-E score for fifth grade students so that I could include them in my analyses. As I entered new data, I reviewed each line of data to insure I was entering the information in the correct cells. The final spreadsheet included the following variables: four demographic variables, including grade level, gender, free/reduced lunch status, special education status, scaled scores, percentiles, NCEs, instructional reading level (IRL), RF-E, and testing time in minutes (TIM). After I completed data entry, I deleted all student numbers and names, which I verified with a district representative prior to leaving the district office.

Data Cleaning

Because I had worked with each line of data multiple times as I entered data manually, the data did not take long to clean. It should be noted that one school did not have diagnostic data for first or fifth grade students, and much of the incomplete student data for all schools was for students designated as Learning Disabled (LD). One school listed no students in special education, which may suggest these students were not tested, rather than indicating that no special education students attend the school. Prior to data cleaning, I had scores for 2094 students. Table 4 illustrates the district demographics for the students by grade and gender. The population of male students in this sample is 6% greater than for females.

Table 4

District Demographics Gender by Grade

	Gender					
Grade	Female	% in Grade	Male	% in Grade	Group Total	Sample %
1	175	47	196	53	371	17.7
2	183	44	233	56	416	19.9
3	221	49	228	51	449	21.4
4	229	47	256	53	485	23.2
5	176	47	197	53	373	17.8
Total	808	47	913	53	2094	100.0

Lunch status indicated 9% of students were eligible for free or reduced lunch, which does not correlate to previously reported district-wide demographics which indicated 45.6% of students in the district were economically disadvantaged. Because of this discrepancy, the lunch status information is not presented here, and the variable was not included in further analyses. Previously discussed demographic information indicated 21% of the sample population were identified as having a disability, however only 9% of students with complete data were identified as having a disability in the final sample. Because of the discrepancy, the data is not presented in this report, and the variable was excluded from further analyses. The Tennessee Department of Education website indicated a total of 3,127 students in the district for 2014-2015, however, the number of students in first through fifth grades who had complete assessment data and demographic information for STAR end of year 2016 was 1723, or 55% of the total number of students in this sample.

Word Caller Variables

I investigated the existence of WCs based on the IRL variable, which represents the reading comprehension score from STAR-R. The STAR-R reports scores as standard scores (SS) which are different from SSs that report age or grade equivalents, making comparisons challenging. Therefore, I used SPSS to convert the RC variable and RF-E variable scores to a standard NCE score because STAR-R did not provide them. The NCEs were used for further analyses. I used the formula NCE=21.06(raw score) + 50, to convert the mean of the raw scores ($Z= (\bar{x}-x)/SD$) to NCE scores (Mertler, 2002). STAR-R does not calculate an RF-E for fifth grade students, but I hoped to include them so I also created a z-score for student's testing time in minutes (TIM). However, when I analyzed the TIM and RF-E data in SPSS to determine whether there was a relationship between TIM and RF-E, I found there was no significant relationship ($p = 0.06$) between the two variables. This finding suggests that these two scores taken directly from STAR-R measure different dimensions of RF. For this reason,

Table 5

Final School Count by Grade

	Grade Level				
School	1	2	3	4	Total
1	77	80	77	67	301
2	3	16	48	48	115
3	81	102	89	95	367
4	25	35	40	40	140
5	30	45	50	35	160
6	59	74	67	67	267
Total	275	352	371	352	1350
%	20.4	26.1	27.4	26.1	100.0

I did not run further analyses for fifth grade students. The omission of fifth grade students decreased the number of scores to 1350. The grade level demographics for the six schools are illustrated in Table 5. School data indicated 52% (n=702) were male, and 48% (n=647) were female.

Chapter 4 Results

I used RF-E and RC components of the STAR-R assessment for students in grades 1-4 to identify potential WCs. After cleaning the data, I analyzed the data using simple descriptive statistics and crosstabs commands in SPSS, and addressed my research questions. It was not ideal to use a single assessment for my research, but it did allow me to focus solely on RC and RF-E. Results of analyses are discussed in the following sections.

Analysis

To analyze the data, I defined the parameters for word caller status based on the 14.04 NCE point discrepancy between on grade-level RF-E and RC. Using this difference score, I identified 34 students (2.5%) who, in theory, would be identified as WCs. I answered demographic questions using crosstabs and frequency commands in SPSS.

Table 6

NCE Variable Descriptive Statistics

Label	N	Min	Max	M	SD
NCE_RF-E	1349	-16.53	100.39	57.34	16.84
NCE_IRL	1579	29.98	136.74	58.23	15.70

N=number of student scores reported; M=mean of scores; SD=standard deviation

I first looked at the NCE ranges for RF-E and IRL, finding that the ranges for the RF-E were substantially lower for minimum and maximum scores, although means and SDs were much closer (see Table 6). When comparing the NCE minimum and maximum ranges by grade level, the ranges were very different. The NCE score ranges were most

similar among first grade students with the largest range being for the IRL NCE difference ranging from a minimum (min) of 43 to a maximum (max) score of 104. There was only a two-point difference between the means for first graders and a five-point difference between SDs. For second grade students, there was a 10 point NCE difference between the min IRL and RF-E scores. Although the max score range for NCE_IRLs was 13 points higher than the NCE_RF-E, the mean scores were close, which suggest few students achieved high scores. The third-grade max. ranges were lower than for first and second grade students, and the max. NCE_IRL score for third grade students was higher than max score of all grades. The lowest NCE_RF-E scores achieved were among fourth grade students, yet the mean NCE scores for both variables were similar (see Table 7).

Table 7

Variable Descriptives by Grade

Grade	Label	N	Min	Max	M	SD
1	NCE_RF-E	275	34.27	100.39	58.385	17.70
	NCE_IRL	275	43.32	104.14	60.32	12.52
2	NCE_RF-E	352	28.04	97.53	55.81	17.05
	NCE_IRL	352	29.98	110.76	56.11	14.69
3	NCE_RF-E	371	29.27	85.59	56.43	16.23
	NCE_IRL	371	35.42	136.74	56.27	15.44
4	NCE_RF-E	351	-16.53	82.98	59.04	16.43
	NCE_IRL	352	37.78	132.34	58.62	17.52

N=number of student scores reported; M=mean of scores; SD=standard deviation.

Using the NCE scores for the students in grades 1-4, I used SPSS to run frequency reports for the comprehension variable, and to calculate the frequency of students who would be identified as WCs based on the 14.04 NCE point discrepancy. Table 8 provides the frequency, percent of students, and cumulative percent for the NCE point differences.

I determined differences by using SPSS to subtract the NCE score for the IRL variable from the NCE score for RF-E. To analyze the range of NCE difference scores, I used SPSS to subtract the individual student NCE scores from the mean scores, which gives the difference ranges. The NCE score differences range from 14 (rounded) to 48 NCE points. Most scores, (94%) fall between 14.04 and 27 NCE point differences. There were two outliers, one at 38 and one at 48 NCE point differences. Table 8 illustrates the ranges and frequencies of scores in the sample.

Table 8

Frequency and Range of Word Caller Scores by Normal Curve Equivalent

NCE Diff.	Frequency	%	Cum. Per.
14	11	32.4	32.4
15	1	2.9	35.3
16	4	11.8	47.1
17	1	2.9	50.0
18	4	11.8	61.8
19	5	14.7	76.5
20	-	-	-
21	2	5.9	82.4
22	-	-	-
23	1	2.9	85.3
24	1	2.9	88.2
25	-	-	-
26	1	2.9	91.2
27	1	2.9	94.1
28	-	-	-
29	-	-	-
30	-	-	-
31	-	-	-
32	-	-	-
33	-	-	-
34	-	-	-

Table 8 Continued

NCE Diff.	Frequency	%	Cum. Per.
35	-	-	-
36	-	-	-
37	-	-	-
38	1	2.9	97.1
39	-	-	-
40	-	-	-
41	-	-	-
42	-	-	-
43	-	-	-
44	-	-	-
45	-	-	-
46	-	-	-
47	-	-	-
48	1	2.9	100.0
49	-	-	-
50	-	-	-
51	-	-	-
Total	34	100.0	100.0

NCE Diff= NCE_ RF-E score – NCE_ RC.
*NCE Diff score of 14.04 rounded to 14. Other scores rounded to nearest whole number.

WC Prevalence by Grade

The prevalence rate of the WCs falls within the ranges reported in previous research. The highest prevalence was in the first grade where 18 students, or 1.3% of this sample were identified as WCs. In second grade, 10 students were identified (2.9%); in third grade, 2 were identified (0.6%); in fourth grade, 4 were identified (1.2%). The highest incidence of WCs was among first grade students. Rates declined nearly 50% among second grade students. As illustrated in Table 9, WC rates continued to drop in

third and fourth grade, although the rate for fourth grade students was slightly higher than for third grade students.

Table 9

Word Caller Prevalence by Grade

		Grade				
		1	2	3	4	Total
Variable	Label	(n=275)	(n=352)	(n=371)	(n=351)	(N=1350)
WC	Number identified	18.0	10.0	2.0	4.0	34.0
	Within Grade %	6.5	2.9	0.5	1.1	2.5
	Within WC Group %	52.9	29.4	5.9	11.8	100.0
	Within Total N %	1.3	0.7	0.1	0.2	2.5

Word Caller by Gender

The NCE range for the RF-E and IRl variables by gender, illustrated in Table 10, illustrate males achieving the lowest scores. The widest range between the minimum and maximum scores were among males.

Table 10

Descriptives by NCE Variable, Gender

Label	Gender	N	Min.	Max.	M	SD
NCE_RF-E	F	646	28.67	100.39	57.73	16.82
	M	702	-16.53	100.39	56.99	16.88
NCE_IRL	F	757	29.98	132.34	58.60	15.80
	M	821	29.98	136.74	57.89	15.62

N=number of student scores reported; M=mean of scores; SD=standard deviation.

The NCE variable scores ranges in the sample were significantly lower for males for the NCE_RF-E variable than for females. The means and SDs for males and females were similar for both NCE variables. The prevalence rate for females was higher than for males, particularly among first grade students. However, the prevalence rate for males

was higher than females in second grade. WC rates for third and fourth grade students were very low, 0.3-0.6% with no gender differences (see Table 11).

Word Callers versus Non-Word Callers by Grade, Gender

Using data mining procedures, I found very few WCs. Rates of WCs was highest among first and second grade students, where 6.5% of first grade students in this sample were identified, and 2.9% of second grade students in this sample were identified. Rates for third and fourth grade students were 0.6 and 1.2% respectively.

Table 11

Word Callers by Grade, Gender

	Grade							
Label	1	%	2	%	3	%	4	%
Female WC	13	4.7	3	0.9	1	0.3	2	0.6
Male WC	5	1.8	7	2.0	1	0.3	2	0.6
Total	18	6.5	10	2.9	2	0.6	4	1.2

Word Callers versus Non-Word Callers

The prevalence rates of WCs were very low, overall, and in this large sample, there was little difference between male and female students although females were identified at a slightly higher percentage than males (see Table 12). For this reason, the NWC differences between the genders was also negligible.

Table 12

Word Caller versus Non-Word Caller Status by Gender

	Gender			
Label	Female	%	Male	%
WC	19	2.9	15	2.0
NWC	628	97.1	688	98
Total	647	100	703	100

Summary

Overall, the results of my analyses suggest there is a group of students who would be identified as WCs, based on the simple definition using the 14.04NCE minimum difference between on grade-level RF-E and RC variable when using the STAR-R data. Prevalence rates differed by grade level with the highest prevalence rate among first grade students. The rate decreased as grade level increased, particularly in third and fourth grade. The gender analysis suggests a slightly higher prevalence rate for females than for males. Female WCs outnumbered males by nearly 3:1 in first grade. In second grade, the rate of male WCs was higher than for females. The results, overall, are unexpected and merit further discussion.

Chapter 5 Conclusions and Recommendations

The purpose, and first question of my investigation was to determine if WCs exist based on a simple definition using two measures, RF and RC, where RF is on students' grade level, and RC is 14.04 NCE points lower than RF. Using data mining techniques, I identified 2.5% of students in my sample as WCs. Although some researchers found significantly higher rates of WCs, my results showed slightly higher prevalence rates than those reported by Meisinger and colleagues (2009), who identified 1.4% in their diverse population. Spencer, et al. (2014) also identified less than 2% of first and second grade students as WCs in a large, diverse sample of Florida's Reading First schools. It is interesting that these two investigations included more ethnically diverse samples of students than my sample, yet the prevalence rate I found was slightly higher. Both research groups suggested that verbal ability plays a role in RC and, therefore, verbal ability should be included in RC evaluations. If verbal ability is a component of RC, then this may highlight potential deficits in verbal abilities of white children that is overlooked in universal screening practices. Buly and Valencia (2002) identified a significantly larger prevalence of WCs (18%) in their investigation identifying 60% of their automatic WCs as poor, ELL students, reflecting back to the language component.

Analysis for the second question, regarding whether the frequency of WCs increases across grade levels, yielded inconsistent results. More students were identified in first grade than in later grades, where they were essentially non-existent. This finding contradicts previous research suggesting WCs emerge in later elementary grades (Catts et al., 2012; Chall & Jacobs, 2003; Compton et al., 2012; Knight-Teague, et al., 2014;

Leach et al., 2003; Meisinger, et al., 2009). These previous investigations consistently found an increase in the number of students identified as grade level increased, particularly among fourth grade students as the increased use of informational text places greater cognitive demands on students at this level. Meisinger and colleagues surmised that the use of a silent reading fluency measure would have increased students' comprehension scores. However, the use of STAR-R Assessment's RF-E, which is an SRF measure, may not support this notion. Furthermore, the STAR-R Assessment's data, which illustrates a decrease in the number of WCs, may be an effective tool for identifying students who struggle with reading comprehension. If so, it is being used to alert teachers to their students' RC difficulties, which can then be effectively targeted and ameliorated using specific RC instruction. In order to make this determination, the 34 students identified as WCs would need to be isolated so that the type of intervention(s) these children received that improved RC can be evaluated, as it would be worthy of sharing with others.

The final question included two variables, gender and grade level, as the lunch status and disability status variables were excluded. Overall, the differences were not significant for males and females within WC groups, however females were identified at slightly higher rates than males, similar to findings of Catts and colleagues (2013) who found insignificant gender differences. Few of the investigations I reviewed included gender as a factor in predicting WCs.

Stanovich (1993) suggested that using "word caller" to refer to these students was too general and overlooked students with specific learning difficulties. What is

interesting, though, is when using the two salient features, RF and RC, to identify WCs among a predominantly white population, excluding potential factors, such as free and reduced lunch status, identification of learning disability, and limited English-speaking ability, WCs are still identified.

Limitations

My focus on RF-E and RC prevents me from making any assumptions regarding relationships or causes for word calling, such as receptive or expressive listening or language abilities. There were several limitations of my research which likely influenced the results. One limitation of my research was that I had only one standardized assessment, which means I had no way to validate my results. As a result, I share the concerns of other researchers regarding the validity of the information gleaned from any single assessment (Koretz, 2003). The explanation of how RF-E is determined was unclear. I would have enjoyed more transparency as I could have attempted to include fifth grade data in my analysis. Additional limitations in the study include time and access to student data. I was informed up front that I would have one opportunity to gather student data. Once I accessed the data, I found scores and demographic information were missing for approximately one third of the students. One school reported very limited data for first and fifth grade students.

Recommendations

Looking forward, this research should be repeated with multiple assessments including one true measure of ORF, including prosodic elements, and at least one

additional standardized RC assessment to validate these results. Because the STAR-R's comprehension score is comprised of multiple component skills specifically related RC as is relates to literary and informational texts, vocabulary, and low level skills, it would be interesting to investigate possible correlations among these domain variables. Another interesting investigation would be a comparison of STAR-R and other online assessments that provide similar data in terms of identifying WCs.

Identifying struggling readers is important, but identifying specific area(s) of weakness is crucial for planning literacy instruction. The domain scores for literature and informational text may be informative and beneficial for teachers to use for identifying WCs and other struggling readers, but teachers must take time to analyze each student's domain scores. The accuracy of the RC and RF-E should be compared to other measures. Unfortunately, the school district did not have scores for the state standardized test because of technical issues at the state level.

Contribution to Research

My results confirm the existence of WCs and further suggest the STAR-R may be used to identify them. Word callers have been identified in investigations exploring correlations of specific skills and different demographic variables, while utilizing various research methods, assessment formats, and statistical analyses. The fact that my sample was mostly white suggests further investigations need to be conducted regarding additional components of RC. Although I did not embark on my research journey planning for the data limitations, it necessarily focused my attention on the two-variable definition of the WC resulting in a clear picture of this small population of students with

no learning disabilities who are from a predominantly Caucasian, English-speaking, rural community. These kids struggle with reading comprehension. While STAR-R looks as though it has been used effectively in this district, I am skeptical based on a visual review of the data as the time students spent testing did not consistently correlate with higher or lower scores. This suggests a need for more transparency in how the STAR-R calculates comprehension.

During my research process, I reviewed numerous universal assessments and the skills they measure. The STAR-R measures are very common among other universal assessments, and while they are useful, they should not be used as the only measure to assess students' individual skills, or as the only basis for planning instruction. Time is important, and it is not unlimited, therefore we should not waste it by conducting assessments that provide repetitive or limited information about individual student needs abilities. We know that every child is different, but we also know there are some fundamental skills necessary to set strong foundations for reading, math, and comprehension. We have walked around our current educational path many times since the inception of public education. I believe we need to take a more pragmatic view of educating children, set our expectations higher for them, and for ourselves as educators. We need to crack open our theoretical frameworks, mix things up a bit, and let in some fresh air. It is likely we need to follow some of those less traveled to improve the reading comprehension ability of some struggling comprehenders, such as the word caller. My research does not negate the potential contribution of factors, such as expressive and language comprehension, life experiences, exposure to print, and even imagery. In fact, I

would encourage the development of protocols that can be used by teachers to help determine the causes and solutions to some children's learning problems.

References

Abbott, M., Wills, H., Miller, A., & Kaufman, J. (2012). The relationship of error rate and comprehension in second and third grade oral reading fluency. Reading Psychology, 33(1/2), 104–132. http://doi.org/10.1080/02702711.2012.630613

Adams, G., & Brown, S. (2007). The Six-Minute Solution: A reading fluency program. Sopris West Educational Services.

Airasian, P. W., & Madaus, G. F. (1983). Linking testing and instruction: Policy issues. Journal of Educational Measurement, 20(2), 103–118.

Alber-Morgan, S. R., Ramp, E. M., Anderson, L. L., & Martin, C. M. (2007). Effects of repeated readings, error correction, and performance feedback on the fluency and comprehension of middle school students with behavior problems. Journal of Special Education, 41(1), 17–30.

Allington, R. L. (1977). If they don't read much, how they ever gonna get good? Journal of Reading, 21(1), 57–61.

Allington, R. L. (1983). Fluency: The neglected reading goal. The Reading Teacher, 36(6), 556–561.

Allington, R. L., & Brown, S. (1979). Fact: A multi-media reading program. Milwaukee, Wis.: Raintree Publishers.

Anderson, B. (1981). Missing ingredient: Fluent oral reading. Education Digest, 46, 53-55.

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. Technical Report No. 306. U.S. Department of education:

Center for the Study of Reading. ED239236. Accessed 1/4/2015 http://eric.ed.gov/?id=ED239236

Armbruster, B. B. (2003). Put reading first: The research building blocks of reading instruction: Kindergarten through grade 3. Washington, D.C.: National Institute for Literacy. Retrieved from http://babel.hathitrust.org/cgi/pt?id=pur1.32754074659057;view=1up;seq=3

Baker, D., Biancarosa, G., Park, B., Bousselot, T., Smith, J.-L., Baker, S., … Tindal, G. (2015). Validity of CBM measures of oral reading fluency and reading comprehension on high-stakes reading assessments in Grades 7 and 8. Reading and Writing, 28(1), 57–104. https://doi.org/10.1007/s11145-014-9505-4

Barone, D. (2013). Important revelations about school reform: looking at and beyond Reading First. School Effectiveness and School Improvement, 24(4), 392–416. http://doi.org/10.1080/09243453.2012.693104.

Barth, A. E., Stuebing, K. K., Anthony, J. L., Denton, C. A., Mathes, P. G., Fletcher, J. M., & Francis, D. J. (2008). Agreement among response to intervention criteria for identifying responder status. Part of a Special Issue Entitled The promises and potential challenges of RTI: Data-based evaluations of the concept and related practices, 18(3), 296–307. https://doi.org/10.1016/j.lindif.2008.04.004

Bashir, A., & Hook, P. (2009). Fluency: a key link between word identification and comprehension. Language, Speech & Hearing Services in Schools, 40(2), 196–200. https://doi.org/10.1044/0161-1461(2008/08-0074)

Begeny, J. C., & Martens, B. K. (2006). Assisting low-performing readers with a group-based reading fluency intervention. School Psychology Review, 35(1), 91–107.

Block, C. C., Parris, S. R., Reed, K. L., Whiteley, C. S., & Cleveland, M. D. (2009). Instructional approaches that significantly increase reading comprehension. Journal of Educational Psychology, 101(2), 262–281. http://doi.org/10.1037/a0014319.

Buly, M. R., & Valencia, S. W. (2002). Below the bar: Profiles of students who fail state reading assessments. Educational Evaluation and Policy Analysis, 24(3), 219–239. http://doi.org/10.3102/01623737024003219

Burns, M. K., Kwoka, H., Lim, B., Crone, M., Haegele, K., Parker, D. C., … Scholin, S. E. (2011). Minimum reading fluency necessary for comprehension among second-grade students. Psychology in the Schools, 48(2), 124–132. https://doi.org/10.1002/pits.20531

Cain, K., & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. British Journal of Educational Psychology, 76(4), 683–696. https://doi.org/10.1348/000709905X67610

Cain, K., & Oakhill, J. (2011). Matthew effects in young readers reading comprehension and reading experience aid vocabulary development. Journal of Learning Disabilities, 44(5), 431–443. http://doi.org/10.1177/0022219411410042.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills.

Journal of Educational Psychology, 96(1), 31–42. https://doi.org/http://dx.doi.org.proxy.lib.utk.edu:90/10.1037/0022-0663.96.1.31

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. Memory & Cognition, 29(6), 850–859. https://doi.org/10.3758/BF03196414.

Campbell, D. E. (1976). Democratic preference functions. Journal of Economic Theory. 12, 259-272.

Carretti, B., Cornoldi, C., De Beni, R., & Romanò, M. (2005). Updating in working memory: A comparison of good and poor comprehenders. Journal of Experimental Child Psychology, 91(1), 45–66. https://doi.org/10.1016/j.jecp.2005.01.005.

Catts, H. W. (1997). The early identification of language-based reading disabilities. Language, Speech & Hearing Services in Schools, 28(1), 86–89.

Catts, H. W., Compton, D., Tomblin, J. B., & Bridges, M. S. (2012). Prevalence and nature of late-emerging poor readers. Journal of Educational Psychology, 104(1), 166–181. http://doi.org/http://dx.doi.org.proxy.lib.utk.edu:90/10.1037/a0025323

Catts, H. W., Petscher, Y., Schatschneider, C., Bridges, M. S., & Mendoza, K. (2009). Floor effects associated with universal screening and their impact on the early identification of reading disabilities. Journal of Learning Disabilities, 42(2), 163–176. http://doi.org/10.1177/0022219408326219

Chall, J. S. (1983). Literacy: Trends and explanations. Educational Researcher, 12(9), 3–8. https://doi.org/10.3102/0013189X012009003.

Chall, J. S., & Jacobs, V. A. (2003). Poor children's fourth-grade slump | Reading topics A-Z. Retrieved December 23, 2013, from http://www.aft.org/newspubs/periodicals/ae/spring2003/hirschsbclassic.cfm

Compton, D. L., Fuchs, D., Fuchs, L. S., Elleman, A. M., & Gilbert, J. K. (2008). Tracking children who fly below the radar: Latent transition modeling of students with late-emerging reading disability. Learning and Individual Differences, 18(3), 329–337. https://doi.org/10.1016/j.lindif.2008.04.003

Cornoldi, C., Rigoni, F., Tressoldi, P. E., & Vio, C. (1999). Imagery deficits in nonverbal learning disabilities. Journal of Learning Disabilities, 32(1), 48–57.

Cornoldi, C., Vecchia, R. D., & Tressoldi, P. E. (1995). Visuo-spatial working memory limitations in low visuo-spatial high verbal intelligence children. Journal of Child Psychology and Psychiatry, 36(6), 1053–1064. https://doi.org/10.1111/j.1469-7610.1995.tb01350.x

Cramer, K., & Rosenfield, S. (2008). Effect of degree of challenge on reading performance. Reading & Writing Quarterly, 24(1), 119–137. https://doi.org/10.1080/10573560701501586

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension Is measured. Scientific Studies of Reading, 10(3), 277–299. https://doi.org/10.1207/s1532799xssr1003_5.

Daly, E. J., III, Bonfiglio, C. M., Mattson, T., Persampieri, M., & Foreman-Yates, K. (2006). Refining the experimental analysis of academic skills deficits: Part II. Use

of brief experimental analysis to evaluate reading fluency treatments. Journal of Applied Behavior Analysis, 39(3), 323–331.

Danks, J. H., & Fears, R. (1976, June). Oral reading: Does it reflect decoding or comprehension? Learning Research and Development Center. Retrieved from http://files.eric.ed.gov/fulltext/ED155644.pdf

Denckla, M. B. (1979). Childhood learning disabilities. In Denckla, M. B. Clinical Neuropsychology (1979): 535-573. (pp. 535–573). New York: Oxford University Press.

Dennis, D. V. (2013). Heterogeneity or homogeneity: What assessment data reveal about struggling adolescent readers. Journal of Literacy Research, 45(1), 3–21. https://doi.org/10.1177/1086296X12468431

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

Deno, S. L. (2003). Developments in curriculum-based measurement. The Journal of Special Education, 37(3), 184–192.

Dewey, J. (2004). Democracy and education. Courier Corporation.

Dowhower, S. L. (1991). Speaking of prosody: Fluency's unattended bedfellow. Theory into Practice, 30(3), 165–175.

Duke, N. K. (2000). 3.6 minutes per day: the scarcity of informational texts in first grade. Reading Research Quarterly, 35(2), 202–224. http://doi.org/10.1598/RRQ.35.2.1

Durkin, D. (1979). What classroom observations reveal about reading comprehension instruction. Reading Research Quarterly, 14(4), 481–533.

Dymock, S. (1993). Reading but not understanding. Journal of Reading, 37(2), 86–91.

Eason, S. H., Sabatini, J., Goldberg, L., Bruce, K., & Cutting, L. E. (2013). Examining the relationship between word reading efficiency and oral reading rate in predicting comprehension among different types of readers. Scientific Studies of Reading, 17(3), 199–223.

Eason, S. H., Goldberg, L. F., Young, K. M., Geist, M. C., & Cutting, L. E. (2012). Reader–text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. Journal of Educational Psychology, 104(3), 515–528. https://doi.org/http://dx.doi.org.proxy.lib.utk.edu:90/10.1037/a0027182.

Eldredge, J. L. (2005). Foundations of Fluency: An Exploration. Reading Psychology, 26(2), 161–181. http://doi.org/10.1080/02702710590930519

Etmanskie, J. M., Partanen, M. & Siegel, L. S. . (2016). A Longitudinal Examination of the Persistence of Late Emerging Reading Disabilities. Journal of Learning Disabilities, 49(1), 21–35. https://doi.org/10.1177/0022219414522706

Falk-Ross, F., Szabo, S., Sampson, M. B., & Foote, M. M. (2009). Literacy issues during changing times: A call to action. The thirtieth yearbook: A peer reviewed publication of the College Reading Association. [Papers from the College Reading Association Conference, 2007]. College Reading Association. Retrieved from http://eric.ed.gov/?id=ED507687

Feinberg, A. B., & Shapiro, E. S. (2009). Teacher accuracy: An examination of teacher-based judgments of students' reading with differing achievement levels. The Journal of Educational Research, 102(6), 453–462.

Fisher, C. W., & Berliner, D. C. (1985). Perspectives on instructional time. Longman.

Fleisher, L. S., Jenkins, J. R., & Pany, D. (1979). Effects on poor readers' comprehension of training in rapid decoding. Reading Research Quarterly, 15(1), 30–48. https://doi.org/10.2307/747430.

Fletcher, J. M., Stuebing, K. K., Barth, A. E., Miciak, J., Francis, D. J., & Denton, C. A. (2014). Agreement and coverage of indicators of Response to Intervention. Topics in Language Disorders, 34(1), 74–89. https://doi.org/10.1097/TLD.0000000000000004

Freeland, J. T., Skinner, C. H., Jackson, B., McDaniel, C. E., & Smith, S. (2000). Measuring and increasing silent reading comprehension rates: Empirically validating a repeated readings intervention. ResearchGate, 37(5), 415–429. https://doi.org/10.1002/1520-6807(200009)37:5<415::AID-PITS2>3.0.CO;2-L

Fuchs, L. S., & Fuchs, D. (1992). Identifying a measure for monitoring student reading progress. School Psychology Review, 21(1), 45.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? Reading Research Quarterly, 41(1), 93–99.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis.

Scientific Studies of Reading, 5(3), 239–256. http://doi.org/10.1207/S1532799XSSR0503_3

Ghelani, K., Sidhu, R., Jain, U., & Tannock, R. (2004). Reading comprehension and reading related abilities in adolescents with reading disabilities and attention-deficit/hyperactivity disorder. Dyslexia, 10(4), 364–384. https://doi.org/10.1002/dys.285

Good, R. H., & Kaminski, R. A. (2002). DIBELS oral reading fluency passages for first through third grades (No. 10, pp. 6-10). (Technical Report) Eugene, OR: University of Oregon.

Good, R. H., III, Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. Scientific Studies of Reading, 5(3), 257–88.

Gough, P. B. (1996). How children learn to read and why they fail. Annals of Dyslexia, 46(1), 1–20. https://doi.org/10.1007/BF02648168

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. Remedial and Special Education, 7(1), 6–10.

Guszak, F. J. (1967). Teacher questioning and reading. An introduction to elementary reading. The Reading Teacher 21(3), 227-234.

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

Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., … Tonks, S. (2004). Increasing reading comprehension and engagement through Concept-Oriented Reading Instruction. Journal of Educational Psychology, 96(3), 403–423. http://doi.org/10.1037/0022-0663.96.3.403

Hammill, D., Wiederholt, J., & Allen, E. (2012). Test of Silent Contextual Reading Fluency–Second Edition. Austin, TX: Pro-Ed, Inc.

Hamilton, C., & Shinn, M. R. (2003). Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skills. School Psychology Review, 32(2), 228.

Harcourt Brace Educational Measurement. (1996). Stanford Achievement Test (9th ed.). San Antonio, TX: Harcourt Assessment.

Harnadek, M. C., & Rourke, B. P. (1994). Principal identifying features of the syndrome of nonverbal learning disabilities in children. Journal of Learning Disabilities, 27(3), 144–154.

Hasbrouck, J. E., & Tindal, G. (1992). Curriculum-based oral reading fluency norms for students in grades 2 through 5. Teaching Exceptional Children, 24, 41–44.

Holliman, A., Williams, G. , Mundy, I. , Wood, C. , Hart, L. ., & Waldron, S. . (2014). Beginning to disentangle the prosody-literacy relationship: a multi-component measure of prosodic sensitivity. Reading & Writing, 27(2), 255–266. https://doi.org/10.1007/s11145-013-9443-6

Homan, S. P., Klesius, J. P., & Hite, C. (1993). Effects of repeated readings and nonrepetitive strategies on students' fluency and comprehension. The Journal of Educational Research, 87(2), 94–99.

Hoover, W., & Gough, P. (1990). The simple view of reading. Reading and Writing, 2(2), 127–160. https://doi.org/10.1007/BF00401799

Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for at-risk readers in a Response to Intervention framework. School Psychology Review, 36(4), 582–600.

Johnson, E. S., Jenkins, J. R., Petscher, Y., & Catts, H. W. (2009). How can we improve the accuracy of screening instruments? Learning Disabilities Research & Practice, 24(4), 174–185. https://doi.org/10.1111/j.1540-5826.2009.00291.x

Jorm, A. F. (1983). The psychology of reading and spelling disabilities. London; Boston; Routledge & Kegan.

Judge, S., & Bell, S. M. (2010). Reading achievement trajectories for students with learning disabilities during the elementary school years. Reading & Writing Quarterly, 27(1–2), 153–178.

Keenan, J. M., & Betjemann, R. S. (2006). Comprehending the Gray Oral Reading Test without reading It: Why comprehension tests should not include passage-independent items. Scientific Studies of Reading, 10(4), 363–380. https://doi.org/10.1207/s1532799xssr1004_2

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral

comprehension. Scientific Studies of Reading, 12(3), 281–300. https://doi.org/10.1080/10888430802132279

Kena, G., Hussar, W., McFarland, J., de Brey, C., Musu-Gillette, L., Wang, X., … Diliberti, M. (2016). The Condition of Education 2016. NCES 2016-144. Jessup, MD: National Center for Education Statistics.

Kim, Y.-S., Wagner, R. K., & Foster, E. (2011). Relations among oral reading fluency, silent reading fluency, and reading comprehension: A latent variable study of first-grade readers. Scientific Studies of Reading, 15(4), 338–362.

Kim, Y.-S., Wagner, R. K., & Lopez, D. (2012). Developmental relations between reading fluency and reading comprehension: A longitudinal study from Grade 1 to Grade 2. Journal of Experimental Child Psychology, 113(1), 93–111. https://doi.org/10.1016/j.jecp.2012.03.002

Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. Journal of Educational Psychology, 100(2), 310.

Klein-Braley, C. (1997). C-Tests in the context of reduced redundancy testing: an appraisal. Language Testing, 14(1), 47–84. https://doi.org/10.1177/026553229701400104

Knight-Teague, K., Vanderwood, M. L., & Knight, E. (2014). Empirical investigation of word callers who are English learners. School Psychology Review, 43(1), 3+.

Koretz, D. (2003), Using multiple measures to address perverse incentives and score inflation. Educational Measurement: Issues and Practice, 22: 18–26. doi:10.1111/j.1745-3992.2003.tb00124.x

Kuhn, M. R., Schwanenflugel, P. J., & Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. Reading Research Quarterly, 45(2), 230–251. http://doi.org/10.1598/RRQ.45.2.4

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

Leach, J. M., Scarborough, H. S., & Rescorla, L. (2003). Late-emerging reading disabilities. Journal of Educational Psychology, 95(2), 211.

Leslie, L., & Caldwell, J. S. (1995). Qualitative Reading Inventory-II. Glenview, IL: Harper Collins College Publishers.

Leslie, L., & Caldwell, J. S. (2010). Qualitative Reading Inventory-V. Pearson

Lipson, M. Y., & Wixson, K. K. (1986). Reading disability research: An interactionist perspective. Review of Educational Research, 56(1), 111–136.

Madaus, G. F. (1988). The distortion of teaching and testing: High-stakes testing and instruction. Peabody Journal of Education, 65(3), 29–46.

Madaus, G., & Russell, M. (2010). Paradoxes of high-stakes testing. Journal of Education, 21–30.

Mammarella, I. C., & Pazzaglia, F. (2010). Visual perception of children at risk of nonverbal learning disabilities. Child Neuropsychology, 16(6), 564–576. https://doi.org/10.1080/09297049.2010.485125

Mather, N., Hammill, D. D., Allen, E. A., & Roberts, R. (2014). Test of Silent Word Reading Fluency-Second Edition. Austin, TX: PRO-ED.

McBride, N. (2005). The effectiveness of Second Shot and/or Lindamood-Bell on reading achievement of elementary students. ProQuest, UMI Dissertations Publishing.

McInnes, A., Humphries, T., & Hogg-Johnson, S. (2003). Listening comprehension and working memory are impaired in Attention-Deficit Hyperactivity Disorder irrespective of language impairment. Journal of Abnormal Child Psychology, 31(4), 427–443.

Meisinger, E. B., Bloom, J. S., & Hynd, G. W. (2010). Reading fluency: Implications for the assessment of children with reading disabilities. Annals of Dyslexia, 60(1), 1–17.

Meisinger, E. B., Bradley, B. A., Schwanenflugel, P. J., Kuhn, M. R., & Morris, R. D. (2009). Myth and reality of the word caller: The relation between teacher nominations and prevalence among elementary school children. School Psychology Quarterly, 24(3), 147–159. http://doi.org/http://dx.doi.org.proxy.lib.utk.edu:90/10.1037/a0017191

Morris, R. D. (1998). Subtypes of reading disability: Variability around a phonological core. Journal of Educational Psychology., 90(3), 347–73.

National Center for Education Statistics, U.S. Department of Education (2014). Institute of Education Sciences. In Digest of education statistics (2015). https://nces.ed.gov/programs/coe/pdf/Indicator_CGE/coe_cge_2015_05.pdf.

National Reading Panel. (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. US Dept of Health and Human Services, Public Health Service, National Institutes of Health, National Institute of Child Health and Human Development. Retrieved from https://www.nichd.nih.gov/publications/pubs/nrp/Documents/report.pdf

Nation, K., Cocksey, J., Taylor, J. S., & Bishop, D. V. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. Journal of Child Psychology and Psychiatry, 51(9), 1031–1039.

Nation, K., & Snowling, M. J. (2000). Factors influencing syntactic awareness skills in normal readers and poor comprehenders. Applied Psycholinguistics, 21(02), 229–241. https://doi.org/null

Nation, K., Snowling, M. J., & Clarke, P. (2007). Dissecting the relationship between language skills and learning to read: Semantic and phonological contributions to new vocabulary learning in children with poor reading comprehension. International Journal of Speech-Language Pathology, 9(2), 131–139.

Neddenriep, C. E., Hale, A. D., Skinner, C. H., Hawkins, R. O., & Winn, B. D. (2007). A preliminary investigation of the concurrent validity of reading comprehension

rate: A direct, dynamic measure of reading comprehension. Psychology in the Schools, 44(4), 373–388.

O'Connor, R. E., White, A., & Lee Swanson, H. (2007). Repeated reading versus continuous reading: Influences on reading fluency and comprehension. Exceptional Children, 74(1), 31–46.

Palladino, P., Cornoldi, C., Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. Memory & Cognition, 29(2), 344–354. https://doi.org/10.3758/BF03194929.

Perfetti, C. A., & Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. Journal of Educational Psychology, 67(4), 461–469. https://doi.org/http://dx.doi.org.proxy.lib.utk.edu:90/10.1037/h0077013

Perfetti, C. A., & Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences. Retrieved from http://eric.ed.gov/?id=ED145400

Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. Merrill-Palmer Quarterly (1982-), 283–319.

Pimperton, H., & Nation, K. (2010). Suppressing irrelevant information from working memory: Evidence for domain-specific deficits in poor comprehenders. Journal of Memory and Language, 62(4), 380–391. https://doi.org/10.1016/j.jml.2010.02.005.

Pinnell, G. S., National Assessment of Educational Progress, P., NJ., Educational Testing Service, P., NJ. Center for the Assessment of Educational Progress., & Others.

(1995). Listening to children read aloud: Data from NAEP's Integrated Reading Performance Record (IRPR) at Grade 4 (No. 0-88685-167-X).

Poncy, B. C., Skinner, C. H., & Axtell, P. K. (2005). An investigation of the reliability and standard error of measurement of words read correctly per minute using curriculum-based measurement. Journal of Psychoeducational Assessment, 23(4), 326–338.

Popplewell, S. R., & Doty, D. E. (2001). Classroom instruction and reading comprehension: A comparison of one basal reader approach and the four-blocks framework. Reading Psychology, 22(2), 83–94.

Pressley, M., & Afflerbach, P. (2009). Verbal protocols of reading: The nature of constructively responsive reading. New York: Routledge.

Pressley, M., & Gaskins, I. W. (2006). Metacognitively competent reading comprehension is constructively responsive reading: how can such reading be developed in students? Metacognition & Learning, 1(1), 99–113. http://doi.org/10.1007/s11409-006-7263-7

Pressley, M., Hilden, K., & Shankland, R. (2005). An evaluation of end-grade-3 Dynamic Indicators of Basic Early Literacy Skills (DIBELS): Speed reading without comprehension, predicting little. Unpublished Manuscript, Michigan State University, Literacy Achievement Research Center.

Pressley, M., Johnson, C. J., Symons, S., McGoldrick, J. A., & Kurita, J. A. (1989). Strategies that improve children's memory and comprehension of text. The Elementary School Journal, 90(1), 3–32.

Price, K. W., Meisinger, E. B., Louwerse, M. M., & D'Mello, S. (2016). The contribution of oral and silent reading fluency to reading comprehension. Reading Psychology, 37(2), 167–201.

Read Naturally, I., & Ihnot, C. (2007). Read Naturally Masters Edition. Read Naturally, Inc.

Renaissance Learning, Inc. (2015). STAR Reading Technical Manual (Technical Manual). Renaissance Learning, Inc. Retrieved from http://doc.renlearn.com/KMNet/R004384910GJF6AC.pdf

Reutzel, D. R., & Hollingsworth, P. M. (1993). Effects of fluency training on second graders' reading comprehension. The Journal of Educational Research, 86(6), 325–331.

Roehrig, A., Petscher, Y., Nettles, S., Hudson, R. F., & Torgesen, J. K. (2008). Accuracy of the DIBELS oral reading fluency measure for predicting third grade reading comprehension outcomes. Journal of School Psychology, 46(3), 343–66. http://doi.org/10.1016/j.jsp.2007.06.006

Rosenblatt, L. M. (1993). The transactional theory: Against dualisms. College English, 55(4), 377–386. https://doi.org/10.2307/378648

Rourke, B. P. (1989). Nonverbal learning disabilities: The syndrome and the model. New York: Guilford Press.

Rourke, B. P. (1995). Syndrome of nonverbal learning disabilities: Neurodevelopmental manifestations. New York: Guilford Press.

Ruddell & N. Unrau (Eds.). (2004). Theoretical models and processes of reading -5th edition. pp. 1570-1613). Newark, DE: International Reading Association.

Samuels, S. J. (1976). Automatic decoding and reading comprehension. Language Arts, 53(3), 323–325.

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

Scarborough, H. S. (1989). Prediction of reading disability from familial and individual differences. Journal of Educational Psychology, 81(1), 101–108. https://doi.org/http://dx.doi.org.proxy.lib.utk.edu:90/10.1037/0022-0663.81.1.101

Schrank, F. A., Mather, N., & McGrew, K. S. (2014). Woodcock-Johnson IV Tests of Achievement. Rolling Meadows, IL: Riverside.

Schwanenflugel, P., Westmoreland, M., & Benjamin, R. (2015). Reading fluency skill and the prosodic marking of linguistic focus. Reading & Writing, 28(1), 9–30. https://doi.org/10.1007/s11145-013-9456-1.

Silberglitt, B., Burns, M. K., & Madyun, N. H. (2006). Relationship of reading fluency assessment data with state accountability test scores: A longitudinal comparison of grade levels. Psychology in the Schools, 43(5), 527–535. https://doi.org/10.1002/pits.20175

Silverman, R. D., Speece, D. L., Harring, J. R., & Ritchey, K. D. (2013). Fluency has a role in the Simple View of Reading. Scientific Studies of Reading, 17(2), 108–133. http://doi.org/10.1080/10888438.2011.618153

Snow, C. (2002). Reading for understanding. Towards an R & D program in reading comprehension. Santa Monica, CA: RAND Corporation, 2002. http://www.rand.org/pubs/monograph_reports/MR1465.html.

Snow, C. E., Burns, M. C., & Griffin, P. (1998). Preventing reading difficulties in young children: Committee on the prevention of reading difficulties in young children. National Research Council. Retrieved from http://site.ebrary.com.proxy.lib.utk.edu:90/lib/utennk/reader.action?docID=10040951

Spear-Swerling, L. (2004). Fourth graders' performance on a state-mandated assessment involving two different measures of reading comprehension. Reading Psychology, 25(2), 121–148. https://doi.org/10.1080/02702710490435727

Spencer, M., Quinn, J. M., & Wagner, R. K. (2014). Specific reading comprehension disability: Major problem, myth, or misnomer? Learning Disabilities Research & Practice, 29(1), 3–9. https://doi.org/10.1111/ldrp.12024

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

Stanovich, K. E. (1993). Romance and reality. Reading Teacher, 47, 280–280.

Stothard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children. Reading and Writing, 4(3), 245–256. https://doi.org/10.1007/BF01027150.

Stothard, S. E., & Hulme, C. (1995). A comparison of phonological skills in children with reading comprehension difficulties and children with decoding difficulties.

Journal of Child Psychology and Psychiatry, 36(3), 399–408. https://doi.org/10.1111/j.1469-7610.1995.tb01298.x.

Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading: A meta-analysis. Remedial and Special Education, 25(4), 252–261. http://doi.org/10.1177/07419325040250040801

Therrien, W. J., Wickstrom, K., & Jones, K. (2006). Effect of a combined repeated reading and question generation intervention on reading achievement. Learning Disabilities Research & Practice (Wiley-Blackwell), 21(2), 89–97. http://doi.org/10.1111/j.1540-5826.2006.00209.x

Torgesen, J. K. (2002). The prevention of reading difficulties. Journal of School Psychology, 40(1), 7–26. https://doi.org/10.1016/S0022-4405(01)00092-9

Torppa, M., Tolvanen, A., Poikkeus, A.-M., Eklund, K., Lerkkanen, M.-K., Leskinen, E., & Lyytinen, H. (2007). Reading development subtypes and their early characteristics. Annals of Dyslexia, 57(1), 3–32.

Valencia, S. W., & Buly, M. R. (2004). Behind test scores: What struggling readers really need. Reading Teacher, 57(6), 520–533.

Veenendaal, N. J., Groen, M. A., & Verhoeven, L. (2015). What oral text reading fluency can reveal about reading comprehension. Journal of Research in Reading, 38(3), 213–225. https://doi.org/10.1111/1467-9817.12024

Vellutino, F. R. (1991). Introduction to three studies on reading acquisition: Convergent findings on theoretical foundations of code-oriented versus whole-language

approaches to reading instruction. Journal of Educational Psychology, 83(4), 437–443. https://doi.org/http://dx.doi.org/10.1037/0022-0663.83.4.437

Wiederholt, J., & Bryant, B. R. (2001). Gray Oral Reading Tests— Fourth Edition. Pro-Ed, Inc.

Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., Wolf, M., Kuhn, M., … Schwanenflugel, P. (2010). The relationship between different measures of oral reading fluency and reading comprehension in second-grade students who evidence different oral reading fluency difficulties. Language, Speech & Hearing Services in Schools, 41(3), 340–348. https://doi.org/10.1044/0161-1461(2009/08-0093)

Wolf, M., & Katzir-Cohen, T. (2001). Reading Fluency and Its Intervention. Scientific Studies of Reading, 5(3), 211–239.

Woodcock, Schrank, McGrew, & Mather, 2014. Woodcock–Johnson III-Reading Fluency subtest

Woodcock, R. W. (1998). Woodcock Reading Mastery Test–Revised/Normative Update. Circle Pines, MN: American Guidance Service.

Universal Screening Center on Response to Intervention. Retrieved from http://www.rti4success.org/essential-components-rti/universal-screening; 01-01-2017;14:33:31.

Yuill, N., & Oakhill, J. (1991). Children's problems in text comprehension: An experimental investigation. New York, NY: Cambridge University Press.

Vita

Rhonda DeJournett was born in Macomb, Illinois and raised in Colorado. She attended Manchester College in Indiana graduating with a Bachelor of Arts degree in English in 1985. Although Rhonda's primary occupational experience has been in management, her positions required her to develop teaching skills as she trained, supervised, and mentored her staff in various industries. Rhonda lacked personal fulfillment as a manager and eventually pursued a career as a private instructor at the Fort Wayne Center for Learning where she found working with struggling students to be personally rewarding.

Working intensively with struggling students inspired her to pursue a Master's Degree in Special Education-Mild Interventions and K-12 teaching license in 2010 at the University of Saint Francis, in Fort Wayne, Indiana. After accepting a teaching position at a Title I high school in Fort Wayne, and successfully coordinating a case load of nearly 50 students, she was determined to make a difference in the identification and teaching of students with special needs and education policy. Rhonda investigated the Philosophy Degree in Education, at the University of Tennessee, applied, and was accepted in the fall semester of 2012 concentrating on Literacy Studies. While at UT, she completed her dissertation on the existence of word callers, and researched the role of public library story-tellers in promoting early literacy skills. She also assisted with research projects including comparing the difference between providing students with summer reading versus summer tutoring, and the prevention of maltreatment of children with special needs. She graduated in 2017 hoping to work with schools, teachers and students,

emphasizing literacy development, teacher training, closing the achievement gap, and other social justice issues. Additional interests are teaching students to image as it relates to listening and reading comprehension, and writing.