

AN EXAMINATION OF DIFFERENCES BETWEEN LOW- AND
AVERAGE-ACHIEVING MIDDLE SCHOOL READERS ON A SET OF INFORMAL
READING MEASURES

A Dissertation
by
ALLISON M. WILSON

Submitted to the Graduate School
at Appalachian State University
in partial fulfillment of the requirements for the degree of
DOCTOR OF EDUCATION

August 2016
Educational Leadership Doctoral Program
Reich College of Education

AN EXAMINATION OF DIFFERENCES BETWEEN LOW- AND
AVERAGE-ACHIEVING MIDDLE SCHOOL READERS ON A SET OF INFORMAL
READING MEASURES

A Dissertation
by
ALLISON M. WILSON
August 2016

APPROVED BY:

Darrell Morris, Ed.D.
Chairperson, Dissertation Committee

Woodrow Trathen, Ph.D.
Member, Dissertation Committee

Jennifer McGee, Ed.D.
Member, Dissertation Committee

Audrey Dentith, Ph.D.
Director, Educational Leadership Doctoral Program

Max C. Poole, Ph.D.
Dean, Cratis D. Williams School of Graduate Studies

Copyright by Allison M. Wilson 2016
All Rights Reserved

Abstract

AN EXAMINATION OF DIFFERENCES BETWEEN LOW- AND AVERAGE-ACHIEVING MIDDLE SCHOOL READERS ON A SET OF INFORMAL READING MEASURES

Allison M. Wilson
B.A., Colgate University
M.Ed., University of Washington
Ed.D., Appalachian State University

Dissertation Committee Chairperson: Darrell Morris

This study examined reading-related differences between 82 low- and average-performing middle school students across a battery of informal reading measures (including word recognition, passage-reading, spelling, and vocabulary tasks). In addition, the study explored reading-related differences within the low-performing group of students ($n = 52$). Lastly, the study sought to develop a shortened informal reading assessment that would yield diagnostic information to inform placement and instructional decisions for older struggling readers.

Results showed that there were clear reading-related differences between the low-performing and average-performing groups of students, specifically in the area of print processing (e.g., accuracy and rate). In addition, an analysis of the low-performing group (those scoring between the 15th and 40th percentile on a standardized comprehension test) revealed that they were a heterogeneous lot who demonstrated various reading strengths and weaknesses. The most remarkable differences in the low group were again found to be in the

area of reading rate or fluency. Finally, a shortened informal reading inventory was developed and tested in order to aid middle school teachers in efficiently collecting information about the print-processing needs of their struggling readers.

The major findings in the study validated past research (Dennis, 2013; Hock et al., 2009; Morris et al., 2014), showing that many older readers still struggle with poor print-processing skills (e.g., accuracy and rate) that can inhibit their ability to read and comprehend grade-level materials. The shortened passage-reading assessment described in the study may prove helpful to teachers in assessing and intervening on behalf of their older struggling readers.

Acknowledgments

This dissertation was not a solo project. There were so many people that made the completion of this work possible. First and foremost, I want to thank my dissertation chair, Dr. Darrell Morris, for his guidance and support. Throughout our meetings over the past three years, Dr. Morris has taught me the importance of clarity in speaking, writing, and thinking. He has come to every one of our meetings with patience, thoughtfulness, and logic. Even after forty years of work in the reading field, Dr. Morris greets his work with enthusiasm and energy. I continue to be inspired by his dedication to investing in teachers, and I hope to bring this type of positivity and gusto to all that I do. I also want to thank the other two members of my committee, Dr. Woodrow Trathen and Dr. Jennifer McGee, for their encouragement and continued commitment to make my dissertation better.

This study involved a half year of data collection at Avery County Schools. I want to thank Dr. Burleson and Ellis Ayers for greeting me with open arms and providing me the tools and resources needed to camp out in their middle schools so that I could collect data. Thank you to principals Justin Carver and Ricky Ward for helping me set up an assessment schedule and a place to work with students. Even with the craziness of the Appalachian winter school schedule (students sometimes missed more than two weeks off in a row), the school leaders and teachers welcomed my work with the students, despite any disruption it may have caused. During my time in their schools, I have learned so much from Avery County teachers, who believe that their students can achieve the highest success.

Finally, this work would not have been possible without the enduring love and flexibility from my network of family and friends, near and far. After my husband and I welcomed the addition of our sweet twin boys in November 2015, the idea of completing my dissertation seemed close to impossible. But, here we are, six months later, and I am about to defend this work. I want to thank all of my family members who believed that I could make it to this point. Thanks to my brother Ethan who encouraged this work by reminding me of its importance. Thank you to Nora Vines for listening and advising me when I needed it most. I am forever in debt to my mom, Gail Delfin (“Gigi”) who happily traversed the 1000+ miles from Philadelphia to Boone multiple times a month to care for her grandsons so that I could steal time to work. I want to thank my husband for his unconditional love and support. I cannot imagine a more patient and caring partner than Eben Wilson. Eben continues to amaze me on a daily basis—he juggles cooking dinner, changing diapers, and cleaning the house with a smile, all the while humming a Beatles tune. Finally, thanks to Oren and Heath, our little men, for providing daily entertainment and constantly reminding me of the goodness of life.

Dedication

I dedicate this dissertation to all the middle school teachers who believe in the power of reading and understand what a single book can do to change the life of a student.

Table of Contents

Abstract	iv
Acknowledgments.....	vi
Dedication.....	viii
List of Tables	x
Chapter One: Introduction	1
Chapter Two: Literature Review	7
Chapter Three: Methodology.....	28
Chapter Four: Results	38
Chapter Five: Discussion and Implications	49
References.....	63
Appendix A.....	71
Appendix B	72
Appendix C.....	77
Vita.....	78

List of Tables

Table 1. Participants.....	29
Table 2. Performance Means and Standard Deviations for Low-Performing and Average-Performing Middle School Readers	40
Table 3. Performance by “Simple View” Quadrant (Print Processing, Vocabulary, and Comprehension)	42
Table 4. Performance Means and Standard Deviations for “Low” Print-Processing and “High” Print-Processing Groups	45
Table 5. Performance Means and Standard Deviations for the Full-Length and Shortened IRI	46
Table 6. Cronbach Alpha Reliabilities for the 10-item Word Recognition and Spelling Tasks.....	48

Chapter One: Introduction

As a former public middle school reading teacher and special educator in a low-income school district, I often faced multiple barriers in teaching my struggling students to read. All of my students read significantly below grade level and, as the teacher, I was expected to somehow close the two- to five-year reading gap without basic resources such as adequate assessment and appropriate reading materials.

In graduate school I had been trained to teach the special education curriculum, but during my first teaching job it was apparent that there was no prescribed reading program for my students. Instead, I was supposed to modify a grade-level core curriculum designed for English Language Arts classrooms. I questioned the appropriateness of this seventh-grade level instruction for my struggling special education students, who sometimes read at the third- or fourth-grade level.

Because my students qualified for specialized services in reading, it was essential for me to understand their reading strengths and weaknesses. I tried to make sense of their standardized test performance, but all I could discern was that they were all reading significantly below grade level. These students required differentiated instruction to make progress in reading, but I did not have the descriptive information I needed in order to teach them effectively. Therefore, I decided to administer three *informal* assessments that would help me make instructional decisions: a passage-reading inventory (Qualitative Reading Inventory; Leslie & Caldwell, 2011), a spelling inventory (Words Their Way; Bear,

Invernizzi, Templeton, & Johnston, 2004) and a 1-minute oral reading fluency probe (DIBELS ORF; Good et al., 2011).

After assessing all my students, I found that my four classes, as comprised, would be nearly impossible to teach effectively as a homogeneous group, given the variety of student reading levels within each class. For example, in a given class, one student would be learning how to read basic short vowels (a first-grade skill) while another was reading fluently but with little understanding in fourth-grade text.

With the support of my principal, I rescheduled all of my students into four groups (or classes) using the data I had collected. Although no two readers were exactly alike, I was able to generate basic student profiles from the data and use this information to identify the most appropriate instructional focus for each group.

Using informal assessments to place and teach my special education middle school students was highly successful, and all of my students began to make progress in their areas of need. So, over the next few years I worked with our school leaders to develop a system for assessing, placing, and instructing *all* students who were reading more than one year below grade level (over 50% of the school population). Our school eventually adopted an informal reading inventory so that we could use diagnostic data to inform instruction, for both general education and intervention classes.

In 2013, I decided to go back to graduate school so I could learn more about affecting change from a systems level. As I reflected upon my time in the classroom, I wondered if my own experience teaching struggling middle school readers was similar to that of others. Were the issues that I faced – assessment, placement, and instruction for older struggling readers – prevalent, or had other middle schools developed efficient and effective approaches for

serving this population? This question became a focal point during my doctoral work and eventually led to my dissertation research.

I learned that this situation was not unique to my students. In fact, the 2015 Nation's Report Card (National Assessment of Educational Progress [NAEP], 2015) revealed that two-thirds of our adolescents scored below the *Basic* level, which requires students to locate information in a passage, infer word meanings, identify main idea, and provide supporting information. Moreover, eighth-grade students today are scoring at an even lower level than in 2013 (NAEP, 2013). After a year and a half in graduate school, a recurring question continued to cross my mind: *How should educators best support middle school students who perform below the Basic level of literacy?*

In order for students to be College and Career Ready (as defined by the Common Core Standards), they must be able to read and analyze complex grade-level texts. Using critical thinking skills, they must evaluate text using evidence and make inferences (National Governors Association for Best Practices, 2016). Higher-level skills depend on a basic foundation (accuracy, fluency, vocabulary knowledge) that many struggling readers lack. Nonetheless, current instructional strategies used in middle schools focus on using grade-level materials only. As a result, struggling students have little opportunity to practice reading skills in texts that they can actually “read” (i.e., read with accuracy and reasonable fluency). For these below-grade-level readers, a gap exists between the opportunity for success and the task that is demanded in the classroom.

Several researchers have recently demonstrated that upper-elementary and middle school students who score low on standardized reading comprehension tests are a heterogeneous lot with different skill profiles (Buly & Valencia, 2002; Dennis, 2013). For

example, using informal reading measures and analyzing student performance with a cluster analysis technique, Buly and Valencia (2002) identified six profiles (or types) of struggling readers. These profiles included *automatic word callers* who read fairly fluently but comprehended poorly, *slow word callers* who struggled with both fluency and comprehension, and *slow and steady comprehenders* who read slowly but comprehended fairly well. Although number and type of reader profiles can differ in such cluster analysis studies, this research line clearly shows that older readers who score poorly on standardized tests are not a homogeneous group. Their skill sets differ and these differences need to be assessed.

To address the achievement gap between low- and average-performing middle school students, a valid initial assessment of reading ability is a needed first step. Although an adequate assessment is only the first of several essential steps (including finding quality materials, using appropriate instructional strategies, and developing an effective management scheme), it is the foundation upon which all other steps are built.

A quality reading assessment provides valuable discriminative information about a student's cognitive reading process. The Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990) theory purports that reading comprehension (RC) is the product of two separate, measureable variables: decoding (D) and linguistic comprehension (LC), expressed in an equation ($RC = D \times LC$). For example, if a student cannot decode words efficiently—and thus read at a reasonable rate—then his or her reading comprehension will suffer. Likewise, a student who can decode words quickly and accurately but has limited knowledge of the meaning of key vocabulary may also have trouble deriving meaning from the text. At the middle school level and beyond, students are expected to read (or print

process) quickly and accurately so that they can devote cognitive energy to understanding the text. However, research has shown that many older readers (fourth grade and above) still struggle with the decoding or print-processing side of the Simple View (Buly & Valencia, 2002; Dennis, 2013; Hock et al., 2009). It is thus imperative that print-processing skill as well as comprehension be accurately assessed (Morris et al., 2011).

There is scant published research about reading assessments that address print-processing (word-reading accuracy and reading rate) issues for middle school students. In fact, I found only five studies that include print-processing assessments with older readers (Buly & Valencia, 2002; Dennis, 2013; Hock et al., 2009; Morris et al., 2014; Rupp & Lesaux, 2006). Possibly the most revealing findings in this area were published by Hock et al. (2009). By giving a battery of standardized assessments, Hock et al. found that eighth- and ninth-grade struggling readers (defined as reading below the 40th percentile on an end-of-grade standardized reading comprehension test) scored consistently below their proficient peers across several component reading skills (e.g., word level, fluency, vocabulary, and comprehension). The researchers also reported that greater than 60% of the struggling readers demonstrated difficulty with word-level skills.

The Hock et al. (2009) study is important. It is one of the few that focuses on the reading component skills of older students (eighth- and ninth-graders). The study highlights the fact that struggling readers demonstrate multidimensional reading needs that require careful assessment. However, Hock et al. state that the cumbersome standardized assessments used in their study are difficult to implement in schools. The researchers call for an informal, shortened reading assessment that is effective at diagnosing component skills and can be used by classroom teachers for placement and instructional decisions.

The purpose of the present study was to build on the work of Hock et al. (2009) by examining an informal reading assessment battery—parallel to Hock et al.’s standardized battery—that can be used by middle school teachers. After assessing sixth- and seventh-grade students with a battery of informal reading measures (word recognition, passage reading, and spelling), I compared, as did Hock et al., the scores of a low-performing group with an average-performing group. Then I examined specific reading profiles among the low readers. Finally, I attempted to reduce the length of the various tests in my informal battery so that they might prove to be useful to busy classroom teachers.

Chapter Two: Literature Review

Results in the 2015 Nation's Report Card (National Assessment of Educational Progress [NAEP], 2015) revealed that approximately one-fourth of our adolescents are failing in reading; that is, performing below a basic level of literacy. As of 2015, 24% of eighth-grade students scored below the Basic level in reading. Clearly a call to arms is needed in order to move older struggling readers forward. These students, even if they manage to graduate eventually, will not be prepared to meet the demands of a competitive global economy, nor will they possess the skills needed to participate in a literate society.

Since 2002, there has been a focus on literacy for primary-grade children (K-3), and this effort has yielded promising results (No Child Left Behind, 2002). Nonetheless, many students in grades four and above still lack basic reading ability. By the time these students get to middle school (grades 6-8), foundational reading skills are no longer taught or sometimes even assessed in a meaningful way.

For example, most state departments of education mandate that all students take end-of-grade (EOG) standardized tests to assess reading ability. Many middle schools rely on these standardized assessments to place students into remedial classes despite the fact that the tests provide limited diagnostic information and simply classify students into a category without reference to reading skill (e.g., *Below Basic*, *Basic*, *Proficient*, and *Advanced*). Such a categorical approach for placing students is, to a large degree, ineffective. It fails to capture

the complexity of older struggling readers, providing limited information regarding correct instructional placement or specific reading weaknesses.

A question remains: Why are so many adolescent readers continuing to perform poorly? Meeting the needs of older students is a complex and daunting task. Many struggling adolescent readers lack positive experiences and shy away from tasks that require them to read. Many read two or more years below grade level and may have difficulty recognizing words, reading fluently, or understanding the passage's meaning (Morris et al., 2014). If these students are asked to read difficult, grade-level texts in the classroom, they have little chance to succeed—that is, to improve their reading skill. It is paramount that these struggling readers do not continue to be left behind. In order to help them move forward as readers, we need to better understand the nature of adolescent reading skills (Hock et al., 2009).

Profiles of Older Struggling Readers

In order to better understand the characteristics of older struggling readers, several researchers have focused on upper-elementary and middle school students who scored poorly (e.g., below the 50th percentile) on end-of-grade reading assessments (Buly & Valencia, 2002; Dennis, 2013; Rupp & Lesaux, 2006). These researchers administered a battery of reading-related tasks to students and developed diagnostic profiles that had instructional significance.

Rupp and Lesaux (2006) investigated the relationship between performance on standards-based reading assessments and performance on ten, diagnostic, reading-related measures such as the WRAT-3 word recognition subtest (Wilkinson, 1993), a letter identification task (Lesaux & Siegel, 2003), the Stanford-Binet memory for sentences subtest

(Thorndike, Hagen, & Sattler, 1986), and other tasks that focused on phoneme manipulation. The researchers assessed 1,111 students two times, during kindergarten and then again during fourth grade.

Rupp and Lesaux (2006) conducted a factor analysis to develop profiles of the fourth-grade readers. The authors found their data could be used to separate the students into four distinct profiles of readers. The four profiles included: (a) *low word recognition, low memory* (34%); (b) *low word recognition, high memory* (11%); (c) *high word recognition, low memory* (16%); and (d) *high word recognition, high memory* (39%). In addition, Rupp and Lesaux conducted a multiple univariate analysis of variance (ANOVA) on performance on the standardized reading assessment and the four reader profiles. The researchers found that there was little to no relationship between performance on the standardized EOG test and the profiles derived from the diagnostic measures. Based on their results, Rupp and Lesaux cautioned against the use of high-stakes standardized tests to make decisions about student performance or to guide the instruction of struggling readers.

Buly and Valencia (2002) examined the reading behaviors of fourth-grade students who had scored *below Proficient* on the end-of-grade standardized assessment in Washington state. The researchers administered a battery of reading assessments to 108 students. These assessments aimed to identify an individual's strengths and weaknesses in component skills of reading (e.g., phonological awareness, decoding, fluency, comprehension). Basic phonological awareness and decoding skills were measured by the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1994) and the Woodcock-Johnson Psycho-Educational Battery-Revised (WJ-R; Woodcock & Johnson, 1989). The two tasks from the WJ-R included letter-word identification and word attack. An informal

reading inventory, the Qualitative Reading Inventory-II (QRI; Leslie & Caldwell, 1995), was used to assess fluency (accuracy and rate) as well as comprehension. The final measure, the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997) served to measure a student's receptive vocabulary.

Using factor analysis, Buly and Valencia (2002) identified three variables that influenced reading ability: word identification, meaning, and fluency. Data from these three factors were then analyzed using cluster analysis to create six distinct profiles of students who had failed the state assessment. These profiles demonstrated strengths and weaknesses across reading components. For example, *automatic word callers* (18% of the sample) could identify words quickly and read the passage fluently (with strong accuracy and rate), but had below average comprehension compared to the rest of the sample. Another profile, *slow comprehenders* (24% of the sample) read with below average accuracy and rate, but demonstrated stronger comprehension skills than their peers. Out of the entire sample ($n = 108$), only half of the students demonstrated difficulties with word identification specifically. Yet, when older students fail state assessments, they are often placed into remedial classes that focus solely on decoding skills, often minimizing opportunities to actually read appropriate text (Buly & Valencia, 2002). Buly and Valencia's description of six distinct profiles of struggling readers demonstrates how poorly state assessments capture the complexities of reading ability, and how the use of these data to place and instruct upper-elementary struggling readers is not only misinformed, but also potentially harmful.

Dennis (2013) replicated Buly and Valencia's (2002) study to see if the same profiles of reading behaviors were found in a sample ($n = 94$) of middle school students (grades 6-8) who failed the Tennessee state assessment. The reading tests paralleled those in the former

study with two exceptions. Instead of the WJ-R, Dennis (2013) used the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999) to assess decoding skill. The researcher collected additional information about students' word knowledge by giving the Intermediate Spelling Inventory (ISI; Bear, Invernizzi, Templeton, & Johnston, 2004). Finally, Dennis used the updated QRI-IV (Leslie & Caldwell, 2006) instead of the QRI-II.

As in the Buly and Valencia (2002) study, Dennis (2013) used exploratory factor analysis to identify the main variables that influenced reading ability. Three factors were identified that accounted for 75% of the variance: Meaning (Factor 1) explained 32% of the variance; Decoding (Factor 2) explained 31%; and Rate (Factor 3) explained 11% of the variance. Similar to Buly and Valencia, Dennis used hierarchical cluster analysis to see how the three factors influenced student reading ability. She identified four distinct profiles:

- *Slow and steady comprehenders* were students who easily understood the meaning of the passage but had difficulty decoding nonsense words and reading with a quick rate.
- *Slow word callers* also read slowly, but demonstrated relative strength in decoding skills and orthographic knowledge.
- *Automatic word callers* showed strong decoding skills and read quickly. However, these students struggled to understand the meaning of the passage.
- *Struggling word callers* read quickly but inaccurately, and demonstrated weak decoding skills as well as weak comprehension.

By using a battery of diagnostic reading assessments to identify four different profiles of adolescent struggling readers, Dennis's study echoes the need to use high-quality and descriptive reading measures to inform placement and instruction, even at the middle school level.

In summary, Rupp and Lesaux (2006), Buly and Valencia (2002), and Dennis (2013) all studied readers who performed poorly on end-of-grade state assessments. However, when these students were given a battery of diagnostic reading assessments, many demonstrated relative strengths that were not captured by state assessment levels. The researchers then used their data to create distinct profiles of students based on reading performance. These studies further our limited knowledge about struggling adolescent readers and reaffirm that current standardized tests are not good tools for diagnostic or instructional practice. The work of these researchers highlights a disconnect between using standardized assessments (current school practice) and using traditional informal assessments that are often advocated by experts in the reading field.

The Simple View of Reading

Generating reading profiles from assessment data has been one way researchers have attempted to understand the characteristics of adolescent struggling readers. Examining readers through the perspective of the Simple View of Reading (SVR; Gough & Tunmer, 1986; Hoover & Gough, 1990) is another way. The Simple View posits reading comprehension (RC) to be the product of decoding (D) and linguistic (or listening) comprehension (LC), or $RC = D \times LC$. In this equation, decoding is defined as automatic word recognition, or the ability to quickly and accurately read words. Linguistic comprehension is defined as the ability to receive lexical information by ear and derive meaning at the word or sentence level. Gough and Tunmer (1986) asserted that these two components (print processing and linguistic comprehension) make separate contributions to reading comprehension; they can be described and measured separately. By applying such a

model, it is possible to determine to what degree the individual factors (D and LC) influence a student's reading comprehension.

Inherent in the Simple View is the notion that both decoding and language knowledge are necessary to effectively comprehend text. Gough and Tunmer (1986) stated, "No reasonable proponent of decoding has ever equated decoding and reading, for we recognize that what is decoded must also be understood" (p. 7). Examining the process of learning a foreign language helps to illustrate this principle. A newcomer to the United States may understand English because of his experience watching television. However, if the TV were muted, he could not read the closed captioning. On the other hand, a person learning Hebrew may be able to decode the script but not understand the meaning. These examples illustrate how both components (decoding and linguistic comprehension) are essential to reading. Because the SVR is a multiplicative model, if a person is unable to decode, then he is unable to read. Conversely, if a person cannot understand the words he decodes, he cannot effectively read.

The SVR (Gough & Tunmer, 1986) helps to highlight three important points when examining struggling adolescent readers: (a) print processing and linguistic comprehension are distinct processes that can be assessed and taught separately, (b) effective reading assessments include both components, and (c) effective reading instruction uses the results of these assessments to meet student needs.

Morris et al.'s (2014) Simple View study. Using the Simple View as their framework, Morris et al. (2014) studied fifth and sixth graders ($n = 65$) who scored below the 50th percentile on the North Carolina state end-of-grade (EOG) assessment. The informal reading measures used in this study were similar to those used by Buly and Valencia (2002)

and Dennis (2013). The test battery consisted of a timed word recognition task, an informal reading inventory including both oral and silent passage reading, and the PPVT, a measure of vocabulary (Dunn & Dunn, 2007). After assessing the low readers, the researchers placed the students into one of four *a priori* categories (or profiles) defined by the Simple View. The profiles were: (a) *high print processing, high vocabulary* (HH group); (b) *high print processing, low vocabulary* (HL group); (c) *low print processing, high vocabulary* (LH group); and (d) *low print processing, low vocabulary* (LL group).

The researchers set cut-off scores for print processing and vocabulary in order to determine a student's profile. Importantly, the cut scores were chosen to represent "the lower limit of an 'average grade-level range'" (Morris et al., 2014, p. 9). Although all students assessed scored below the 50th percentile on the EOG, the cut scores indicated if a student was high or low within the given sample.

For print processing, the dual cut scores were 94% for oral reading accuracy and 105 words per minute (wpm) for oral reading rate. Students who scored at or above cut scores on *both* measures were considered high in print processing. Those who did not were designated as low in print processing. To enhance stability of the print-processing scores, the researchers combined the grade-level and one grade-level-below scores on the passage reading assessment.

For vocabulary, one cut score was used. Students achieving at the 40th percentile or above on the PPVT were considered high. Those achieving below this score were considered low in vocabulary.

The application of the cut scores for print processing and vocabulary enabled Morris et al. (2014) to identify four Simple View categories or profiles. The *first profile* (48% of the

sample) was low on both print processing and vocabulary (LL). This LL group demonstrated a slow reading rate, scored below the 20th percentile on the PPVT, and scored between the 20th and 26th percentile on the EOG reading test. The *second profile* (25% of the sample) was high on print processing and low on vocabulary (HL). The HL group demonstrated relative strength in print processing compared to vocabulary knowledge. The HL group could read accurately and quickly, but scored at or below the 25th percentile on the PPVT. These students scored between the 40th and 49th percentile on the EOG. The *third profile* (14% of the sample) was high in both print processing and vocabulary (HH). The HH group read accurately and quickly, scored above the 40th percentile on the PPVT assessment, and between the 43rd and the 51st percentile on the EOG. The *fourth profile* (14% of the sample) was low in print processing and high in vocabulary (LH). The LH group demonstrated relative strength in vocabulary despite low scores in print processing. These students scored low on accuracy and rate, but scored around the 70th percentile on the PPVT and between the 36th to 53rd percentile on the EOG.

The Morris et al. (2014) study, though exploratory in nature, demonstrated that it is possible to examine differences among older struggling readers by using Simple View categories. The study also showed that a reader's strength in one area may compensate for weakness in another area, as demonstrated by the HL group who scored highly on the EOG despite low scores in vocabulary. Finally, the Morris et al. study has implications for instructional practice. If students can print process at grade level, they can be taught with grade-level materials. However, those groups who are not able to process grade-level text with adequate accuracy and rate (LL and LH) should be given the opportunity to read below-grade-level texts. As Allington (2002) has stated, it is hard for students to learn from books

that are too difficult for them to read (or print process). Morris et al. (2014) conclude with three recommendations: (a) assess low readers, (b) provide materials at different grade levels, and (c) differentiate and individualize instruction within the classroom.

Low and Average Readers

The aforementioned studies focused solely on low readers as defined by low scores on state end-of-grade (EOG) reading assessments. Hock et al. (2009) went further and documented key differences between low and average readers as defined by standardized assessments. In their study, Hock et al. assessed eighth- and ninth-grade students ($n = 345$) attending seven schools in the Midwest. Students were selected to participate in the study based on the Kansas Reading Assessment (KRA). The researchers aimed to sample 60 students who scored in each category on the KRA (i.e., unsatisfactory, basic, proficient, advanced, and exemplary). From this sample, they wanted to analyze the component skills (e.g., word level, fluency, vocabulary, and comprehension) of adolescent struggling readers and proficient readers. Hock et al. defined struggling readers as students reading below the 40th percentile on the state test. They chose this cut score because it describes students scoring one third of a standard deviation below the expected mean.

Hock et al. (2009) chose assessments that would provide insight into four component skills, including word level, fluency, vocabulary, and comprehension. All assessments used by Hock et al. were standardized, norm-referenced measures, unlike the previously reviewed studies, which used a mix of standardized and informal measures. Each participant was assessed individually during one 2.5 hour session. For word level, they used two subtests (word identification and word attack) from the Woodcock Language Proficiency Battery-Revised (WLPB-R; Woodcock, 1991). For fluency, the researchers used four different

assessments. They used two subtests (sight word efficiency and phonemic decoding efficiency) from the TOWRE (Torgesen et al., 1999), and two subtests (rate and accuracy) from the Gray Oral Reading Test-4 (GORT-4; Wiederholt & Bryant, 2001). For vocabulary, the researchers measured both oral vocabulary and reading vocabulary. For oral vocabulary, they used the PPVT-III (Dunn & Dunn, 1997). For reading vocabulary, they used the reading vocabulary subtest from the WLPB-R. Finally, they assessed both listening comprehension and reading comprehension. For listening comprehension, the researchers used the listening comprehension subtest from the WLPB-R. For reading comprehension, they used two separate instruments, the passage comprehension subtest (read orally) from the GORT-4 and the passage comprehension subtest (read silently) from the WLPB-R.

To analyze their data, Hock et al. (2009) first conducted a principal-components analysis to confirm that the assessments measured the four different reading components. Next, the researchers divided the sample into struggling readers and proficient readers by using a 40% cut score based on the comprehension composite score. Students scoring above the 40th percentile were labeled as proficient readers ($n = 145$) and those scoring at or below the cut score were labeled as struggling readers ($n = 202$).

Hock et al. (2009) reported three major findings. First, the researchers found that struggling readers scored significantly below proficient readers on each of the four component skills (word level, fluency, vocabulary, and comprehension). Second, they found that the struggling readers scored about one standard deviation below the mean in each component skill, while proficient readers scored above the mean on each component. Third, Hock et al. reported that more than 60% of struggling readers were low in word-level skills.

The Hock et al. (2009) study is important for several reasons. It was carefully conducted using standardized measures; focused on upper middle school students (eighth- and ninth-graders); and showed that, along with fluency, vocabulary, and comprehension deficits, struggling readers also had weaknesses at the word level. The study clearly indicates the need for valid, multi-component reading assessments in the middle school and upper grades in order to understand low readers' relative strengths and weaknesses. However, Hock et al. acknowledged that their standardized assessments were too lengthy and cumbersome for classroom teachers to administer. It is important to remember that each student was assessed using a complicated test battery that took over two hours to administer. The researchers noted that although this type of assessment is doable in a research study, it is nearly impossible to conduct in a school setting. Hock et al. (2009) concluded:

Therefore, there is a pressing need for development and validation of instruments that are efficient for screening, placement, and diagnostic purposes at the secondary level....educators need fewer instruments that require less time to administer and result in a single report providing student results in a form that is easy to interpret and use. (p. 35)

Middle School Reading Assessments

Currently, reading measures used to assess middle school students vary widely. Many, if not most, schools rely on end-of-grade (EOG) assessments to place students into remedial classes (Buly & Valencia, 2002). Often, this standardized test is the *only* tool many middle and secondary schools use. Although the EOG may be effective as a rough screening assessment—that is, to see which students are achieving at grade level and above and which

are not—it is a poor instrument for understanding students’ instructional needs (Torgesen & Miller, 2009).

There is currently a lack of research regarding how reading is, or should be, assessed in the middle grades. Nonetheless, a recent focus on struggling adolescent readers (particularly the Response to Intervention [RTI] initiative) may provide some direction in this area.

Response to Intervention. The goal of RTI is to “ensure that all children have access to high-quality instruction and learning opportunities and that struggling learners are identified, supported, and served early and effectively” (Center on Response to Intervention at American Institutes for Research, 2014, para.1). The components of RTI include: use of academic screening tools to identify struggling students, multi-tiered instruction and intervention, progress monitoring, and data-based decisions. Thus far, research on RTI shows that, when implemented with fidelity, the RTI model can support students of all ability levels (National Center for Response to Intervention [NCRTI], 2010b).

Assessment is one of the major components of RTI. Screening, diagnostic, and progress monitoring tools are all important pieces of this model. Although an effective RTI system will identify and serve many students in the early grades, many older students will still require intervention in reading (Johnson, Pool, & Carter, 2009). In order to identify older struggling readers, an effective assessment system must be in place through the middle and secondary grades.

Screening tools. A universal screening system for all students is at the heart of any comprehensive literacy assessment system (Torgesen & Miller, 2009). Johnson et al. (2009) describe a suggested sequence when implementing a RTI screening process for reading. This

sequence includes four steps: (a) review end-of-year assessment results for all students, (b) identify students who do not meet benchmarks, (c) determine the severity of performance discrepancy, and (d) conduct targeted assessments to inform intervention.

The Center on Response to Intervention (NCRTI, 2010a) compiled a list of tools intended to serve as “screeners” for middle and secondary students. Members of the Center’s Technical Review Committee (TRC) listed ten instruments for screening reading ability in middle and secondary schools. The website (rti4success.org) supplies the name of the tool, the area it assesses (e.g., reading accuracy, fluency, and comprehension), as well as ratings of generalizability, reliability, and predictive validity. The website also gives information about who administers the assessment, how long it takes, and how the test is scored (by computer or with a scoring key).

What follows is a brief description of three of the screening tools for reading recommended by the NCRTI (2010a). These screeners include the Standardized Test for the Assessment of Reading (STAR Reading), Discovery Education Predictive Assessment, and AIMSweb Reading-Curriculum Based Measurement.

STAR Reading. The STAR Reading assessment (Renaissance Learning, 2015) is meant to estimate a student’s understanding of grade-level state standards, predict performance on an end-of-grade state reading test, determine appropriate instructional level, and monitor student progress.

Students (grades 1-12) take this 15-minute computer test individually. The student reads short passages silently and answers multiple-choice comprehension questions. In this computer-adaptive test, the computer adjusts the difficulty of the questions based on the student’s previous responses. After completing the assessment, the student receives a

computer-generated scaled score that is used to report three types of information: national norm scores, instructional reading level, and zone of proximal development (ZPD). The national norm scores include information about grade equivalent and percentile rank; the instructional-level score provides a grade level at which the student can comprehend text with 80% proficiency; and the ZPD information provides a grade-level band in which the student should choose reading books (e.g., Matthew should read books from the 2.6 to 3.7 range).

The STAR assessment receives high ratings from the TRC. According to their review process, it is reliable and has good predictive validity (NCRTI, 2010a). However, a major weakness of the STAR test is that it does not measure a student's print-processing skill. Because each student is reading silently, there is no information collected about oral reading accuracy and rate. This is important because the previously mentioned studies (Dennis, 2013; Hock et. al, 2009; Morris et al., 2014) have reported that low readers often show processing problems at the word or sentence level, problems that would be undetected by tests such as STAR.

Discovery Education Predictive Assessment. The Discovery Education Assessment (ThinkLink Learning, 2006) is a benchmark reading test designed to screen students at risk and predict how well they will score on state reading assessments. Students (grades 1-12) take this 40-minute test individually on computers. The student reads silently a number of passages written at grade level (i.e., a seventh-grade student reads seventh-grade-level passages) and answers 30-40 multiple-choice questions. The number of correct responses is scored automatically.

The student receives two scores: number of test items correct and a Rasch scale score, which estimates reading ability. The score report allows teachers to see which reading standards the student has mastered, as well as those not yet learned.

The Discovery Education Predictive Assessment also receives high ratings from the TRC. They claim that this test, too, is reliable and has predictive validity (NCRTI, 2010a). However, the Discovery Assessment also has major weaknesses. Like the STAR test, this measure does not account for a student's print-processing skill. In addition, it gives no indication of how students may have scored if they were to read texts at other levels (e.g., a seventh-grade student reads a sixth- or fifth-grade-level passage). Giving students, especially struggling readers, texts at lower grade levels would provide valuable diagnostic information about both print-processing and comprehension skill (Barr, Blachowicz, Bates, Katz, & Kaufman, 2007; Morris et al., 2014).

AIMSweb Reading- Curriculum Based Measure (R-CBM). R-CBM is designed to screen, monitor, and report student progress (K-12), as well as identify students at risk of academic failure (Pearson Education, 2012). The R-CBM is administered individually, on paper or computer. The student reads three grade-level passages aloud for one-minute each. As the student reads, the examiner marks the number of word-reading errors. Accuracy and rate for each passage are combined, yielding a words correct per minute (WCPM) score. The median WCPM score for the three passages is used.

The student's accuracy and rate scores are compared to national norms and percentiles. This information can be used to approximate the student's instructional reading level and compare his or her performance to a peer group.

The R-CBM receives moderate ratings from the TRC. It is rated high in reliability, but only moderate in predicting performance on end-of-grade state achievement tests (NCRTI, 2010a). The R-CBM is a quick, efficient measure of grade-level print-processing skill, but it does not measure comprehension skill. Although oral reading assessments, like the R-CBM, are often good predictors of oral reading success for students in kindergarten through third grade (Fuchs, Fuchs, Hosp, & Jenkins, 2001; Wiley & Deno, 2005), older students require additional assessments in the area of comprehension. Baker et al. (2015) reported that oral reading fluency (ORF) alone was limited in predicting whether or not a student will pass an EOG reading comprehension test. However, when an ORF score was paired with a brief reading comprehension measure, the combined measure explained 55-58% of the variance in middle school students' performance on an EOG state test (Baker et al., 2015).

Of these three assessments, two of them, the STAR Reading and Discovery Education Predictive Assessment, are strong at measuring comprehension. The third, AIMSweb Reading, is strong at measuring grade-level print processing. However, none of the three assessments mentioned examines a low reader's print-processing skills in below-grade-level passages.

The Informal Reading Inventory. While the Center for Response to Intervention has reviewed multiple screening instruments, ironically the Center has not reviewed or commented on the oldest of the reading assessments — the Informal Reading Inventory (IRI). First designed by Emmett Betts in 1946, the IRI has long been used in the reading field to informally diagnose reading ability. This inventory is appropriate for older readers because it can be used to diagnose both print-processing and comprehension skills.

The IRI is composed of a series of graded passages (150-250 words in length) that are used to determine a student's instructional reading level (IRL). The assessment begins with the student reading an easy, below-grade-level passage aloud and answering several questions about its content. As the student reads, the examiner marks oral reading errors and records the reader's rate. The student proceeds to read successive passages, each a grade level higher, until he or she reaches a frustration level, at which point the testing is stopped (*Note*. At this point, a series of silent passages is sometimes administered).

Three important scores are derived for each IRI passage read: oral reading accuracy (IRL criterion: 95%); oral reading rate (IRL criterion: varies by grade level); and oral reading comprehension (IRL criterion: 75%). Note that these scores provide measures of both print-processing and comprehension skills.

Multiple versions of the IRI are currently being used in schools, such as the Benchmark Assessment System (BAS; Fountas & Pinnell, 2011), the Next Step Guided Reading Assessment (Richardson & Walther, 2013), and the Qualitative Reading Inventory (QRI; Leslie & Caldwell, 2011). These IRIs are important because, in contrast to the aforementioned screening instruments, they provide a fuller picture of component reading skills at varying reading levels.

IRIs do present problems. First, some examiner training is necessary to ensure reliable, valid results. Second, administration time can be significant, especially when multiple passages are administered (*Note*. This time factor is exacerbated when the IRI [e.g., BAS] includes multiple passages at each grade level). Third, some IRIs (e.g., BAS and Next Step) fail to include reading rate in setting a student's instructional level. By considering only oral reading accuracy and comprehension scores in setting instructional level, these IRIs omit

a crucially important factor in older students' reading skill, that is, rate or fluency (Baker et al., 2015; Morris et al., 2014).

Although the IRI can be effective in diagnosing reading skills and placing an older reader at the appropriate instructional level, the assessment is rarely used in middle schools. Through my participation on multiple professional discussion boards and my reading of journal articles, I have found few middle schools that systematically use IRIs to screen low readers for instructional placement. A reading specialist or special education teacher might use an IRI within his or her classroom for a particular student, but broad use of this assessment appears to be rare, probably for the reasons mentioned previously (teacher training and administration time).

Direct versus indirect measurement of reading skills. A useful distinction when examining assessments is the difference between direct and indirect measurement of a construct (e.g., fluency, comprehension, etc.) (Thorndike, 1971; Ward, Stoker, & Murray-Ward, 1996). Measuring exactly the construct as it happens is a direct measurement, whereas measuring the construct by measuring something else (as a proxy for that construct) is an indirect measurement. Carver (2000) argues that in assessing reading ability, direct measures are more valid than indirect measures. Examples of measuring reading directly include rate and accuracy scores from a passage-reading assessment. Indirect examples of reading assessments include computing scores (or reading level) based on comprehension questions that are answered *after* reading has occurred. The scores are then used as a proxy to infer the quality of the reading behavior, but are not a direct measurement of that behavior. At best, these indirect scores are indicants of comprehension, but only after the process has happened. Print-processing behaviors, on the other hand, are measurable as they occur in real time.

The Simple View emphasizes the importance of assessing print processing, which is the one part of the reading process that can be assessed directly. These distinctions highlight the usefulness of the Simple View as a framework to understand and evaluate the validity of reading assessment instruments.

The Present Study

As indicated in the previous section, there is a pressing need for a multi-component reading assessment for middle school students. Reliance on one-dimensional, categorical test scores does not adequately inform instruction, especially for those students who are reading below grade level.

The profile studies (Buly & Valencia, 2002; Dennis, 2013; Rupp & Lesaux, 2006) showed that when struggling students were assessed with diagnostic reading measures, many demonstrated areas of strength that were not captured by state assessment data. By forming profiles of reader types, these studies highlighted the differences among reading component skills in upper-elementary and middle school readers, differences that are crucial to inform placement and instructional decisions.

Simple View studies (e.g., Morris et al., 2014) have shown that it is possible to use a priori categories to help interpret the reading skill of older struggling readers. These four categories or groups (HH, HL, LH, LL), based on print-processing and oral language knowledge scores, can provide useful information about students' strengths and weaknesses, which can, in turn, inform instructional practice.

Finally, the Hock et al. (2009) study, the impetus for this dissertation, pointed to specific needs in the area of middle school reading assessment. First, using standardized assessments, Hock et al. showed significant differences between low and average middle

school readers across a variety of component skills. Second, the researchers verified distinct differences in component skills in the low-performing group. Third, the researchers called for development of valid, yet easy-to-administer assessments that would be appropriate for classroom use.

In an effort to build on Hock et al.'s (2009) research findings and recommendations, the present study examines the following questions:

1. Are there significant differences between sixth- and seventh-grade low and average readers who are administered a battery of informal, as opposed to standardized, reading/spelling measures?
2. Are there distinct differences in component reading skills among the low-performing students who are administered the informal test battery? If so, do they align with previous research (e.g., Hock et al., 2009; Morris et al., 2014)?
3. Is it possible to shorten the informal reading/spelling battery in order to create a quick, efficient, and teacher-friendly reading assessment for middle school readers?

Chapter Three: Methodology

The major goal of the present study was to better understand the nature of struggling adolescent readers based on their test performance. A secondary goal was to develop a valid, efficient assessment battery that can be used by teachers. The research focused on three major gaps in the literature: (a) understanding if and where significant differences exist between sixth- and seventh-grade low and average readers; (b) identifying distinct differences in component reading skills among the low-performing students; and (c) developing a shortened informal reading battery that is reliable, valid, and teacher-friendly.

Context of the Study

Data used in this study were originally collected by Appalachian State University reading faculty and graduate assistants during Spring 2015 as part of a study that examined older struggling readers (IRB Study #15-0176). Portions of data collected in the original study were used in the current study, including information from the following assessment tasks: word recognition-timed, oral and silent passage reading, spelling, and the Peabody Picture Vocabulary Test (PPVT).

Participants. The sample consists of 82 sixth- and seventh-grade students (39 females, 43 males) who attended two public middle schools located in a rural county in western North Carolina. The total population of sixth- and seventh-grade students in the two middle schools was 326 students (175 females, 151 males). Most of the students in the two schools were Caucasian (89%), followed by Latino (9%), and other (2%). The percentage of

students receiving free and reduced lunch was 58% for School A and 64% for School B. Although these numbers reflect all the students in sixth and seventh grade in the district, the demographic sample used in this study is representative of the population.

Participants were chosen for this study based on their performance on the state end-of-grade (EOG) reading test administered the previous spring. Since one goal of this study was to compare low- and average-performing students, cut scores were drawn to designate two groups. *Low-performing* students scored between the 15th and 40th percentile and *average-performing* students scored between the 55th and 70th percentile. The 40th percentile cut score has been used by a number of researchers to indicate students who are struggling and may be unlikely to progress without support (Hock et al., 2009; Torgesen et al., 2008; Wilson, 2005). A 15-point buffer between low-performing students and average-performing students was used in order to ensure the samples were different. All students falling within these two ranges received a letter from the principal detailing involvement in the study, including an informational letter and a consent form. A total of 100 students were invited to participate. All students who returned consent forms ($n = 82$) participated in the study (see Table 1), including 52 low-performing students (24 sixth graders and 28 seventh graders) and 30 average-performing students (15 sixth graders and 15 seventh graders).

Table 1

Participants

Grade level	Total	Low-performing	Average-performing
6 th grade students	39	24	15
7 th grade students	43	28	15
Total	82	52	30

Assessment tasks. Testing was conducted in February and March, 2015. The informal reading assessments in the study paralleled the formal reading assessments administered in the Hock et al. (2009) study. That is, the middle school students were assessed in the areas of word recognition, fluency, and comprehension. In addition, a spelling measure was used to gain another look at the students' orthographic knowledge.

Overall, participants were assessed with five reading measures. Four of these assessments were administered over two sessions in a one-to-one context. Session one included three measures: word recognition-timed, oral passage reading, and silent passage reading. Session two was devoted to the Peabody Picture Vocabulary Test (PPVT) (Dunn & Dunn, 2007). The two sessions combined lasted approximately 60 minutes. The fifth assessment, a spelling inventory, was administered whole-group within the students' regular classrooms.

One university professor and two graduate assistants conducted research for this study. Each member of the team was familiar with all assessment measures used in the study and had extensive experience administering the IRI (oral and silent passage reading tasks). Before administering the assessment tasks, the team met to clarify any questions about assessment administration. After assessment began, the team met and discussed any issues that arose.

Word Recognition–timed (WR-t). In order for a student to focus on the meaning of a text, he or she must be able to efficiently decode words. This first task, WR-t, measured a student's automatic word recognition, or the ability to read isolated words quickly and accurately (Morris et al., 2011; Perfetti, 2007).

Test instrument. The WR-t task is a series of 20-word lists (see Appendix A) that were developed by randomly sampling grade-level words in *Basic Reading Vocabularies* (Harris & Jacobson, 1982). Each list contains words that increase in difficulty (third grade, fourth grade, fifth grade, etc.). (*Note.* The word recognition-timed task is available for free at www.fcids.org/academic/jac/asureading.) These word lists are graded in difficulty in two areas: word frequency (how often the word appears in text) and orthographic complexity (number of syllables). As established in a previous study, KR-21 coefficients for the WR-t task averaged .86, and stability coefficients for the task were around .86 (see Morris et al., 2011).

Administration. To measure automatic word reading, individual words were flashed on a computer screen for one half second and the student had to read the word without hesitation (Barr et al., 2007). If the student read the word correctly, the examiner flashed the next word on the screen. If he or she was unable to read the flashed word successfully, the word reappeared on the screen and the student was given five seconds to decode it. However, in this study, only responses on the flash presentation were scored. All told, students read four, graded, 20-word lists. Sixth graders read third- to sixth-grade lists, and seventh graders read fourth- to seventh-grade lists.

Scoring. On each list, the student received a score ranging from 0 – 100%. Cut-off scores for interpreting reading level (90% – independent; 70% — instructional; 50% or below — frustration) aligned with previous literature using the word recognition-timed measure (Barr et al., 2007; Stauffer, Abrams, & Pikulski, 1978).

Oral passage reading. Similar to WR-t, oral passage reading focused on a student's ability to print process effectively. When a student reads a passage aloud, it opens a “window”

into the developmental reading process. By listening to a student read passages of varying difficulty, an examiner can gain insight into a student's accuracy, rate, and comprehension while reading in context.

Test instrument. The oral passage-reading task consisted of short passages (see Appendix B) taken from the Morris Informal Reading Inventory (grades three to seven) (Morris, 2015). There are three forms (A, B, and C) of the Morris inventory. The oral passages used in this study were all from Form A. The length of the passages varied from 147 words on the third-grade passage to 254 words on the seventh-grade passage. Mean passage length was 210 words.

Each passage was written in narrative form and most passages were based on historical events (e.g., the Gold Rush and the Lost Colony). A set of six comprehension questions followed each passage, and these questions were passage-dependent and either explicit or implicit in nature.

Morris (2015) used readability formulas to establish the difficulty level of each passage: the Spache formula (Spache, 1953) for the third grade passage, and the New Dale-Chall (Chall & Dale, 2000) formula for the fourth- to seventh-grade passages. The Dale-Chall formula uses sentence complexity and word difficulty to determine readability. Using this formula, Morris reported that both average sentence length and average percentage of difficult words increased from one grade to the next.

Field testing was then used to show whether the different forms of the Morris inventory (A, B, or C) were consistent in difficulty and hierarchical in nature. The results showed that the different forms were approximately equivalent in grade level difficulty. The field testing also confirmed that the passages showed hierarchical ordering; that is, each

grade-level passage was more difficult than the one below (e.g., the fourth-grade passage was more difficult than the third-grade passage and so on) (Morris, 2015).

Administration. All students read Form A passages during the oral reading, which was audio-recorded. Before beginning each passage, the examiner read a one-sentence introduction to the student (e.g., “This story is about a young Native American woman.”). Then the student was asked to read the passage aloud at his or her normal speed. The reading was timed with a stopwatch, and the examiner marked the passage for errors as the student read. After the student finished the passage, the examiner asked six passage-related comprehension questions.

Each student read at least three passages. Sixth graders read fourth-, fifth-, and sixth-grade passages and seventh graders read fifth-, sixth-, and seventh-grade passages. On occasion, some students struggled with below-grade-level passages. They were then asked to read only the first 100 words of subsequent passages, and comprehension questions were not asked. Note that accuracy and rate scores could still be obtained on these 100-word readings. Out of the 82 students assessed, only nine students (11%) read the shortened 100-word passages instead of the full passages.

Scoring. Passage reading was scored for oral reading accuracy, oral reading rate, and oral reading comprehension. Oral reading accuracy (ORA) is the percentage of words read correctly (0 - 100%). Five types of errors were scored as the student read orally: substitutions, insertions, omissions, self-corrections, and teacher helps. To calculate ORA, the number of words read accurately was divided by the total number of words in the passage (e.g., 140/148 to yield an accuracy percentage of 95%).

Oral reading rate (ORR) is the number of words read per minute (wpm). To obtain this score, the number of words in the passage was multiplied by 60 and then divided by the number of seconds it took to read the passage. For example, if a student read a 200-word passage in 90 seconds, his or her ORR would be: $200 \times 60/90 = 133$ wpm.

Oral reading comprehension (ORC) is the percentage of questions answered correctly. This score was obtained by dividing the number of questions answered correctly by the total number of questions asked (e.g., if a student answered 5 out of 6 questions correctly, his ORC would be 83%).

Silent passage reading. Although oral passage reading provides useful information about print-processing skills (oral reading accuracy and rate), most students, from mid-second grade on, begin to read faster silently because they do not have to say each word aloud. By the time students get to middle school, most reading (in and out of school) is conducted silently so that attention can be focused on comprehension. Therefore, it is important to gain information about silent reading rate and comprehension during this third task, silent passage reading.

Test instrument. The Morris IRI (2015) was also used for silent passage reading. However, an alternate form (B) was used so that students were not rereading the same passage they had read orally. For all silent passages except one, Form B was used. Form C was used for the seventh-grade passage. This decision was made intentionally because of the content and word choice of the Form B passage.

Administration. As in oral passage reading, each student read a series of graded passages silently. The passages were similar to those used for oral reading (e.g., sixth grade students read fourth-, fifth-, and sixth-grade passages). The only difference was that if a

struggling reader read only 100 words on any of the oral passages (e.g., grade 6), he or she did not read the corresponding silent passage. This decision was made because if a student could not read and comprehend the oral passage, he or she would likely struggle even more with silent comprehension.

Scoring. Silent reading was scored for silent reading rate (SRR) and silent reading comprehension (SRC) by the examiner. These scores were obtained in the same manner as in the oral passages.

Peabody Picture Vocabulary Test – IV (PPVT). The PPVT was administered to gain information about the students' oral receptive vocabulary, one aspect of linguistic comprehension (see The Simple View). This measure was also administered in the Hock et al. (2009) and Morris et al. (2014) studies.

Test instrument. The Peabody Picture Vocabulary Test-IV (Dunn & Dunn, 2007) is a standardized assessment that was developed to measure receptive vocabulary. The test is untimed and consists of a series of vocabulary items that steadily increase in difficulty.

Administration. This task was administered on the second day of testing. A starting point for the PPVT was established based on the student's age. Then testing began when a basal level was established (8 or more correct responses in a row). During this assessment, the examiner showed the student a card that contained four different pictures. The examiner then pronounced a word (e.g., *buffalo*) and the student had to tell the examiner which picture corresponded to the word. The task got progressively harder over time and ended when a ceiling was reached (eight or more consecutive mistakes were made).

Scoring. To score the PPVT, the examiner counted the number of correct responses below the ceiling. This raw score was then converted to a percentile to indicate where the

child placed within the normal distribution. This assessment has a strong test-retest reliability of .93 (Semel, Wiig, & Secord, 2003).

Spelling. Research shows that spelling and word recognition are highly correlated across grade levels (Swanson, Trainin, Necochea, & Hammill, 2003; Zutell & Rasinski, 1989). According to Perfetti (1992), as students learn to read, they develop mental representations of words, and these representations can be measured by spelling. Therefore, this final task provided an alternative view of a student's orthographic or word knowledge through a spelling inventory.

Test instrument. All students received a spelling task (see Appendix C) called the Qualitative Inventory of Word Knowledge (QIWK; Schlagal, 1992). The QIWK was designed to establish a student's developmental spelling level (e.g., third grade, fourth grade, fifth grade, and so on). Schlagal developed this assessment by sampling grade-level words in the Houghton-Mifflin Spelling Program (Henderson, Templeton, Coulter, & Thomas, 1990). The QIWK has been used in previous studies to measure student's orthographic knowledge (Morris et al., 2011; Rasinski & Zutell, 1996).

Administration. Each student was asked to spell 20 words at each grade level, third through sixth grade. The spelling assessment was administered whole-group by the language arts teacher, beginning with the third-grade list. The teacher pronounced the word, read the word in a sentence, and then repeated the word one final time. Each student wrote the word on his or his individual sheet. Every student completed the spelling of all four lists.

Scoring. On each spelling list, the student received a score ranging from 0 – 100%. This score was calculated by dividing the number of correct responses by the total number of

spelling words (e.g., if a student spelled 7 of 20 words correctly, he obtained a score of 35% on the list).

Design

There were four groups of participants in this study: low-performing sixth-grade students, low-performing seventh-grade students, average-performing sixth-grade students, and average-performing seventh-grade students. Because of the small sample size, I collapsed the low-performing and average-performing students across grade levels. This yielded two groups: a low-performing group (sixth and seventh graders) and an average-performing group (sixth and seventh graders).

The first research question examined the performance of the low middle school readers ($n = 52$) compared to that of their average-achieving peers ($n = 30$). The second research question examined possible performance differences within the low-performing group. Descriptive statistics and t -tests, where appropriate, were used to address this question.

A third research question explored the viability of shortening these assessments (oral passage reading, word recognition, and spelling measures) in order to make them more efficient in terms of administration time. To this end, for all 82 students, I compared IRI accuracy and rate scores on the first 100 words of the grade-level passage with the same scores on the full-length version (230-250 words). Regarding the word recognition and spelling assessments, in each case I employed an item analysis procedure (Cronbach Alpha) to see whether the 20-word lists could be successfully reduced to 10-word lists.

Chapter 4: Results

The present study built on the work of Hock et al. (2009) by using informal measures to test the reading ability of low- and average-performing middle school students. The study also examined the viability of shortened forms of the assessments. The present study posed three research questions:

1. Are there significant reading-related differences between low and average middle school readers?
2. Are there distinct differences in component reading skills among the low-performing students (sixth and seventh graders) on the informal reading battery, and do they align with previous research (e.g., Hock et al., 2009; Morris et al., 2014)?
3. Can the informal reading/spelling battery be shortened in order to create a quick, efficient, and teacher-friendly reading assessment for middle school readers?

Data Analysis

Question 1. The first research question asked if significant reading-related differences exist between sixth- and seventh-grade low-performing students and their average-performing peers. To address this question, I created variables in the dataset for each component of the informal reading battery (Word Recognition-Timed [WR-t], Oral Reading Accuracy [ORA], Oral Reading Rate [ORR], Oral Reading Comprehension [ORC], Silent

Reading Rate [SRR], Silent Reading Comprehension [SRC], and Spelling) and also for two standardized measures (a vocabulary test [PPVT] and a silent reading comprehension test [EOG]). In all, there were nine distinct variables.

Because of the small sample size, the sixth- and seventh-grade students were combined to produce a low-performing group ($n = 52$) and an average-performing group ($n = 30$). These groups were then compared across the nine variables using a *t*-test analysis. *T*-test results (see Table 2) showed that the means for seven of the nine variables were significantly different between groups: Word Recognition-timed, Oral Reading Accuracy, Oral Reading Rate, Silent Reading Rate, Spelling, PPVT, and EOG. Regarding the students' performance on the IRI, all of the print-processing measures (WR-t, ORA, ORR, and SRR) were statistically significantly different. Only the IRI comprehension measures (ORC and SRC) were not statistically significantly different. Note, however, that there were large differences between the two groups (low and average readers) on the standardized comprehension measure (EOG) (35th vs. 67th percentile) and the vocabulary measure (PPVT) (40th vs. 62nd percentile).

Regarding the print-processing differences between the low and average readers, the WR-t and ORA differences (8% and 2%, respectively), though statistically significant, do not put the low group at a severe disadvantage. However, the reading rate differences (ORR and SRR) between the two groups are large and of educational consequence. The low group read orally 24% slower and silently 18% slower than the average group. Reading rate (or fluency) appears to be the major print-processing problem for the low readers, a problem that certainly affects their ability to complete reading assignments and also might hinder their reading comprehension.

Table 2

Performance Means and Standard Deviations for Low-Performing and Average-Performing Middle School Readers

Variable	Low-Performing (<i>n</i> =52)		Average-Performing (<i>n</i> =30)		<i>t</i> -test
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Word Rec-Timed (%)	70	12.9	78	12.2	.011*
Oral Reading Accuracy(%)	95	2.2	97	2.2	.003*
Oral Reading Rate (wpm)	112	20.2	147	19.1	.000*
Oral Reading Comp (%)	66	22.2	70	22.4	.531
Silent Reading Rate (wpm)	131	37.9	160	41.9	.003*
Silent Reading Comp (%)	48	29.8	54	26.5	.421
Spelling (%)	55	22.1	72	18.5	.000*
PPVT (%ile)	40	29.6	62	19.8	.000*
EOG 2015 (%ile)	35	16.2	67	15.3	.000*

Note. * $p < .05$. Spelling mean is the average of the fourth-grade, fifth-grade, and sixth-grade level scores.

Question 2. The second research question asked if distinct reading-related differences exist within the low-performing group of middle school readers ($n = 52$). This question builds on the work of Hock et al. (2009) and Morris et al. (2014), both of whom examined reading behaviors of students who scored below the 50th percentile on a standardized comprehension measure.

Reading skills according to the Simple View. To answer this question, the data for the low readers were divided into four quadrants using cut scores set a priori (see Morris et al., 2014). There were two cut scores for print processing and one for vocabulary. The idea was that students who scored below the cut scores might be at risk in terms of reading grade-level material.

Print processing had dual cut scores, one for Oral Reading Accuracy (ORA) (94%) and one for Oral Reading Rate (ORR) (115 wpm). The ORA cut score was assigned because 95% accuracy is generally viewed as an indication that students can read (or print process) at that level with minimal support (Barr et al., 2007; McKenna & Stahl, 2003). Using 94% accuracy as the cut score instead of 95% (see Morris et al., 2014) allowed some flexibility. Although there is little research about grade-level reading-rate minimums for older students, a cut score of 115 wpm was chosen. In studying fifth- and sixth- grade readers, Morris et al. had used a rate cut score of 105 wpm. (This score approximated the 30th percentile in previous studies [e.g., Hasbrouck & Tindal, 2006; Morris et al., 2011].) Because the students in the present study were sixth and seventh graders, an ORR score of 115 wpm seemed appropriate. In summary, to be high (H) in print processing, a student had to read with 94% accuracy and at 115 wpm or higher. Otherwise, he or she was considered low (L).

The cut score for *vocabulary* was the 40th percentile on the PPVT (same as Morris et al., 2014). A student who scored at the 40th percentile or above on vocabulary was considered high (H); a student who scored below this cut-off was considered low (L).

Students ended up in one of four quadrants: (a) High Print Processing, High Language (HH); (b) High Print Processing, Low Language (HL); (c) Low Print Processing, High Language (LH); or (d) Low Print Processing/Low Language (LL) (see Table 3).

Table 3

Performance by “Simple View” Quadrant (Print Processing, Vocabulary, and Comprehension)

Grade	ORA <i>M SD</i>	ORR <i>M SD</i>	WR-t <i>M SD</i>	PPVT <i>%ile</i>	EOG <i>%ile</i>	Grade	ORA <i>M SD</i>	ORR <i>M SD</i>	WR-t <i>M SD</i>	PPVT <i>%ile</i>	EOG <i>%ile</i>
<u>High print processing/Low language (H/L)</u>						<u>High print processing/High language (H/H)</u>					
Sixth (<i>n</i> =2)	98 (0.7)	147 (4)	80 (0)	36	46	Sixth (<i>n</i> =3)	97 (1.0)	134 (5)	80 (13)	52	35
Seventh (<i>n</i> =4)	98 (1.4)	143 (26)	78 (14)	23	49	Seventh (<i>n</i> =9)	96 (1.9)	126 (14)	63 (16)	47	43
<u>Low print processing/Low language (L/L)</u>						<u>Low print processing/High language (L/H)</u>					
Sixth (<i>n</i> =14)	94 (2.2)	104 (12)	72 (14)	30	27	Sixth (<i>n</i> =5)	94 (1.3)	90 (11)	63 (10)	67	32
Seventh (<i>n</i> =11)	94 (2.7)	99 (14)	66 (10)	19	20	Seventh (<i>n</i> =4)	94 (1.4)	96 (2.7)	65 (11)	53	40

Note. ORA = Oral reading accuracy (%); ORR = Oral reading rate (wpm); WR-t = Word recognition-timed (%); PPVT= Peabody Picture Vocabulary Test; EOG = North Carolina End-of-Grade Reading Test.

Similar to the Morris et al. (2014) study, the largest number of low readers ($n = 25$) fell into the LL group. These students read slowly (less than 105 wpm). They also scored poorly on the PPVT (sixth graders = 30th percentile; seventh graders = 19th percentile), as well as on the EOG standardized test that was given later in the school year (all scored below the 27th percentile).

The HH quadrant held the second highest number of students ($n = 12$). These students could read accurately (above 96%) and with more speed (above 126 wpm); they also scored around the 50th percentile on the PPVT and around the 40th percentile on the EOG test.

Nine students fell into the LH group quadrant. These students scored at 94% accuracy, but struggled with rate (96 wpm or less). The LH group demonstrated strength with language skills, scoring above the 50th percentile on the PPVT. However, their EOG comprehension scores were lower; sixth graders = 32nd percentile; seventh graders = 40th percentile.

The final quadrant, HL, contained only six students. These students showed strong print-processing skills, with high accuracy (98%) and reading speed (143 wpm); however, they showed weak language skills (sixth graders = 36th percentile; seventh graders = 23rd percentile). Somewhat surprisingly, the HL group, despite low vocabulary performance, scored near the 50th percentile on the EOG (sixth graders = 46th percentile; seventh graders = 49th percentile).

In general, these results replicate the findings of Morris et al. (2014), who found students within each of the four quadrants, with the highest concentration of students in the LL quadrant. However, students in the present study tended to score higher on most measures (e.g., ORA, ORR, WR-t, and PPVT) than did those in the Morris et al. study. This was probably due to the fact that Morris et al.'s sample included *all* students who scored below

the 50th percentile on a standardized reading test. The sampling in the present study (15th to 40th percentile) eliminated the very lowest readers (i.e., 0 to 14th percentile), thereby producing higher average scores in the low-reading group.

Although the quadrant results shown in Table 3 are suggestive and parallel results reported by Hock et al. (2009) and Morris et al. (2014), the small sample sizes in three of the four quadrants (HL = 6; HH = 12; and LH = 9) present problems regarding interpretation and generalizability. To address this issue, I performed an additional analysis.

Print processing by group (low and high). Based on the low reader data in Table 3, I combined groups (or quadrants) based on print-processing skills. These combinations produced a low group (LL + LH = 34) and a high group (HL + HH = 18). (*Note.* The terms “low” and “high” in this context are relative, because all 52 students had originally scored between the 15th and 40th percentile on a standardized reading test.)

Table 4 shows the print-processing performance of the low and high groups across three measures: Word Recognition- Timed (WR-t), Oral Reading Accuracy (ORA), and Oral Reading Rate (ORR). *T*-test results showed significant differences between the two groups on ORA and ORR, but not on WR-t. In other words, the low and high groups did not differ in reading isolated words (72% vs. 70%) but they did differ on two contextual reading measures: accuracy (94% to 97%), and rate (100 wpm vs. 134 wpm). On first look, the ORA differences do not appear large. But consider a student reading a sixth-grade text with 200 words on the page. Reading with 97% accuracy means six word-reading errors on the page; on the other hand, reading with 94% accuracy means twelve misread words on a single page, a performance that could negatively affect rate or comprehension.

Table 4

Performance Means and Standard Deviations for “Low” Print-Processing and “High” Print-Processing Groups

Variable	Low Print-Processing (<i>n</i> =34)		High Print-Processing (<i>n</i> =18)		<i>t</i> -test
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Word Rec- Timed (%)	72	13.2	70	14.2	.627
Oral Reading Accuracy(%)	94	2.1	97	1.2	.000*
Oral Reading Rate (wpm)	100	12.4	134	11.4	.000*

Note. **p* < .05

The Oral Reading Accuracy difference notwithstanding, the most notable difference between the low and high print-processing groups involved reading rate or fluency (ORR). The high group read orally 34 wpm (or 25%) faster than the low group.

Question 3. The final research question asked if it was possible to create a shortened assessment that would yield similar results to the full-length informal reading battery. Four measures (Oral Reading Accuracy, Oral Reading Rate, Word Recognition-timed, and Spelling) were examined.

Oral reading accuracy (ORA) and Oral reading rate (ORR). First, each student’s (*n* =81) IRI was re-examined. Accuracy and rate scores were calculated using only the first 100 words of each passage (full-length passages ranged from 230-250 words). These scores were obtained by re-listening to the audio-recordings of the students reading. This analysis yielded two new variables: Oral Reading Accuracy-Shortened (ORA-s) and Oral Reading Rate-Shortened (ORR-s). To calculate the ORA-s score, the number of errors a student made was

subtracted from 100 (e.g., if a student make five errors, the ORA-s score would be 95%, $100 - 5 = 95$). ORR-s was calculated by obtaining a wpm score for the first 100 words.

Regarding the accuracy scores (Table 5), the means of the full-length and shortened assessments were the same. Both measures (ORA and ORA-s), regardless of passage length, yielded a mean accuracy score of 96%. Regarding the rate scores, the means of the full-length and shortened assessments were similar (125wpm vs. 132wpm). With only a seven-point wpm difference between the two rate measures, the shortened assessment appears to be an appropriate substitute for the longer assessment. (*Note.* *T*-tests were not run on the ORA vs. ORA-s means or the ORR vs. ORR-s means because the mean scores were derived from the same performance sample, not independent samples.)

Table 5

Performance Means and Standard Deviations for the Full-Length and Shortened IRI

Measure	<i>n</i>	<i>M</i>	<i>SD</i>
ORA (%)	81	96	2.4
ORA-s (%)	81	96	2.4
ORR (wpm)	81	125	25.9
ORR-s (wpm)	73	132	29.4

Note. ORR-s has an *n* of 73 because 5 of the oral reading passages were not recorded and 3 more were not clear enough to get an accurate rate score when replayed.

Next, I examined correlations between student performance on the full-length and shortened versions of the IRI. Results showed that both shortened measures, ORA-s and ORR-s, were strongly correlated with their full-length counterparts: for example, ORA vs. ORA-s ($r = .77$); ORR vs. ORR-s ($r = .91$).

Finally, linear and stepwise regressions were run to see if IRI performance (accuracy and rate) predicted later performance on an end-of-grade reading comprehension test (EOG 2015). Results showed that the full-length and shortened IRI rate scores were equally good predictors of EOG performance, with both accounting for 32% of the variance on the end-of-grade standardized reading test. However, accuracy scores on the IRI (both full-length and shortened) were not significant predictors of EOG reading performance, accounting for less than 10% of the variance (ORA = 9%; ORA-s = 6%).

Word Recognition-timed and spelling. Two more measures, Word Recognition-timed (WR-t) and Spelling, were also examined to see if they could be shortened. Both of these measures involved grade-level lists, each containing 20 words. Item analysis was used to see whether shortened 10-word lists could predict, or reliably represent, the full 20-word list.

Table 6 shows the Cronbach alpha reliabilities for the reduced 10-item Word Recognition and Spelling lists. For *word recognition*, the reliabilities for the sixth-grade students' performance across the grade levels, 3-6, ranged from .67 to .79. These reliability coefficients are somewhat lower than the hoped for .80 (J. Perney, personal communication, February 28, 2016). The reliabilities for the seventh-grade students' performance across the grade levels, 4-7, are even lower (.47 to .73).

For *spelling*, the reliabilities for the sixth-graders' performance, grades 3-6, ranged from .84 to .89. The reliabilities for the seventh-grade students' performance, grades 4-7, also met the .80 criterion (range = .81 to .83).

Table 6

Cronbach Alpha Reliabilities for the 10-Item Word Recognition and Spelling Tasks

Grade	Word Recognition					Spelling			
	3	4	5	6	7	3	4	5	6
6	.73	.79	.75	.67		.86	.84	.87	.89
7		.47	.73	.68	.49		.81	.83	.83

Given these results, it appears that the 10-item spelling lists, at grade level and below, are a more reliable measure of orthographic knowledge than are the 10-item word recognition lists. At the least, it can be said that the shortened spelling lists do a better job in representing the full-length spelling lists, than do the shortened word recognition lists in representing the full-length word recognition lists.

Chapter Five: Discussion and Implications

The purpose of the study was to build on the work of Hock et al. (2009) and Morris et al. (2014) by examining reading behaviors of struggling middle school students. The study had three major goals: (a) explore reading-related differences between low- and average-performing students, (b) examine reading profiles among the low group of students, and (c) develop a short diagnostic reading assessment that can be used by middle school language arts teachers. Eighty-two students (Grades 6 and 7) were administered a battery of informal reading assessments in order to answer these questions. In this section, I discuss the major findings of each question as well as implications for using such assessments with older students.

Major Findings

Question 1. Are there significant reading-related differences between sixth- and seventh-grade low-performing readers and their average-performing peers? Results yielded two major findings. First, there were significant differences between the two groups across print-processing abilities (including word recognition, oral reading accuracy, and oral reading rate) as well as on a spelling measure and a vocabulary measure. Second, print-processing skill, specifically oral reading rate, was the most important factor differentiating the two groups.

Results of the present study validate and extend past research documenting key differences between low- and average- performing older readers (Hock et al., 2009). Similar

to Hock et al. (2009), the present study found that older struggling readers, in this case, sixth and seventh graders, performed below an average group in two areas:

(a) meaning vocabulary, and (b) accurately and fluently reading the words and sentences on the page.

It was expected that the low reader group (15th to 40th percentile on a standardized reading comprehension test) would show deficits in meaning vocabulary. Their print-processing results, however, were more varied and in some areas surprising. For example, the low-achieving group performed significantly below the average-achieving group when reading isolated words across two measures. On the Word Recognition-timed assessment, a list of grade-level words, the low-performing group read words with less accuracy (70% vs. 78%). On the Spelling assessment, the low-performing group (55% correct) also scored significantly below their average-achieving peers (72% correct), a measure of orthographic knowledge. When it came to reading words in context, the low reader group scored below the average reader group on two measures. In Oral Reading Accuracy, the low-performing group read less accurately on grade-level passages than their average-achieving peers (95% vs. 97%, respectively). They also scored lower on Oral Reading Rate, reading much slower than their peers (112 wpm vs. 147 wpm, respectively).

It was in the area of Oral Reading Rate (ORR) that the low-performing students showed clear deficits. When compared to their average-performing peers, low-performing students read *24% slower orally* (112 wpm vs. 147 wpm) and *18% slower silently* (131 wpm vs. 160 wpm). This finding parallels past research (Hock et al., 2009; Morris et al., 2014) and has major implications for instruction and intervention. For example, a student who struggles with fluency or rate, yet is placed into an intervention program that focuses on decoding

individual words, will not learn to read more quickly because he or she is not getting practice reading text. A student must practice reading interesting material at the correct instructional level if reading rate is to improve (Allington, 2002).

Print-processing data of the kind mentioned above is not revealed on EOG state reading assessments, which afford only categorical (e.g., Levels 1-5) or percentile information. However, in order to read and comprehend grade-level text, middle school students must be able to process the words and sentences on the page in an efficient manner (Perfetti, 1985; Samuels, 2006). Accuracy and rate data can only be provided through individual informal reading assessments.

Question 2. Within the low-performing group ($n = 52$) in this study, were there reading-related differences among the students? To answer this question, I examined possible differences in two ways. First, I used the Simple View model to place students into one of four quadrants set a priori: High Print-Processing, High Vocabulary (HH); High Print-Processing, Low Vocabulary (HL); Low Print-Processing, High Vocabulary (LH); or Low Print-Processing, Low Vocabulary (LL). Similar to the Morris et al. (2014) study, I found that low readers, when examined through the Simple View lens, were indeed a heterogeneous group with different reading strengths and weaknesses (e.g., reading accuracy, reading rate, and vocabulary). Regarding print processing and vocabulary, 71% of the students were either relatively low in both areas (LL; $n = 25$) or relatively high in both areas (HH; $n = 12$). On the other hand, 29% of the students were low in one area but high in the other (either HL or LH).

Next, I analyzed the low-group data by print-processing level. Because of the small sample ($n = 52$), I collapsed the four quadrants into two quadrants: low print-processing (LL and LH; $n = 34$) and high print-processing (HH and HL; $n = 18$). Even within the low-

performing group (again, $n = 52$), some students read significantly slower than other students also categorized as low readers. That is, “low” print-processors ($n = 34$) demonstrated significantly lower reading rates than the “high” print-processors ($n = 18$). These students read at an average of 100 wpm, a rate that is appropriate *in third grade* (Morris et al., 2011). However, the high print-processing group read orally at 134 wpm, an acceptable rate in sixth and seventh grade. (*Note.* The designations, “low” and “high” are relative in this case. All students were part of the low-performing group and had scored below the 40th percentile on the EOG.)

Such a difference in rate is alarming for a few reasons. First, low print processors will not be able to catch up to their peers if they continue to read only grade-level texts. Since they read at a much slower pace, they cannot read the same volume of pages as their classmates, even some peers who are also categorized as low-performing. Second, current assessments used in middle schools do not use rate as a factor in determining reading level. (*Note.* See Chapter 2 for a sample of current assessments used in middle schools.)

Determining a student’s reading level according to accuracy and comprehension scores *only* is problematic and does not give a full picture of the student’s reading proficiency. In addition, if assessments that lack rate information are used for intervention placement, the student may end up receiving instruction that is misaligned with his or her needs.

Question 3. Is it possible to develop a shortened reading assessment that can be used by middle school teachers to screen and place struggling readers? In their study, Hock et al. (2009) called for a quick, teacher-friendly assessment that yields useful diagnostic information about older struggling readers. Results in the present study indicate that a shortened 100-word version of the IRI passages, focusing on accuracy and rate, can be

substituted for the full-length version (approximately 250 words). In addition, a shortened spelling measure can be used as a quick and reliable screener to assess a student's word knowledge.

The present study showed that student performance on the first 100 words of the IRI was significantly related to their performance on the entire assessment. Across the whole sample ($n = 81$), both assessments yielded similar means for oral reading accuracy (96% for both) and oral reading rate (full-length mean = 125 wpm; shortened mean = 132 wpm). The seven-word rate difference between the full-length and shortened versions of the IRI, although small, can be easily explained. While the mean for oral reading rate for the full-length assessment was calculated based on the entire sample ($n = 81$), the shortened version included fewer participants ($n = 73$) because of procedural issues (i.e., five of the oral reading passages were not recorded and three more were not clear enough to get an accurate rate score when replayed). The mean rate score of the eight students not included in the analysis was 118 words per minute. This rate score is lower than the mean of both the full-length and shortened assessments (125 wpm and 132 wpm, respectively). Had these eight students been part of the analysis—had I been able to obtain an ORR-s score them—the mean rate scores of the full-length and shortened assessments would have been closer.

In addition, the rate component on both measures (full-length and shortened) was equally powerful in predicting later performance on an end-of-grade (EOG) standardized reading comprehension test. Although oral reading accuracy was not shown to be a significant predictor of the EOG, oral reading rate on both measures predicted 32% of the variance. This result highlights the importance of including rate as a measure of student

reading ability. As discussed in question two, current middle school assessments do not account for rate when calculating a student's reading level.

In summary, performance on the shortened version of the IRI was closely related to performance on the full-length version. In addition, the rate measure on both versions predicted performance on a later standardized comprehension measure. Because of these two factors, I conclude that the 100-word IRI is a valid substitute for its full-length counterpart and can yield reliable and valid print-processing (accuracy and rate) information about older readers.

The present study also examined Word Recognition- timed (WR-t) and Spelling lists to see if they could be shortened from 20 to 10 words. While the 10-word WR-t lists showed insufficient reliability coefficients (below the .80 threshold), the 10-word spelling list met this reliability criterion (i.e., .80). This finding suggests that a 10-word spelling list may be used to approximate a middle school student's orthographic knowledge. Such an assessment can serve as a screener to quickly and efficiently gather valuable information about a student's print-processing level.

Implications

Results in the present study have implications for assessing the reading ability of students in middle school, a neglected population for reading research. The findings showed that the profiles of low readers (15th to 40th percentile on a standardized test administered the previous spring) were different from the profiles of their average-achieving peers (55th to 70th percentile). Moreover, within the low or struggling reader group, there were significant print-processing differences. These results parallel findings in previous research (Dennis, 2013; Hock et al., 2009; Morris et al., 2014) and further the call for quality, needs-based reading

assessment (and instruction) in the middle grades. In this section, I first offer practical suggestions about how middle school teachers can use diagnostic assessments in their classrooms, and then I discuss how teachers can use data from these assessments to make instructional decisions.

How can diagnostic assessments be used effectively in the middle school and beyond? Restructuring reading instruction and intervention at the middle school level is no small feat. The first challenge deals with assessment. Middle school readers are a heterogeneous group, possessing a wide variety of reading strengths and weaknesses. However, typical assessments used in middle school (e.g., EOG) are unable to capture the complexity of struggling adolescent readers. Nonetheless, schools often use this imprecise data to place students scoring below a certain threshold into intervention programs. A major shift in assessment practice must occur if struggling readers are to get the differentiated instruction they truly need. This will entail a move from reliance on EOG data to the use of more sensitive diagnostic assessments (e.g., IRI).

However, such a shift does not come without complications. Middle school teachers are often responsible for planning and teaching five or more instructional periods, which means that the teacher sometimes interacts with 80 or more students a day. Without outside help, individually assessing all 80 students would be a major undertaking. Therefore, teachers need a way to effectively screen students in order to know which students to assess first. Data from the present study suggest that teachers may be able to use a 10-word, grade-level spelling assessment as a screener. This short spelling test can be given to the entire class. If students who scored low on the EOG (e.g., below the 40% percentile) also score below a certain threshold on the spelling test (e.g., below 50%), then these students should be

given the shortened IRI. Thus, a teacher may end up having to assess 20 students instead of all 80. Because the shortened IRI only takes approximately five minutes per student to administer, a teacher may have to dedicate only two planning periods in order to obtain valuable diagnostic information about his or her most at-risk students. If the teacher gives this assessment three times per year (beginning, middle, and end), then the resulting data can also be used to guide instruction over time.

In order to administer and interpret the shortened reading assessment effectively, middle school teachers will require specific training. Such training might include two or three afterschool workshops where teachers learn how to score and interpret performance on the IRI passages. First, teachers would learn how to code and score five types of oral reading errors, including substitutions, omissions, insertions, teacher helps, and self-corrections. Then, they would learn how to put all information from the testing on a coversheet and use this coversheet to interpret a student's reading strengths and weaknesses. For example, an IRI coversheet might indicate that a student scored poorly when reading a sixth-grade passage (e.g., 92% accuracy and 91 wpm), but read with appropriate accuracy and speed on a fifth-grade passage (e.g., 95% accuracy and 130 wpm).

Part of the teacher training would include practice examining such data in order to make meaningful instructional decisions. In the given example, the teacher might choose to allow the student to practice reading fifth-grade texts during independent reading, or choose easy sixth-grade texts when teaching the student in a guided reading group. In another example, a teacher may notice that a student demonstrates a gap between accuracy and rate on the cover sheet (e.g., reads with 97% accuracy but only at 94 wpm on a grade-level

passage). Accordingly, the teacher may decide to supplement this student's instruction with fluency-building activities (e.g., repeated readings of some kind).

At present, even when schools do use diagnostic assessments, they often determine a student's reading level by using accuracy and comprehension scores only. However, accuracy scores do not always yield discriminative information. In the present study, the at-risk group, who had scored low on the EOG, still achieved a 95% accuracy score when reading a grade-level passage. Therefore, word-reading accuracy, by itself, is not sufficient in diagnosing reading level. Oral reading rate, which measures the fluency of a student's print processing, must be considered as well. This conclusion is supported by other research as well (Morris et al., 2011). The inclusion of reading rate in the assessment is important in making instructional and intervention decisions. For example, if a poor reading rate (lack of fluency) is the major contributor to a student's comprehension problem, then an intervention focusing on comprehension strategy instruction alone will not get at the underlying problem (i.e., rate). Middle school teachers can best serve their students when they are able to understand these nuances about student reading skill.

How can these assessments be used to guide instruction and intervention? Even with the assessment issue resolved, a second challenge remains—instruction. Typically, middle schools offer only a single reading intervention program, and students are usually placed into this intervention by their score on a standardized end-of-grade measure. This type of “one-size-fits-all” model does not meet the needs of all struggling readers. Many middle school intervention programs focus on either decoding skills (e.g., System 44; Houghton Mifflin Harcourt, 2013) or comprehension skills (e.g., Boldprint; Houghton Mifflin Harcourt, 2006). Such skill-based programs oftentimes give students limited access to texts that they

can actually read. When students cannot read with acceptable accuracy and rate, it is likely that their comprehension will suffer (Perfetti, 1985; Samuels, 2006). Students who are continually asked to read texts that are too hard for them may also suffer emotional and behavioral consequences as well.

In order to support students in both narrative and content materials, it is paramount that middle school teachers begin to rethink instruction for struggling readers. To accomplish this task, teachers must first acquire appropriate reading materials. Struggling readers need access to texts that are authentic, engaging, and at their instructional level. Such texts should include multiple genres (poems, short stories, novels, informational books, articles, etc.) and be at various grade levels in difficulty. Accumulating such a library can be time-consuming and potentially expensive, but free online resources (e.g., Newsela, n.d.) and teacher-friendly donation programs (e.g., Donorschoose, n.d.) can allow teachers to obtain materials at a low cost so that they can build a rich multilevel library over time. Adopting instructional-level materials in the general education classroom is a clear shift from current middle school practice, which focuses on *grade-level* instruction. However, in order for struggling students to make progress, they need interesting materials that match their print-processing skills and comprehension capabilities.

After acquiring materials, teachers must implement innovative instructional strategies to address individual differences. These pedagogical changes should occur in both the general education English Language Arts (ELA) setting as well as in supplemental intervention programs. In ELA, teachers could incorporate thematic text sets into the curriculum (e.g., Civil War, Westward Expansion; Weather, Gravity) so that struggling readers can build content knowledge by reading content-related texts at the appropriate

difficulty level. For example, imagine that an ELA class was discussing the theme of racism in a class novel, *Huck Finn*. The teacher could gather multi-genre texts at various levels so that struggling students could address this theme in books they *can* read. A thematic text set might include excerpts from *Jackie Robinson* (biography, third- grade reading level) or *Day of Tears* (play, fifth-grade reading level). In addition, the teacher might incorporate a news article about recent events in Ferguson, Missouri from Newsela, a web-based resource. Using Newsela, the students could adapt the lexile, or readability level, of the text so that they could read at their respective instructional levels. Using thematic text sets can give students access to rich content while also improving their print-processing skills.

In addition to creating thematic text sets, middle school teachers could also group students as an instructional strategy for reaching low readers. In a given class of 30 students, a teacher might have five students who require materials one grade-level below, and three students who require materials two or more grade levels below. The teacher could use assessment information to diagnose their reading strengths and weaknesses so that he or she could teach them in small groups while the rest of the class is engaged with independent work. Although this task can be difficult to manage (in regard to both materials and behavior), small instructional-level reading groups can be very effective in middle school. The teacher-guided discussions in such groups allow the students to hear their peers' ideas and afford the teacher insight into the students' thinking styles. Moreover, when students are taught with books that they can actually read, they build the confidence needed to attack harder texts.

Limitations of the study

There are four major limitations of the present study. First, generalization of the findings is tentative because the sample of struggling readers ($n = 52$) in this study was smaller than the number of low readers in both the Morris (2014) ($n = 65$) and Hock et al. (2009) ($n = 345$) studies.

A second limitation was that the study was conducted in a predominantly rural, high-poverty setting. Therefore, the specific results may not generalize to the reading performance of students in suburban or urban settings. However, the issues of assessment and instruction in such settings remain the same.

A third limitation was that the informal comprehension measures used in this study were compromised in that the questions were passage-dependent as opposed to higher-level thinking. Nonetheless, the focus in the current study was on assessing print processing rather than comprehension.

A fourth limitation was how the full-length and shortened IRI assessments were analyzed. Because they were both part of the same sample (the shortened assessment was taken from the full-length assessment), an independent samples t -test analysis could not be conducted. If these were two separate samples, a t -test analysis could have provided further validation of the shortened measure.

Future Research

Although the present study yielded promising results for assessing middle school students' reading skill, further research is needed. First, the study should be replicated with a larger sample size. The present study concluded that the shortened IRI can be a valid

assessment to measure print-processing skills of the middle school students in this sample, but more research is needed to determine if the same results hold true for a larger sample.

Second, the study could be designed to incorporate elements of experimental design so that *t*-tests could be run across the full-length and shortened assessments. For example, participants could read both versions A and B of the IRI (instead of just one version or the other), and then *t*-tests could be run within and across passages and participants.

Finally, further research could focus on which words in a 10-word spelling list are the best predictors of orthographic knowledge. The present study concluded that a 10-word spelling test can serve as a screener for a student's print-processing knowledge, but future research could examine which specific words are the best to use at each grade level.

Conclusions

The purpose of the study was threefold: (a) to examine differences in reading-related behaviors between low-performing and average-performing middle school students; (b) to understand what reading-related differences existed within the low-performing group; and (c) to develop a shortened, teacher-friendly diagnostic reading assessment that had instructional significance. My research strategy involved two steps. First, I administered a battery of reading measures to 82 low- and average-achieving students and analyzed the results across and within groups. Second, I tried to shorten the assessments to see if I could develop a measure that yielded similar results but took less time to administer.

There were three major findings. The first finding was that the struggling middle school readers were different from their average-achieving peers, most prominently in reading rate or fluency. Second, within the low-performing group ($n = 52$), there were also large reading differences, again most significantly in the area of reading rate. And third, a

shortened IRI (100 words) proved to be an effective way of identifying struggling students' print-processing or instructional level. Because of the small sample size, the results in this study need to be replicated. Nonetheless, these findings add to the growing evidence that reading fluency (or print-processing skill) is a serious problem for older struggling readers. The results also point to promising informal assessments that can be used with this group of readers.

References

- Allington, R. L. (2002). You can't learn much from books you can't read. *Educational Leadership*, 60(3), 16-19.
- Baker, D. L., Biancarosa, G., Park, B. J., Boussetot, T., Smith, J. L., Baker, S. K., Kame'enui, E. J., Alonzo, J., & Tindal, G. (2015). Validity of CBM measure of oral reading fluency and reading comprehension on high-stakes reading assessments in Grades 7 and 8. *Reading and Writing*, 28, 57-104.
- Barr, R., Blachowicz, C., Bates, A., Katz, C., & Kaufman, B. (2007). *Reading diagnosis for teachers: An instructional approach* (5th ed.). Boston, MA: Allyn & Bacon.
- Bear, D. R., Invernizzi, M., Templeton, S., & Johnston, F. (2004). *Words their way: Word study for phonics, vocabulary, and spelling instruction* (3rd ed.). Upper Saddle River, NJ: Pearson.
- Betts, E. (1946). *Foundations of reading instruction*. New York: American Book Company.
- 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.
- Carver, R. P. (2000). *The causes of high and low reading achievement*. Mahwah, NJ: Erlbaum.
- Center on Response to Intervention at American Institutes for Research. (2014). *National Center for Learning Disabilities: RTI action network position statement on RTI*. Retrieved from <http://www.rtinetwork.org/about-us/position-statement>

- Chall, J. S., & Dale, E. (2000). *Readability revisited: The New Dale-Chall readability formula*. Brookline, MA: Brookline Books.
- Dennis, D.V. (2013). Heterogeneity or homogeneity: What do assessment data reveal about struggling adolescent readers? *Journal of Literacy Research*, 45(3), 3-21.
- Donors Choose [Website]. (n.d.). Retrieved from <https://donorschoose.org>
- Dunn, L. M., & Dunn, D. M. (1997). *Peabody picture vocabulary test* (3rd ed.). Circle Pines, MN: American Guidance Service.
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test* (4th ed.). Minneapolis, MN: Pearson Assessments.
- Fountas, I., & Pinnell, G. S. (2011). *Benchmark assessment system* (2nd ed.). Portsmouth, NH: Heinemann.
- Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, 5(3), 239-256.
- Good, R. H., Kaminski, R. A., Cummings, K., Dufour-Martel, C., Petersen, K., Powell Smith, K., Stollar, S., & Wallin, J. (2011). *Dynamic indicators of basic early literacy skills: DIBELS next assessment manual*. Eugene: University of Oregon. Retrieved from <http://www.dibels.org/>
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7, 6-10.
- Harris, A. J., & Jacobson, M. D. (1982). *Basic reading vocabularies*. New York: Macmillan.
- Hasbrouck, J., & Tindal, G. (2006). Oral fluency norms: A valuable assessment tool for reading teachers. *Reading Teacher*, 59, 636–644.

- Henderson, E., Templeton, S., Coulter, B., & Thomas, J. (1990). *Houghton Mifflin spelling*. Boston, MA: Houghton Mifflin.
- Hock, M. F., Brasseur, I. F., Deschler, D. D., Catts, H. W., Marquis, J. G., Mark, C. A., & Stribling, J. W. (2009). What is the reading component skill profile of adolescent struggling readers in urban schools? *Learning Disability Quarterly*, 32, 21-38.
- Hoover, W., & Gough, P. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127–160.
- Houghton Mifflin Harcourt. (2006). Steck Vaughn BOLDPRINT® anthologies series. Oakville, ON, Canada: Rubicon Publishing. Retrieved from <http://www.hmhco.com/shop/education-curriculum/reading/guided-reading/boldprint-anthologies>
- Houghton Mifflin Harcourt. (2013). *System 44 next generation* [Online software]. New York: Scholastic. Retrieved from <http://www.hmhco.com/products/system-44/>
- Johnson, E. S., Pool, J. L., & Carter, D. R. (2009). Screening for reading problems in grades 1 through 3: An overview of select measures. National Center on Learning Disabilities RTI Action Network. Retrieved from <http://www.rtinetwork.org>
- Lesaux, N. K., & Siegel, L. S. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology*, 39(6), 1005-1019.
- Leslie, L., & Caldwell, J. S. (1995). *Qualitative reading inventory* (2nd ed.). New York: Harper Collins.
- Leslie, L., & Caldwell, J. S. (2006). *Qualitative reading inventory* (4th ed.). Boston, MA: Allyn & Bacon.

- Leslie, L., & Caldwell, J. S. (2011). *Qualitative reading inventory* (5th ed.). Boston, MA: Allyn & Bacon.
- McKenna, M., & Stahl, S. (2003). *Assessment for reading instruction*. New York: Guilford Press.
- Morris, D. (2015). *Morris informal reading inventory: Preprimer through grade 8*. New York: Guilford Press.
- Morris, D., Bloodgood, J. W., Perney, J., Frye, E. M., Kucan, L., Trathen, W., Ward, D., & Schlagal, R. (2011). Validating craft knowledge: An empirical examination of elementary-grade students' performance on an informal reading assessment. *Elementary School Journal, 112* (2), 205-233.
- Morris, D., Meyer, C., Trathen, W., McGee, J., Vines, N., Stewart, T., Gill, T., & Schlagal, R. (2014). The "Simple View," instructional level, and the plight of struggling middle school readers. Manuscript submitted for publication.
- National Assessment of Education Progress. (NAEP). *The Nation's Report Card, 2013*. US Department of Education. Institute of Education Sciences, National Center for Education Statistics. Retrieved from www.nces.ed.gov/nationsreportcard.
- National Assessment of Education Progress. (NAEP). *The Nation's Report Card, 2015*. US Department of Education. Institute of Education Sciences, National Center for Education Statistics. Retrieved from www.nces.ed.gov/nationsreportcard.
- National Center for Response to Intervention. (NCRTI). (2010a). Tools and charts. National Center on Learning Disabilities RTI Action Network. Retrieved from <http://www.rti4success.org/resources/tools-charts>

National Center for Response to Intervention. (NCRTI). (2010b). What is RTI? National Center on Learning Disabilities RTI Action Network. Retrieved from

<http://www.rtinetwork.org/learn/what/whatisrti>

National Governors Association Center for Best Practices. (2016). *Common core state standards, English and language arts, college and career readiness anchor standards for reading*. Council of Chief State School Officers. Washington D.C: Governors Association Center for Best Practices.

Newsela: Nonfiction literacy and current events [Website]. (n.d.). Retrieved from

<https://newsela.com/>

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002).

Pearson Education, Inc. (2012). *Aimsweb: Reading curriculum-based measurement administration and scoring guide*. Bloomington, MN: Author.

Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp.145-174). Hillsdale, NJ: Erlbaum.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383.

Rasinski, T. V., & Zutell, J. B. (1996). Is fluency yet a goal of the reading curriculum? *Reading Research and Instruction, 36*, 237-246.

Renaissance Learning. (2015). *STAR reading* [Computer software]. Retrieved from

www.renlearn.com/se

- Richardson, J., & Walther, M. (2013). *Next step guided reading assessment*. New York: Scholastic.
- Rupp, A. A., & Lesaux, N. K. (2006). Meeting expectations? An empirical investigation of a standards-based assessment of reading comprehension. *Educational Evaluation and Policy Analysis*, 28(4), 315-333.
- Samuels, S. (2006). Toward a model of reading fluency. In S. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 24-46). Newark, DE: International Reading Association.
- Schlagal, R. (1982). *A qualitative inventory of word knowledge: A developmental study of spelling, grades one through six*. Unpublished doctoral dissertation, University of Virginia, Charlottesville.
- Schlagal, R. (1992). Patterns of orthographic development into the intermediate grades. In S. Templeton & D. Bear (Eds.), *Development of orthographic knowledge and the foundations of literacy: A memorial festschrift for Edmund H. Henderson* (pp. 32-52). Hillsdale, NJ: Erlbaum.
- Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical evaluation of language fundamentals, fourth edition (CELF-4)*. Toronto, Canada: The Psychological Corporation.
- Spache, G. (1953). A new readability formula for primary-grade reading materials. *Elementary School Journal*, 53(7), 410-413.
- Stauffer, R., Abrams, J., & Pikulski, J. (1978). *Diagnosis, correction, and prevention of reading disabilities*. New York: Harper & Row.

- Swanson, H. L., Trainin, G., Necochea, D. M., & Hammill, D. D. (2003). Rapid naming, phonological awareness, and reading: A meta-analysis of the correlation evidence. *Review of Educational Research, 73*(4), 407-440.
- Thinklink Learning. (2006). *Discovery education predictive assessment* [Computer software]. Nashville, TN: Discovery Education. Retrieved from http://www.discoveryeducation.com/pdf/assessment/Discovery_Education_Assessment_Research.pdf
- Thorndike, R. L. (Ed.). (1971). *Educational measurement* (2nd ed.). Washington, DC: American Council on Education.
- Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Technical manual: Stanford Binet intelligence scale* (4th ed.). Chicago, IL: Riverside.
- Torgesen, J. K., & Miller, D. H. (2009). *Assessments to guide adolescent literacy instruction*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.
- Torgesen, J. K., Myers, D., Schirm, A., Stuart, E., Vartivarian, S., Mansfield, W., Stancavage, F., Durno, D., Javorsky, R., & Haan, C. (2008). *Closing the reading gap: First year findings from a randomized trial of four reading interventions for striving readers*. Volume II: National Assessment of Title 1: Interim Report to Congress, Institute of Education Sciences, Washington, D.C.
- Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of word reading efficiency*. Austin, TX: PRO-ED.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). *The Comprehensive Test of Phonological Processing (CTOPP)*. Austin, TX: Pro-Ed.

- Ward, A. W., Stoker, H. W., & Murray-Ward, M. (Eds.). (1996). *Educational measurement: Origins, theories, and explications* (Vol. 2). Lanham, MD: University Press of America.
- Wiederholt, J. L., & Bryant, B. R. (2001). *GORT-4: Gray oral reading test*. Austin, TX: PRO-ED.
- Wiley, H. I., & Deno, S. L. (2005). Oral reading and maze measures as predictors of success for English learners on a state standards assessment. *Remedial and Special Education, 26*(4), 207-214.
- Wilkinson, G. S. (1993). *Wide Range Achievement Test: WRAT3*. Wilmington, DE: Wide Range.
- Wilson, J. (2005). *The relationship of Dynamic Indicators of Basic Early Literacy Skills (DIBELS) oral reading fluency to performance on Arizona Instrument to Measure Standards (AIMS)* [Technical Report]. Eugene, OR: University of Oregon.
- Woodcock, R. W. (1991). *Woodcock language proficiency battery-revised*. Allen, TX: DLM Teaching Resources.
- Woodcock, R. W., & Johnson, M.B. (1989). *Woodcock Johnson psychoeducational Battery-revised*. Allen, TX: Psychological Corporation.
- Zutell, J., & Rasinski, T. (1989). Reading and spelling connections in third and fifth grade students. *Reading Psychology: An International Quarterly, 10*(2), 137-155.

Appendix A
Word Recognition Assessment*

Sixth Grade List

1. elevate
2. conservation
3. tenderness
4. barrier
5. adulthood
6. kennel
7. humiliated
8. nonfiction
9. revive
10. wallet
11. depression
12. carvings
13. similarity
14. unanswered
15. fingernail
16. breed
17. marrow
18. starter
19. pedestrian
20. quantity

Seventh Grade List

1. civic
2. shirttail
3. nominated
4. gruesome
5. disadvantage
6. architecture
7. tonic
8. straightforward
9. warrant
10. unthinkable
11. ridicule
12. engulf
13. kindhearted
14. maturity
15. impassable
16. bolster
17. copyright
18. foliage
19. prune
20. persecution

*From Morris, D. (2015). *Morris Informal Reading Inventory: Preprimer through grade 8*. New York: Guilford Press.

Appendix B
Passage Reading Assessment

Sixth- and Seventh-Grade Passages from the
Morris Informal Reading Inventory (2015)

SIXTH GRADE – Oral

Form A (243 words)

“The Lost Colony”

Examiner's Introduction: This story tells about an early settlement in America.

In the summer of 1587, a group of 150 English colonists settled on Roanoke Island off the coast of North Carolina. The colonists, led by Governor White, became friendly with the Indians living on nearby Croatan Island. The Indians showed the colonists how to survive in the New World, but it soon became clear that the newcomers needed more supplies. Governor White quickly sailed back to England to pick up more colonists and bring supplies for the winter. However, his return was delayed because England had gone to war with Spain.

It was two years before White was able to sail back to the New World to check on the colonists at Roanoke Island. When he and his men reached the settlement, they found it deserted. The houses were torn down, the settlement was overgrown with weeds, and there was no sign of life. The only clue they found was the word CROATAN, carved into a large tree.

White thought that CROATAN meant that the settlers were on Croatan Island. He tried to sail there the next day, but bad weather prevented the trip. White had to return to England and never got a chance to search Croatan Island. Future settlers tried to discover the fate of the “lost colonists,” but were unsuccessful. Did the colonists perish during the harsh winter, were they taken in by friendly Indians, or did they meet some other fate? The answer remains a mystery to this day.

Questions

- Who was Governor White?
(The leader of the colonists on Roanoke Island.)
- Why did Governor White return quickly to England?
(To get more supplies for the winter.)
- Describe what the men saw when they came back to the colony after two years?
(Houses torn down; weeds everywhere; no sign of life. [need two for full credit])
- What was the clue they found?
(The word CROATAN carved into a tree.)
- Why didn't Governor White search for the missing colonists on Croatan Island?
(The weather was bad.)
- Give two possible reasons the colonists on Roanoke Island disappeared.
(They may have died during the harsh winter. and The Indians may have taken them in. [need both answers for full credit])

Error Quotient = 100/243 = .41

# of Errors	Accuracy	# of Errors	Accuracy
1	99	16	93
2	99	17	93
3	99	18	93
4	98	19	92
5	98	20	92
6	98	21	91
7	97	22	91
8	97	23	91
9	96	24	90
10	96	25	90
11	95	26	89
12	95	27	89
13	95	28	89
14	94	29	88
15	94	30	88

Total Errors = _____

Meaning Changes = _____

Oral Read. Acc. = _____%

Comprehension = _____%

Rate = _____ wpm
(14,580 ÷ No. of seconds)

From Morris Informal Reading Inventory; Preprimer through Grade 8 by Darrell Morris. Copyright 2015 by The Guilford Press. Permission to photocopy this material is granted to purchasers of this book for personal use only (see copyright page for details).

Note. The slash indicates the 100-word mark or the end of the shortened IRI measure.

SIXTH GRADE - Silent

Form B (231 words)

"The Popham Colony"

Examiner's Introduction: This story tells about an early settlement in New England

Many students learn that the first English settlement in America was founded in 1607 in Jamestown, Virginia. But in August of that same year, two ships from England landed on the coast of Maine with 120 people aboard. Led by George Popham, they started a settlement called the Popham Colony.

As soon as they landed, the colonists built a fort called Fort St. George. They needed the fort for housing and for protection against some of the native people living nearby. Soon after that, the colonists began building a ship so they could explore the coast of Maine and establish trade with England and other countries. Written records show that the Popham colonists were the first to build a sailing ship in the New World. They named it the Virginia. It was about 51 feet long (half the length of a basketball court) and weighed 30 tons.

Unfortunately, the Popham Colony in Maine did not survive. Exactly why the colony failed is not certain, but possible causes include the harsh Maine winter, the lack of supplies, and the deaths of its leaders. Fourteen months after its beginning, the colony disbanded, and all the remaining settlers returned to England aboard the Virginia. Although the Popham Colony failed, the building of the Virginia was a success. It marked the beginning of a long tradition of shipbuilding in Maine that continues to this day.

Questions

1. What was the relationship between the Popham Colony and the Jamestown Colony?
(Both colonies were founded in the same year.)
2. Why did the colonists immediately build a fort?
(For housing, and For protection from the native people. [need both for full credit])
3. Why did the colonists build the ship, the Virginia?
(So they could explore the coastline, and Trade with other countries. [need both for full credit])
4. What was important about the Virginia?
(It was the first sailing ship to be built in the New World.)
5. Give some reasons why the Popham Colony failed after only 14 months.
(Harsh winters; lack of supplies; deaths of its leaders. [need two for full credit])
6. After the colony disbanded, what happened to the remaining settlers?
(They returned to England on the Virginia.)

Error Quotient = 100/231 = .43			
# of Errors	Accuracy	# of Errors	Accuracy
1	99	16	93
2	99	17	93
3	99	18	92
4	98	19	92
5	98	20	91
6	97	21	91
7	97	22	91
8	97	23	91
9	96	24	90
10	96	25	89
11	95	26	89
12	95	27	88
13	94	28	88
14	94	29	88
15	94	30	87

Total Errors = _____

Meaning Changes = _____

Oral Read. Acc. = _____%

Comprehension = _____%

Rate = _____ wpm

(13,860 ÷ No. of seconds)

From Morris Informal Reading Inventory: Preprimer through Grade 8 by Darrell Morris. Copyright 2015 by The Guilford Press. Permission to photocopy this material is granted to purchasers of this book for personal use only (see copyright page for details).

SEVENTH GRADE - Oral

Form A (254 words)

"The Little Mermaid"

Examiner's Introduction: *This story tells about a famous statue.*

On a stone in the harbor of Copenhagen, Denmark, sits a life-sized statue looking out over the water. The statue's head is that of a woman, but her legs turn into a fish's tail. This is "The Little Mermaid," and it is the most photographed statue in the world.

The story of the Little Mermaid has been popular for a long time. Although the animated movie version of the story has a happy ending, the original tale is rather sad. Published in 1837 by Hans Christian Anderson, the legend tells of a mermaid who falls madly in love with ^{100 words} a prince. To win his love, she finds a witch who turns her into a human in exchange for her voice. When the prince eventually marries someone else, the heartbroken mermaid turns into seafoam and disappears.

A wealthy Danish businessman ordered the statue made after seeing a ballet based on the story. The sculptor modeled the statue's head after the leading dancer from the ballet. Cast in bronze, the statue was presented to the city of Copenhagen in 1913.

The Little Mermaid has not had an easy time in the harbor; she has been painted a number of times by vandals, and has had her head cut off more than once. One morning she was found in the water, completely blasted off her stone. So many people have climbed up to take photos with her that her luster has worn away. Yet, still beloved, she continues to be the biggest tourist attraction in Denmark.

Questions

1. Where does the statue of The Little Mermaid sit?
(On a stone in a sea harbor.)
2. What does the statue look like?
(It has the head of a woman but a fish's tail for legs.)
3. In the fairy tale, what happens to the Little Mermaid when her prince marries someone else?
(She turns into seafoam and disappears.)
4. Who paid for the statue to be made?
(A Danish businessman who had seen the ballet.)
5. Who did the sculptor use as a model for the statue's head?
(The lead dancer in the ballet.)
6. Name two things that vandals have done to the Little Mermaid statue.
(Painted; cut off head; blasted into water. [need two for full credit])

Error Quotient = 100/254 = .39			
# of Errors	Accuracy	# of Errors	Accuracy
1	99	16	94
2	99	17	93
3	99	18	93
4	98	19	93
5	98	20	92
6	98	21	92
7	97	22	91
8	97	23	91
9	96	24	91
10	96	25	90
11	96	26	90
12	95	27	89
13	95	28	89
14	95	29	89
15	94	30	88

Total Errors = _____
 Meaning Changes = _____
 Oral Read. Acc. = _____%
 Comprehension = _____%
 Rate = _____wpm
 (15,240 ÷ No. of seconds)

From Morris Informal Reading Inventory: Preprimer through Grade 8 by Darrell Morris. Copyright 2015 by The Guilford Press. Permission to photocopy this material is granted to purchasers of this book for personal use only (see copyright page for details).

Note. See Sixth-grade note (p.73).

SEVENTH GRADE – Silent

Form C (243 words)

“Palm Trees”

Examiner’s Introduction: *This passage tells about palm trees.*

When most people think of England, palm trees and balmy weather do not usually come to mind. Yet there are parts of England where palm trees do grow. How could an island as far north as England be home to palm trees? The two-part answer lies in the ocean and a misconception about palm trees.

The Gulf Stream is one of many currents or “rivers” in the ocean that move enormous amounts of water around the world. The Gulf Stream moves north from the Gulf of Mexico and brings warm water across the Atlantic Ocean toward England. It travels at a speed of 60 miles per day, moving 100 times more water than all the rivers on Earth.

When the Gulf Stream reaches the southwestern corner of England, its warm water makes the climate there considerably milder than the rest of the country. It is there, in the region of Cornwall, that the palm trees can be found in English gardens.

The popular idea of palm trees is that they grow exclusively on tropical islands. However, although palm trees cannot survive a harsh winter, they can grow outside the tropics. The palms that flourish in southwestern England are hardy ones, acquired from places like China, where palms grow high in the mountains.

Palm trees have made Cornwall, England a warm haven in an otherwise gray and chilly country; they have also made it one of the favorite tourist destinations in the British Isles.

Questions

1. Why does it seem unusual to have palm trees in England?
(England is far north and known to be chilly.)
2. What does the Gulf Stream do?
(It moves warm water from the Gulf of Mexico across the ocean.)
3. How does the Gulf Stream affect the climate of England?
(It makes the climate milder or warmer.)
4. What is a misunderstanding about palm trees?
(Most people think they grow only in warm climates.)
5. What two factors allow palm trees to grow in Cornwall, England?
(The mild climate, and The particular kind of palm grown there.)
6. What is one of the major sources of jobs in Cornwall, England?
(Serving the large number of tourists.)

Error Quotient = 100/243 = .41			
# of Errors	Accuracy	# of Errors	Accuracy
1	99	16	93
2	99	17	93
3	99	18	93
4	98	19	92
5	98	20	92
6	98	21	91
7	97	22	91
8	97	23	91
9	96	24	90
10	96	25	90
11	95	26	89
12	95	27	89
13	95	28	89
14	94	29	88
15	94	30	88

Total Errors = _____
 Meaning Changes = _____
 Oral Read. Acc. = _____ %
 Comprehension = _____ %
 Rate = _____ wpm
 (14,580 ÷ No. of seconds)

From Morris Informal Reading Inventory: Preprimer through Grade 8 by Darrell Morris. Copyright 2015 by The Gullford Press. Permission to photocopy this material is granted to purchasers of this book for personal use only (see copyright page for details).

Appendix C Spelling Assessment Lists

QWIK* Sixth Grade Spelling List (given to both sixth and seventh graders)

6th

1. **satisfied** (satisfied with my school grades)
2. **violence** (too much violence in the world)
3. **impolite** (impolite to stare)
4. **musician** (brother is a musician)
5. **illustrate** (illustrate a book)
6. **prosperity** (prosperity and freedom)
7. **accustom** (accustom yourself to new home)
8. **patriotic** (patriotic to hang flag outside)
9. **impossible** (impossible to get all As)
10. **wreckage** (wreckage after car crash)
11. **commotion** (commotion during accident)
12. **mental** (mental focus during test)
13. **conceive** (conceive good idea)
14. **admitted** (admitted cheating test)
15. **introduction** (introduction to the book)
16. **operating** (operating room)
17. **decision** (hard decision to make)
18. **acknowledge** (acknowledge hard work)
19. **connect** (connect with friends)
20. **declaration** (declaration of independence)

* Permission to reprint granted by R. Schlagal.

Schlagal, R. (1982). *A qualitative inventory of word knowledge: A developmental study of spelling, grades one through six*. Unpublished doctoral dissertation, University of Virginia, Charlottesville.

Schlagal, R. (1992). Patterns of orthographic development into the intermediate grades. In S. Templeton & D. Bear (Eds.), *Development of orthographic knowledge and the foundations of literacy: A memorial festschrift for Edmund H. Henderson* (pp. 32- 52). Hillsdale, NJ: Erlbaum.

Vita

Ali Wilson grew up in Mohnton, Pennsylvania. She earned a Bachelor of Arts in Psychology from Colgate University in Hamilton, New York in 2005. She went on to work as a research assistant in psychology but quickly learned that her passion was working with kids. So in 2007, she moved cross-country to Seattle, Washington, to pursue her Masters in Special Education at the University of Washington. Her professors, Dr. Joe Jenkins and Dr. Dixie Massey, sparked her love for literacy, and Ali went on to pursue an endorsement in reading. In Seattle, Ali first taught elementary school and then spent three years as a special educator and reading interventionist at a high-needs middle school in White Center, WA. During these three years, Ali worked closely with Dr. Dixie Massey to set up a summer reading clinic at her school. In order to further develop her understanding of beginning literacy, clinical work, and assessment, Ali moved cross-country again to pursue further education with Dr. Darrell Morris. From 2013-2016, she was enrolled in the Doctoral Program in Educational Leadership at Appalachian State University. She earned her Ed.D. in June, 2016. Ali is currently a Special Education Curriculum Designer and Assistant Professor of Practice for Relay Graduate School of Education in Philadelphia.