SEVENTH-GRADE STUDENTS READING ALOUD: AN EXAMINATION

OF THE RELATIONSHIP BETWEEN RATE AND ACCURACY

IN ORAL READING AND READING COMPREHENSION

by

LORI RAE SMITH

A DISSERTATION

University of Oregon Graduate School

Confirmation of Approval and Acceptance of Dissertation prepared by:

Lori Smith

Title:

"Seventh-Grade Students Reading Aloud: An Examination of the Relationship between Rate and Accuracy in Oral Reading and Reading Comprehension"

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Education degree in the Department of Educational Leadership by:

Gerald Tindal, Co-Chairperson, Educational Leadership
Leanne Ketterlin Geller, Co-Chairperson, Educational Leadership
Joe Stevens, Member, Educational Leadership
Jeffrey Measelle, Outside Member, Psychology

and Richard Linton, Vice President for Research and Graduate Studies/Dean of the Graduate School for the University of Oregon.

June 14, 2008

Original approval signatures are on file with the Graduate School and the University of Oregon Libraries.

An Abstract of the Dissertation of

Lori Rae Smith   for the degree of   Doctor of Education

in the Department of Educational Leadership   to be taken   June 2008

Title: SEVENTH-GRADE STUDENTS READING ALOUD: AN EXAMINATION OF

THE RELATIONSHIP BETWEEN RATE AND ACCURACY IN ORAL

READING AND READING COMPREHENSION

Approved: _____

Dr. Leanne Ketterlin-Geller, Co-Chair

Approved: _____

Dr. Gerald Tindal, Co-Chair

Many adolescents lack a necessary level of reading proficiency. Disaggregated data indicate severe inequities in reading outcomes for students in various demographic groups. This problem has serious implications for individuals and the nation and is difficult to address because there are multiple reasons why adolescents experience reading difficulty. Determining whether these students have acquired adequate basic reading skills is important. However, one of the challenges for educators is the scarce availability of technically adequate measures to inform decision making regarding when basic skills are adequate.

This study examined the relation between students' rate and accuracy on a measure of oral reading administered in the fall and their comprehension, as measured by scores on the Oregon State Assessment of Reading and Literature taken in the spring. A total of 422 students in grade 7 participated. Multiple regression analyses were used to determine whether oral reading rate, and rate in combination with accuracy, predict scores on the state assessment, and whether these findings varied by subgroups. Additionally, a discriminant function analysis was conducted to determine whether rate and accuracy predict students' performance level on the state test while controlling for gender, ethnicity, economic status, and special education identification classification.

Results indicated that rate of oral reading had a moderately strong correlation ($r = .62$) and accuracy of oral reading had a modest correlation ($r = .48$) with scores on the state reading test. Results of the multiple regression analyses indicated that 38% of the variance of scores on the state test was explained by rate. The addition of accuracy increased the variance explained to 40%. The discriminant function analysis resulted in a highly predictive model with 82% of students correctly classified. In addition, the discriminant function was strongly related to rate of oral reading, accuracy of oral reading, and special education classification. Gender, ethnicity, and economic disadvantage were not strongly related to the discriminant function. These results have positive implications for closing the achievement gap between minority and economically disadvantaged students and their peers. Implications for assessment and instruction are discussed.

CURRICULUM VITAE

NAME OF AUTHOR: Lori Rae Smith

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon
Lane Community College
Southern Oregon College

DEGREES AWARDED:

Doctor of Education, Educational Leadership, 2008 University of Oregon
Master of Arts, Special Education, 1996, University of Oregon
Bachelor of Arts, Psychology, 1994, University of Oregon

AREAS OF SPECIAL INTEREST:

Early Reading
Adolescent Literacy
Reading Assessment
System Performance

PROFESSIONAL EXPERIENCE:

Director of Student Achievement, Bethel School District, Eugene, Oregon
2004 -- present

Special Education Teacher, Bethel School District, Eugene, Oregon
1996 - 2004

Research Assistant, Behavioral Research and Teaching, University of Oregon
1994 - 1996

ACKNOWLEDGMENTS

I would like to express my sincere appreciation to my examining committee. I owe a debt of gratitude to Dr. Gerald Tindal for his continued support, encouragement, and confidence in me. In his classes I acquired the conceptual knowledge that has informed my career in education. Dr. Leanne Ketterlin Geller provided academic advice, excellent feedback, and clear direction that helped me plan my next steps and accomplish this task. Dr. Joe Stevens contributed suggestions and expertise instrumental in the design and evaluation of this research and gave helpful and specific feedback. Dr. Jeffrey Measelle's insightful questions and thoughtful input are greatly appreciated.

I also want to thank Dr. Drew Braun, who consistently provided excellent guidance and generously gave his time to help me navigate the complexities of SPSS and statistical analyses, and Dr. Jenelle Braun Monegan, whose friendship and cheerful encouragement helped me continue toward the goal of completion. Thank you also to Dr. Sarah McDonagh, who read early versions and shared thoughtful comments. My son, Bennet Smith, edited this manuscript and, along with the rest of our family, cheered me on. For them, I am most grateful.

This work is dedicated to my husband Bill, who could not be more steadfast in his love and support of me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Adolescent Reading Performance –

Improvement Needed

The *No Child Left Behind Education Act of 2001* (NCLB) ushered in a new era of

accountability for public education in America. While the merits and flaws of this

legislation continue to be debated, the need to ensure better and more equitable outcomes

for all students cannot be dismissed. This is particularly true in the area of adolescent

reading performance.

The need for better reading outcomes for older students is clear. Data from the

National Assessment of Educational Progress (NAEP) indicate the magnitude of the

problem. In 2005, more than two-thirds of the eighth graders tested scored below the

proficient level, and more than one-quarter were unable to read at even a basic level

(Perie, Grigg, & Donahue, 2005). Scores for 12th-grade students in 2002 showed a

significant decrease from the last assessment in 1998, with more than one-quarter of 12th

graders also reading below basic level (Grigg, Daane, Jin, & Campbell, 2003). Students

reading below proficiency on the NAEP are unable to demonstrate competency or solid

academic performance at grade level; students reading below basic level are unable to

demonstrate even literal understanding or identify aspects of overall meaning on grade-

level text.

In increasingly challenging secondary curricula (Kamil, 2003; Snow & Biancarosa, 2003), reading well is a prerequisite for academic success. For the more than 8 million struggling readers in Grades 4-12 (NCES, 2003), "chances for academic success are dismal because they are unable to read and comprehend the material in their textbooks" (Kamil, p. 1). While being a capable reader will not ensure success, the inability to read at or above grade level seriously limits a student's potential in all academic areas.

It is not only school performance that is affected by poor reading skills. As stated in *Becoming a Nation of Readers* (Anderson, Hiebert, Scott, & Wilkinson, 1984), "Reading is a basic life skill. It is the cornerstone for a child's success in school and, indeed, throughout life" (p. 1). Increasingly high levels of literacy are required for living-wage jobs (Barton, 2000). Students who leave school without a level of literacy sufficient to enter the skilled workforce or successfully enter higher education are without means to access the economic and social capital necessary for personal autonomy. For older students with serious reading difficulties, the prospects are grim. Students who drop out or leave school with poor reading skills are disproportionately represented in the juvenile justice system, and large numbers of incarcerated juveniles are illiterate or marginally literate (Coalition for Juvenile Justice, 2001).

The Achievement Gap – A Civil Rights Issue

In addition to the need for improved outcomes in reading for older students, there is an urgent need for more equitable outcomes. Students with below-basic reading skills

are far more likely to have disabilities, lack English proficiency, or be students of color or from economically disadvantaged families (Perie et al., 2005). Not only are these differences in outcomes evident on the NAEP, but they are also reflected on state assessments (McCombs, Kirby, Barney, Darilek, & Maggee, 2004). While not directly comparable, "there is a remarkable consistency across the state and national assessments in the magnitude of the performance gaps between minority and economically disadvantaged students and their peers" (McCombs et al., p. 60).

Students in these demographic groups are also more likely to drop out of school. Nationwide, approximately 70% of the 9th-grade students who enter high school graduate with a regular diploma in 4 years. On-time graduation rates are even lower for urban students (50%) (Joftus, 2002) and minority students: African American (50%), Native American (51%) and Latino (53%) (Greene & Forster, 2003, cited in McCombs et al., 2004).

The disparity of educational outcomes for students of color and poverty has been termed the achievement gap. Closing the achievement gap is a national priority. NCLB mandates that data on educational performance, such as test scores and graduation rates, be disaggregated by student subgroups. For schools to be considered successful they must eliminate the achievement gap in all areas. Furthermore, NCLB requires that all students reach proficiency in reading and mathematics by 2014.

The reading achievement gap between adolescents in different demographic groups leads to an economic and social gap that presents a serious challenge to America's democratic ideals. Oregon School Superintendent Susan Castillo stated that addressing

these inequities "is our moral imperative; it is a civil rights issue, regardless of federal law" (Oregon Department of Education, 2006). Educators have a responsibility and an opportunity to improve outcomes and increase social equity and fairness for all adolescents.

Closing the Achievement Gap:

"With Literacy and Justice for All"

Many secondary schools have undertaken a variety of efforts to increase student success and to close the achievement gap. *Breaking Ranks II: Strategies for Leading High School Reform,* (National Association of Secondary School Principals [NASSP], 2004) features 31 recommendations for high school principals, suggesting that this number allows for greater focus than the still-relevant 82 recommendations from the original report, *Breaking Ranks* (NASSP, 1996). These various efforts are important and may ultimately improve both the context and content of instruction in secondary schools. However, if students cannot read well, they will continue to struggle in even the best situations.

Evidence that closing the achievement gap requires a focus on adolescent literacy is found in the Executive Summary of the Carnegie Corporation's Adolescent Literacy Funders' Forum Report (Snow & Biancarosa, 2003). At the meeting, attended by nine public and private funding organizations, the issue of adolescent literacy was addressed. The Bill and Melinda Gates Foundation reported that many high schools undergoing systemic change through their Small High School and Early College High initiatives have

found that organizational change is not enough and that "Grantees came back to funders with concerns about the challenge of students' lagging literacy skills" (p.18). The William and Flora Hewlett Foundation also reported that addressing adolescent literacy is an important aspect of its grant-making process, as literacy is an area of specific challenge in the Foundation's systemic secondary school reform efforts. The authors of this report conclude that the "minority achievement gap results primarily from the poor literacy accomplishments of African-American, Latino, ELLs (English Language Learners) and low-income children in the United States. Shrinking the gap will require improving literacy instruction for those groups in particular" (p. 31).

Teaching all students to read well and achieve high levels of literacy is the first step toward equity in educational outcomes. However, in the area of adolescent reading, the achievement gap has remained relatively unchanged. According to Kamil (2003), "reading scores of high school students have not improved over the last thirty years," and "in recent years twelfth graders' scores have decreased significantly" (p. 2). Educational equity remains elusive despite the priority of the goal and the mandate of NCLB.

<div align="center">

Improving Adolescent Reading –

From Reading First to Reading Next

</div>

Most efforts to improve reading have focused on beginning reading and early grades. This is due in large part to the well-researched and well-reported work on this topic. Important works such as *Beginning to Read: Thinking and Learning about Print* (Adams, 1990), *Preventing Reading Difficulties in Young Children* (Snow, Burns, &

Griffin, 1998) commissioned by the National Academy of Sciences, and the report of the National Reading Panel (National Institute of Child Health and Human Development [NICHD], 2000) provide a research basis to inform assessment and instruction in early reading. Through the Reading First Initiative, part of NCLB, the U.S. Department of Education will have provided $5 billion over 5 years to states to implement evidence-based practice in beginning reading assessment and instruction.

Investment in improving early reading outcomes is important, but it is only the beginning. It was once thought that students who could decode and understand grade-level text by third grade were ready to transition from learning to read to reading to learn. However, many adequate readers in third grade begin to experience difficulty in fourth or fifth grade. Jeanne Chall (2003) termed this phenomenon the *fourth grade slump* in the American Educator's special issue devoted to the topic. Students who leave third grade with less than adequate reading skills begin fourth grade with an even greater chance of taking an academic plunge as they are expected to read and comprehend academic subject matter in content-area textbooks.

There are a number of possible explanations for the reading difficulty that many students experience in 4th grade and beyond. Texts shift from narrative to expository, and text structure becomes increasingly sophisticated (Armbruster & Anderson, 1988). In addition, unfamiliar vocabulary and a multitude of multisyllabic words appear (Nagy & Anderson, 1984). The difficulty level of these elements compounds with each grade level and continues to increase in college and the workplace. Thus, many students need

continued explicit instruction in navigating the demands of the diverse and challenging text they will encounter.

Without instruction and intervention, readers who are less than proficient in reading the increasingly difficult texts struggle to keep up and gain meaning from what they read. Students who enjoy reading read more than those who don't find it reinforcing. According to Cunningham and Stanovich (2001), when reading is difficult, reluctant readers read less, and a downward spiral begins: "Reading has cognitive consequences ... that are reciprocal and exponential in nature" (p. 137), and this difference in the amount of reading volume produces increasing divergence in reading ability. Stanovich (1986) described these reciprocal effects of reading as *Matthew effects*, a reference to the Biblical passage that describes a rich-get-richer and poor-get-poorer situation.

Struggling readers in 4th grade are in need of reading instruction that will help them gain the skills and knowledge they need to avert or interrupt this downward slide. However, at the same time as expectations for increasingly higher levels of literacy are introduced in 4th grade, there is often a decline in the amount of reading instruction and support services available to struggling readers. The necessity to cover content in a standards-based curriculum may compete with time that might otherwise be available for reading interventions. In addition, due to the emphasis on the prevention of early reading problems, resources are usually allocated to the early grades. Title I funds are typically budgeted to promote reading success through grade 3, with little remaining for struggling readers in grades 4 and beyond.

Given the increased difficulty of texts, the increased demands on comprehension, and the decreased instruction in reading that occurs beyond third grade, perhaps it should not be a surprise that so many readers in grades 4 through 12 experience difficulty becoming proficient readers. According to Biancarosa and Snow (2004), meeting the needs of these older readers will "require expanding the discussion of reading from Reading First – acquiring grade-level reading skills by third grade – to Reading Next – acquiring the reading comprehension skills that can serve youth a lifetime" (p. 8).

Assessment: Determining the Adequacy

of Basic Literacy Skills

A discussion about the need to improve adolescent reading skills has begun. A number of key reports and summaries have been issued stressing the urgency and need for addressing what the National Association of School Boards of Education called "Reading at Risk: The State Response to the Crisis in Adolescent Literacy" (2006). The following reports are filled with recommendations for educators, researchers, and policymakers:

- *Academic literacy instruction for adolescents: A guidance document from the center on instruction.* Torgesen et al. (2007).

- *Adolescent literacy: A position statement.* International Reading Association. Moore, Bean, Birdyshaw, & Rysik (1999).

- *Adolescent literacy and the achievement gap: What do we know and where do we go from here?* Snow & Biancarosa (2003).

- *Adolescents and literacy: Reading for the 21st century.* Kamil (2003).

- *Reading for understanding: Toward a research and development program in reading comprehension.* Snow (2002).

- *Reading next: A vision for action and research in middle and high schools.* Biancarosa & Snow (2004).

- *Achieving state and national literacy goals, a long uphill road: A report to the Carnegie Corporation of New York.* McCombs et al. (2004).

- *Reading to achieve: A governor's guide to adolescent literacy.* Berman & Biancarosa (2005).

These reports illustrate the magnitude of concern for addressing the issue of improving adolescent literacy. However, they offer little in the way of specific guidance in solving the problem at the individual student level. Researcher Catherine Snow, in the foreword to *Reading Next: A Vision for Action and Research in Middle and High Schools* (Biancarosa & Snow, 2004), states that "Educators must figure out how to ensure that every student gets beyond the basic literacy skills of the early elementary grades, to the more challenging and more rewarding literacy of the middle and secondary years" (p. 1). Early reading research provides specific guidance regarding assessment to inform decisions relevant to improving the skills of early readers. As evidenced by Snow's comment above, such specific guidance is less available to educators working to improve the skills of adolescent readers.

Ensuring that students get beyond the basic literacy skills of early reading requires knowing when early reading skills are adequate. Snow, Burns, and Griffin (1998)

reported that students who do not acquire these basic skills are at risk for developing reading difficulties. Secondary students who lack these skills will continue to struggle in reading unless they are identified so that they can receive continued instruction in necessary reading skills. Determining when older students have adequate basic early reading skills is not easy but it is important.

One of the challenges for those working to improve adolescent literacy is the scarce availability of reliable and valid measures designed to inform such decision making. There is very little research addressing this. There is, however, a recommendation that extending the research base by identifying acceptable oral reading rates by grade and identifying "information about performance levels, by grade, that predict success on graduation tests would serve as an important practical function for determining which students require special intervention" (Fuchs, Fuchs, Hosp, & Jenkins, 2001, p. 252).

Having data early in the school year regarding which adolescent readers are at risk of not passing state assessments of reading would provide educators a means of identifying students who may need extra instruction or who are likely to benefit from reading intervention. The lack of such information particularly hampers educators who teach students in junior high or middle schools where there is still time to intervene and help students improve their reading skills before they enter high school.

The purpose of this study is to examine the association between measures of oral reading (rate and accuracy) and a measure of comprehension (defined by performance on a state standardized reading assessment) for students in grade 7. In addition, the study

examines whether consideration of accuracy in oral reading improves the prediction of students' scores on a state test of reading, and whether these results vary for certain groups of students. Finally, this study addresses the question of whether rate and accuracy of oral reading predict student performance level on the state assessment of reading, while controlling for gender, ethnicity, economic status, and special education identification classification. Such information may help identify which adolescent readers are still in need of instruction to improve basic skills. The implications may be important for adolescent readers, whose skill deficits may otherwise go unidentified, and for educators and policymakers working to close the reading achievement gap and provide all students an education that will afford them equal opportunity.

CHAPTER II

LITERATURE REVIEW

Adolescent Literacy

Adolescent literacy is a broad term, used widely in educational research and policy. Any discussion of adolescent literacy must begin with clarification about the terminology. Generally, *adolescent* means someone no longer a child but not fully an adult. In education the term is often used to describe students in junior high, middle school, or high school. Therefore the term *secondary literacy* is sometimes used synonymously with the term *adolescent literacy*. Though students in fourth and fifth grade are not considered to be in secondary school, in terms of literacy they are sometimes included in the definition of adolescents (McCombs et al., 2004). This is largely because, as Berman and Biancarosa (2005) describe, "adolescent literacy refers to the set of skills and abilities that students need in grades four through twelve" (p. 6), the grades in which students are expected to navigate difficult texts while reading to learn.

The literacy that adolescents need "encompasses both *more than* just reading and reading in *many and varied* forms" (Torgesen et al., 2007, p. 2). Some definitions of literacy expand beyond reading and include writing and thinking about text (Berman & Biancarosa, 2005). Heller and Greenleaf (2007) define it even more broadly as "the capacity to draw inferences from academic texts, synthesize information from various sources, and follow complicated directions" (p. 5).

Though literacy may be broadly defined, it is closely linked with the ability to read accurately and understand the words involved. Thus, many of the documents which address adolescent literacy specify the need to improve adolescent students' reading proficiency (Berman & Biancarosa, 2005; Biancarosa & Snow, 2004; Kamil, 2003; Torgesen et al., 2007). Although there are a number of types of literacy, they all depend on the prerequisite skill of being able to read the words.

Reading – A Complex Construct

Helping adolescents improve their reading proficiency can be a challenging endeavor because reading is a complex task consisting of multiple component skills, and a deficiency in any area may impact reading performance. The National Reading Panel (NICHD, 2000) reported five areas essential for reading with understanding: phonemic awareness, decoding, fluency, vocabulary, and comprehension.

1. Phonemic awareness relates to the ability to hear, identify, and manipulate the sounds within words. Phonemic awareness is a predictor of early reading ability and is an important skill for beginning readers.

2. Decoding words requires an understanding of the relationship between sounds (phonemes) and the letter or letters (graphemes) that represent them. Instruction in phonics helps students decode and read words.

3. Fluency is the ability to read connected text accurately, automatically, and with proper expression.

4. Vocabulary is important because for meaning to occur, pronunciation of words in text must be accompanied by an understanding of the concepts that words represent.

5. Comprehension results from a combination of the preceding skills, but comprehension is complex and depends also upon the type of text and terminology, the purpose for reading, and the background knowledge belonging to the reader.

These five component skills are intended to be mastered during intensive reading instruction in kindergarten through third grade. Ideally, students would enter grades 4 and beyond having mastered these components, or sub-skills, of reading and have sufficient interest and background knowledge to read and understand grade-level content texts. If that were the case, adolescent literacy instruction could focus solely on "the core of reading comprehension, learning while reading, reading in the content areas, and reading in the service of secondary or higher education, of employability, of citizenship" (Biancarosa & Snow, 2004, p. 1). However, many adolescents do experience reading difficulties because they have not mastered the important component skills of reading.

Students may have difficulty in one or more of the 5 component skills. Some older students have not had adequate instruction in phonics (Adams, 1990), and as many as one out of ten adolescents has serious difficulty identifying words (Curtis & Longo, 1999). Other students may have adequate basic decoding skills but have not mastered the skill of decoding multisyllabic words (Archer, Gleason, & Vachon, 2003). Or, students may be able to decode all word types but decode words so slowly that lack of fluency

interferes with comprehension (Rasinski et al., 2005). Alternatively, students may have acquired adequate decoding and fluency skills but lack necessary vocabulary and background knowledge (Marzano, 2004), strategies (NICHD, 2000), or experience with various text features and structures (Heller & Greenleaf, 2007) and therefore experience difficulties with reading comprehension. Regardless of the reason for the difficulty, a lack of any of the components of reading skill significantly impacts students' ability to engage with grade-level texts.

<div align="center">How the Components Fit Together</div>

<div align="center">(Theories of Reading)</div>

Deficits in any component skill affect reading proficiency. The exact nature of the interaction of component skills remains a topic of interest and investigation, but a number of researchers endorse a developmental theory of reading (Carnine, Silbert, Kame'enui & Tarver, 2004; Chall, 1983; LaBerge & Samuels, 1974; Pikulski & Chard, 2005). The RAND Reading Study Group, in its publication *Reading for Understanding: Toward a Research & Development Program in Reading Comprehension* (Snow, 2002), sees "achieving reading proficiency as a long-term developmental process; what constitutes 'reading well' is different at different points in the reader's development" (p. 9). These points can be described by a developmental model of reading.

Jeanne Chall (1983) proposed a developmental model of reading that incorporates 6 stages. In Stage 0, the pre-reading stage (ages 0-6), children learn vocabulary, begin to understand the sound structures of words, and engage in reading-like activities such as

holding books and turning pages. In Stage 1 (grades 1 and 2), students learn the letters of the alphabet and their corresponding sounds. Decoding skills and word recognition are evident but typically plodding and laborious. In Stage 2 (grades 2 and 3) students expand decoding skills to more difficult words and practice reading, developing fluency, and automaticity. In Stage 3 (grades 4-8) students are reading to learn new information and gaining vocabulary and background knowledge. Stage 4 (high school) consists of reading in greater depth, often textbooks laden with information and containing multiple perspectives. Stage 5 (ages 18 and above) reading is personalized to meet individual needs, and information is synthesized and integrated with the reader's viewpoint and background knowledge.

The stages of the developmental model correlate with the 5 components of reading identified by the National Reading Panel (NICHD, 2000). According to Carnine et al. (2004) the 5 components are consistent with Chall's (1983) developmental model. They propose that Stages 0 and 1 reflect phonemic awareness and decoding; Stage 2, fluency; Stage 3, vocabulary; and Stages 4 and 5, text comprehension. However, they point out that the components (or stages) are not to be considered discrete categories; rather they are continuous and overlap.

The overlapping nature of this continuum permits simpler models, such as the 2-part paradigm proposed by Shinn, Good, Knutson, Tily, & Collins (1992). This model consists of two major components: decoding and comprehension. At the early stage of reading, individuals must devote significant cognitive resources to decoding words. According to the theory of automatic processing (LaBerge & Samuels, 1974), cognitive

energy is finite. Attention allocated to the lower-level reading process of decoding words detracts from the cognitive energy needed for comprehension. This characterization doesn't account for a transition from the early skill of decoding to the later skill of comprehension.

Pikulski and Chard (2005) utilize this conceptualization but propose that fluency is the bridge between decoding and comprehension. This categorization is consistent with Chall's (1983) stages of reading development if stages 0, 1, and 2 are considered early reading skills and stages 3, 4, and 5 are considered advanced reading skills. It is also consistent with the National Reading Panel model of components, with fluency situated between decoding and comprehension. Recalling the recommendation of Biancarosa and Snow (2004) that educators must ensure that students get beyond the "basic literacy skills of the early grades" and on to "the more challenging and rewarding literacy of the middle and secondary years" (p. 1), it may be helpful to categorize, as do Pikulski and Chard (2005), the basic early skills as decoding and the later secondary skills as comprehension, with fluency being an important factor that facilitates the move from one to the other. This study utilizes this conceptualization, and so it is important to further define the terms *decoding*, *fluency* and *comprehension*.

Decoding

Carnine, Silbert, and Kame'enui (1990) define decoding as the translation of printed words into a representation similar to oral language. Inadequate word recognition skills are believed to be the most debilitating source of reading challenges (Adams, 1990; Share & Stanovich, 1995). Secondary students who have difficulty with decoding fall

into one of two categories (Archer et al., 2003). The first group of students has difficulty

decoding words commonly learned in first and second grade. The number of these

students is relatively small, and they need intensive intervention in the early skills of

reading. The larger group has mastered single-syllable and high frequency words but

struggles with decoding multisyllabic words. The inability to decode multisyllabic words

is highly problematic for secondary students who are expected to learn from intermediate

and secondary textbooks replete with multisyllabic words. Nagy and Anderson (1984)

estimated that from 5th grade on, the average student encounters approximately 10,000

words per year that he or she has never before seen in print.

Secondary students who cannot adequately decode the words they encounter –

short or long – face a formidable challenge. It is important to identify these students so

they can be provided with instruction that will enable them to read words accurately, yet

it is not always easy to recognize which students need this type of instruction. Observing

students' silent reading or evaluating their responses to questions about a passage read

silently provides scant information regarding a student's ability to decode words.

Listening to students read aloud is the best way to assess their ability to decode words

accurately.

Fluency

Listening to a sample of oral reading also provides information regarding the

student's rate of reading. Oral reading fluency is most often defined as the ability to read

accurately, quickly, and with proper expression (NICHD, 2000). Meyer and Felton

(1999) provided more detail, describing fluency as the "ability to read connected text

rapidly, smoothly, efficiently, and automatically with little conscious attention to the mechanics of reading, such as decoding" (p. 284).

According to Pikulski and Chard (2005), a definition of fluency needs to encompass more than just oral reading and include comprehension. They propose a more comprehensive definition of fluency, a synthesis of the definitions in the Report of the National Reading Panel (NICHD, 2000) and The Literacy Dictionary (Harris & Hodges, 1995) and state:

> Reading fluency refers to efficient, effective word-recognition skills that permit a reader to construct the meaning of text. Fluency is manifested in accurate, rapid, expressive oral reading and is applied during, and makes possible, silent reading comprehension. (p. 510)

Other researchers purposely pare down the definition. Archer et al. (2003) adopted what they termed a "pragmatic definition of fluency" defining fluency as "rate plus accuracy" (p. 96). Such parsimonious definitions are useful because they are more measurable and observable. Fuchs et al. (2001) also defined oral reading fluency as "The oral translation of text with speed and accuracy" (p. 239) and found that oral reading fluency (as measured by rate and accuracy) serves as an indicator of overall reading competence. For the purpose of this study fluency is defined as rate and accuracy in oral reading.

Increased fluency (rate and accuracy) is important for secondary students (Archer et al., 2003). Research supports the relationship between fluent oral reading and overall reading ability including comprehension (Meyer & Felton, 1999; Rasinski et al., 2005),

and such understanding is the goal of reading. A lack of fluency impacts work completion because reading assignments may take twice as long for struggling readers. Furthermore, it renders reading slow and difficult and so students read less (Moats, 2001), and this lack of practice causes further delays and difficulties.

Comprehension

According to Lenz (2005), "comprehension is the process of constructing meaning from text" (p.1). Snow (2002) expands this definition, defining reading comprehension as "the process of simultaneously extracting and constructing meaning through interaction and involvement with written language" (p. 11). Constructing meaning from written language requires making a connection to oral language and vocabulary. A printed word must be connected to an individual's phonological memory for the word and be connected to the syntax and meaning of the word or it cannot be fluently decoded or understood (Pikulski & Chard, 2005).

Vocabulary contributes to reading comprehension (Yovanoff, Duesbery, Alonzo, & Tindal, 2005). The NRP acknowledged the importance of vocabulary as a component of comprehension, reporting that reading comprehension "integrates complex skills and cannot be understood without examining the critical role of vocabulary learning and instruction and its development" (p. 4-1). It is this integration of complex skills that results in comprehension and reflects the reading skills that adolescents need to acquire. This is consistent with the advanced reading skills described by Chall (1983) in Stages 3, 4, and 5, which include the ability to read to learn new information, gain vocabulary and background knowledge, and synthesize and integrate information.

Instruction can support and improve students' reading comprehension skills. The value of teaching vocabulary, both content terminology and academic language, to students has been well documented (Marzano & Pickering, 2005; NICHD, 2000). Teaching comprehension strategies to secondary students is important, and research supports the teaching of comprehension strategies as an integral part of improving outcomes for adolescent readers (Biancarosa & Snow, 2004; Kamil, 2003; Mastropieri, Scruggs, & Graetz, 2003; NICHD, 2000).

However, according to Archer et al. (2003), "No comprehension strategies are powerful enough to compensate for not being able to read the words within the text" (p. 90). Reading the words accurately is critical, but reading the words at an adequate rate is highly related to comprehension (Daane, Campbell, Goodman, & Orange, 2005; Pikulski and Chard, 2005; Pinnell et al., 1995; Rasinski et al., 2005). Therefore, educators must be able to identify students who are still lacking the foundational skills of decoding and fluency, described by Chall (1983) as Stages 0, 1, and 2, so that they can receive instruction to progress to the advanced reading level described in Stages 3, 4, and 5. Understanding more about the level of skills represented at the interface between Stages 2 and 3 where the divide between early literacy skills and advanced literacy skills occurs would help educators make decisions about which students need intervention to improve basic reading skills.

Information for Decision Making

As stated in the previous chapter, ensuring that students get beyond the basic literacy skills of early reading requires knowing when early reading skills are adequate (Snow, Burns, & Griffin, 1998). Yet, one of the challenges for educators charged with helping adolescents improve their reading is the scarce availability of reliable and valid measures designed to inform such decision making.

Adolescence and Reading Assessments

In an educational context, assessment, as defined by Salvia and Ysseldyke (1995), is "the process of collecting data for the purpose of making decisions about students" (p. 5). According to the International Reading Association's position statement on adolescent literacy (Moore et al., 1999), "Adolescents deserve assessment that shows their strengths as well as needs and that guides their teachers to design instruction to help them grow as readers" (p. 6). Educators and policymakers also need this type of assessment information to guide their decision making in improving outcomes for struggling readers.

A number of the major documents and reports on adolescent literacy listed in the previous chapter note the importance of having good assessments and data on which to base decision making. However, only two of them provide guidance in this area. The recommendations they make are preliminary in nature.

Reading Next: A Vision for Action and Research in Middle and High Schools (Biancarosa & Snow, 2004) makes it clear that good assessment is critical. The report delineates 15 key elements important for improving adolescent literacy (9 instructional improvements and 6 infrastructure improvements). The authors state that while research

and/or professional opinion supports each of the 9 instructional improvements and 6 infrastructure improvements recommended, the optimal mix of these factors has not been determined. However, the authors emphasize that 3 of the elements are critical, making their point by stating that "15 − 3 = 0" because, "Without a) professional development, b) ongoing formative [assessment] and c) ongoing summative assessment of students and programs as the foundation of any middle or high school literacy program, we cannot hope to effect major change in adolescent literacy achievement, no matter what instructional innovations are introduced" (p. 29).

Two of the 3 essential elements cited by Biancarosa and Snow are assessment-related. The report advocates for systems in which both formative (frequent assessments to inform classroom instruction) and summative (end-of-program or end-of-year tests designed to document achievement and evaluate progress) assessment data are collected and catalogued in databases for inspection by teachers, administrators, and evaluators. Yet, despite emphasis on the pivotal importance of these 2 types of assessment, the authors provide only one paragraph of general information for each of the 2 types of assessment; no specific measures are recommended or mentioned.

The other report that provides some information on assessment contains a short chapter on the topic and lists some specific measures used to assess the reading skills of adolescents. In the report *Improving Adolescent Literacy: Findings from the Research by the Northwest Regional Education Lab* (Davis, Spraker, & Kushman, 2005), the authors provide a chart (Table 1) of available assessments adapted from Howell and Nolet (2000). The chart frames the assessments by types and purposes and lists the names of

specific assessments in each of 3 categories: state assessments, diagnostic assessments, and progress monitoring measures. Screening, an important 4th category of assessment, is notably absent from the chart and will be discussed later in this section.

Table 1

Types and Purposes of Assessments

Characteristic	State assessments	Diagnostic	Progress monitoring
Purpose	To document achievement, evaluate schools, inform policy decisions	To provide specific information on skills or strategies of individual students	To determine whether students are learning at appropriate rates to inform instruction
Examples of individual measures		QRI III, individual reading inventories, Comprehensive Test of Phonological Processing (CTOPP), oral reading fluency measures	Oral reading fluency measures, Qualitative Reading Inventory III, individual reading inventories
Examples of group measures	State Assessments: WA-WASL; OR-OSA; ID-ISAT; MT-MontCAS; AK-Benchmark exams & HSGQE; ITBS; CTBS; ITED; SAT-9	Bader Reading & Language Inventory; Brigance Comprehensive Inventory of Basic Skills; Stanford Diagnostic	Curriculum-based measures: reading, vocabulary timed tests, classroom-based assessments
Primary audience	Policymakers, parents, public	Teachers and students	Teachers and students
Schedule and frequency	Annual, usually in the Spring	When determined that students are not proficient in reading	3-4 times per year (or more) at critical decision-making points
Cost	Very high	Moderate	Low

Source: Adapted from Howell and Nolet (2000)
In Improving Adolescent Reading: Findings from Research. Northwest Regional Educational Laboratory. Davis, Spraker and Kushman, (2005).

State Assessments

According to the chart, the purpose of state tests is to document and evaluate achievement. Most adolescent readers take a state or standardized assessment, if not yearly, then periodically, as part of the monitoring of accountability in this era of NCLB. National concern for the state of adolescent literacy is generated from the results of student performance on such assessments.

State tests, a form of outcome evaluation, are summative assessments intended primarily for policy makers, parents, and the public. Increasingly higher stakes are attached to these tests. Stakes are high for students because in many states the test scores are linked to important outcomes such as graduation. Stakes are high for schools and districts because of the sanctions associated with not making adequate yearly progress (AYP) as required by NCLB. In both cases, the consequences loom large.

These tests, in which students silently read passages and answer multiple choice questions about what they have read, indicate which students are and are not yet proficient on grade-level assessments of silent reading comprehension; they do not provide information about how to move these students to proficiency. Yet, standardized test scores are often the only data available for secondary teachers making decisions regarding the reading ability of their students. Without further information, teachers and administrators are left not knowing how to change outcomes for older students who lack proficiency.

Often secondary school teams mistakenly believe that adopting a new instructional program will help solve the problem of adolescent literacy, but this is

seldom the case. Students continue to struggle with reading because a single-approach solution fails to deliver the continuum of instructional support that is most likely needed. Buly and Valencia (2003) studied 108 4th-grade students who failed to pass a traditional state reading test. They found that "scores on state tests mask distinctive and multifaceted problems having to do with word identification, fluency, and meaning" (p. 3). Educators need more data than state test scores if they hope to change outcomes for students at risk of failure.

Diagnostic Assessments

Diagnostic tests are formative assessments that can be helpful to educators charged with instructional planning and addressing student need. Diagnosis provides specific detailed information about an individual's strengths and weaknesses in reading and therefore requires assessment of a broad range of skills; however, the items on the test represent a limited scope within the specific areas. In addition, diagnostic tests are not sensitive to improvement over short periods of time. Formal diagnostic measures can be expensive and time consuming. It is usually impractical to administer these tests to large numbers of students, and these tests are rarely used with older students.

Progress Monitoring Assessments

If secondary students are receiving an instructional intervention in reading, it is important to administer progress monitoring assessments. This requires frequent brief assessment using alternate forms of a test, measuring student progress weekly or monthly. The Research Institute on Progress Monitoring and the National Center on Student Progress Monitoring provide information to educators and practitioners about the

reliability and validity of progress monitoring tools. Progress monitoring measures help

determine if learning is progressing at an adequate rate but are extraneous if students are

not receiving systematic reading instruction or intervention in basic literacy skills, as is

often the case for struggling readers in secondary schools.

Screening

Kame'enui et al. (2006) evaluated the adequacy of reading measures for use in

kindergarten through 3rd grade. They classified reading assessments according to 4

different decision-making purposes. They identified the same categories as those in Table

1 (referring to state assessments by the broader category of outcome assessment) and

included screening as a 4th type of assessment. Screening measures "target skills that are

strongly predictive of important future reading outcomes" (Kame'enui et al., 2006, p. 4).

Sometimes referred to as "universal screening," these measures are administered

to all students in a grade level at least at the beginning of the school year and as often as 3

times per year: in the fall, winter, and spring. The purpose is to collect data that teachers,

grade-level teams, or schools can use to decide whether more intensive assessment is

needed to identify students who are at risk of failure on the statewide accountability test.

An aspect of screening is the acknowledgement that students' difficulties may be

unrecognized if teachers do not test for them (Howell & Nolet, 2000).

Examples of individual measures appropriate for screening early reading skills are

phonemic segmentation tasks and nonsense word fluency tasks. After the middle of 1st

grade, measures of oral reading fluency or measures of word list reading skills are

appropriate. Examples of group-administered measures used for screening include silent

reading comprehension, vocabulary tests, and maze tasks. Universal screening is used extensively with beginning readers but, as perhaps evidenced by the absence of such a category in the chart in Table 1, has not been widely used with adolescent readers.

Oral Reading Fluency and Reading Assessment

Standardized assessments of silent reading comprehension, in which students read passages and answer multiple-choice questions, are the most common reading tests taken by adolescents, and in many cases they yield the only available data for decision making. However, listening to students read aloud can provide additional helpful information about a student's reading ability. When collected using prescribed standardized procedures, Oral Reading Fluency (ORF) is an assessment that can be used to quantify each student's rate, accuracy, and expressiveness in oral reading. Students read aloud from an unpracticed grade-level passage for one minute while a trained observer listens and notes any errors. At the end of 1 minute the number of errors is deducted from the number of total words read per minute (TWPM) to calculate a score of the number of correct words per minute (CWPM). A large body of research confirms that ORF is a reliable and valid indicator of reading competence (Fuchs et al., 2001; Shinn, 1998).

Depending on how it is used, and for what purpose, ORF can provide information for several types of assessments. For this reason it is listed as an appropriate measure under more than one category in Table 1. ORF can be conducted for the purpose of providing diagnostic information (Hasbrouck & Tindal, 2006; Hosp & Fuchs, 2005). Analysis of the students' rate of reading and the error markings, recorded as students read, can guide instruction because the data can indicate if a student is making errors on

basic words or only on multisyllabic words, or if the student is reading very accurately but very slowly. Such information has important implications for planning instruction.

ORF can also be conducted to provide information for progress monitoring (Hasbrouck & Tindal, 2006; National Center on Student Progress Monitoring, 2008) when that is the purpose of assessment. ORF is an appropriate measure for determining adequate progress because it is easy and quick to administer, and it is possible to create or utilize existing banks of ORF passages that contain alternate forms. These features make it easy to administer ORF frequently, which is an important requirement for progress monitoring. Also, ORF is sensitive to small changes over time, so progress can be charted and decision rules developed to guide decisions about the adequacy of instruction or intervention.

In addition, Hasbrouck and Tindal (2006) propose that ORF may also be used for the purpose of screening. They provide an analogy of "ORF as a thermometer" to describe how this single brief measure can provide accurate, reliable, and important data as "one indicator of a student's reading ability" (p. 640). The availability of such an indicator enhances the decision-making process for those attempting to determine which adolescents are at risk of needing additional support to meet reading benchmarks. Although ORF must be individually administered, the assessment is easy to administer, relatively brief, and results are a global indicator of a student's reading ability. In order to use ORF for this purpose, scores are compared with established grade-level norms or benchmarks to assist in determining which students may need targeted assistance. Screening using ORF is common at the elementary level.

For students beyond elementary school, screening has been less common. This is in part due to a previous lack of established norms for students beyond 5th grade. Hasbrouck and Tindal (1992) reported ORF norms and percentile scores for students in 1st through 5th grade. However, in their reassessment of national fluency norms, Hasbrouck and Tindal (2006) report norms for a large sample of students through 8th grade and provide guidelines, based on their field experience, for interpreting scores using the ORF norms. They suggest that for students in grades 2-8, teachers should compare the students' CWPM scores to the grade-level norms for the same time of year and consider scores within 10 words above or below the score for the 50th percentile to be within the normal range. The availability of guidelines and norms for students in the middle grades facilitates the use of ORF for screening decisions through 8th grade.

ORF as a Predictor of Reading Comprehension

NAEP

Two large studies have compared students' oral reading ability to their performance on the NAEP test, a standardized reading achievement test conducted at the national level. Pinnell et al. (1995) assessed a random sample of 4th-grade students ($n = 1,136$) comparing oral reading as it relates to reading ability as measured by the NAEP. In this study, oral reading was measured by the components of rate, accuracy, and fluency. Fluency, in this case, referred to prosody or expressiveness of reading. Positive relationships were found between a student's rate, accuracy, and fluency in oral reading and the student's performance of the NAEP assessment.

A follow-up study of 4th-grade students ($n = 1,779$) also compared oral reading and performance on the NAEP (Daane et al., 2005). This study examined the relationship between the same three aspects of oral reading (rate, accuracy, and fluency) and found similar results to the first study. Results indicated that, for the 4th-grade students in the study, accuracy, rate, and fluency are related to each other, and all 3 are related to reading comprehension.

The 2 NAEP studies are not comparable to each other because different reading passages and different administration procedures were used. Furthermore, the studies are not comparable with other studies examining the relationship between oral reading and performance on a standardized reading test because both studies employed procedures not typical of most ORF data collection. In each instance, students had the opportunity to read the passage silently twice before reading it orally. Students read the passage silently during the NAEP assessment one week earlier and were given an opportunity to read the same passage silently again just prior to the oral reading of the passage. In contrast, standardized ORF procedures require that students read an unpracticed copy of a passage. This variance from the standardized administration procedures for ORF makes it inappropriate to compare results of these 2 studies with other studies that utilize standard procedures. In addition, the internal validity of these studies may have been compromised as students were allowed to pre-read the passages, and it may have been the repeated reading that affected the relationship between fluency, accuracy, and comprehension.

However, both studies are of interest because they provide specific information about students' reading performance in comparison with their score on a standardized

assessment of silent reading comprehension. Also both studies examined student performance beyond the traditional reporting of rate, or words per minute, and included the specific measure of accuracy of oral reading. Pinnell et al. (1995) found an overall positive relationship between accuracy, rate, and fluency. This study categorized levels of reading accuracy by grouping the number of total errors, or deviations from print, into the following percentages of accuracy: 99, 97, 96, 94, and less than 94. They reported that 57% of the students read with at least 96% accuracy. The researchers described this as a relatively high level of accuracy in oral reading.

Daane et al. (2005) found that students who read with the fewest errors demonstrated the greatest comprehension as measured by their higher scores on the NAEP reading assessment. In contrast, the occurrence of oral reading errors was negatively related to comprehension. These findings indicate that word identification errors negatively impact comprehension and are consistent with the recommendations made by many experts that accuracy in word identification impacts comprehension and reading performance.

State Assessments

In their rationale for using ORF norms for screening decisions to identify which students are likely to achieve future success in reading and which need extra assistance, Hasbrouck and Tindal (2006) described screening measures as being "developed from research examining the capacity of an assessment to predict future, complex performance based on a current, simple measure of performance" (p. 638). A number of studies have

confirmed that ORF can predict performance on outcome measures such as state assessments.

One of the first studies to examine the predictive ability of ORF in regard to state achievement tests was conducted by Crawford, Tindal, and Stieber (2001), who observed moderate correlations between students' rate of oral reading in 2nd and 3rd grade and their scores on a statewide achievement test of reading taken in 3rd grade. Correlations between scores on the state test and ORF were .60 for students in 3rd grade and .66 for students in 2nd grade. They reported that perhaps the most important finding of their study was that 2nd graders who read at least 72 CWPM and 3rd graders who read at least 119 CWPM in January passed the state test in reading, taken in March, at rates of 100% and 94% respectively. Such data provide clear and important information to classroom teachers.

Stage and Jacobsen (2001) also provided evidence that ORF scores, collected in September, January and May, reliably predicted the reading performance of 173 4th-grade students on the reading portion of the Washington Assessment of Student Learning taken in May. In a study designed to replicate the findings of Stage and Jacobsen, McGlinchey and Hixson (2004) studied extant data, examining the scores from a more diverse sample of 1,362 4th-grade students over 8 years on a different state test. They investigated the predictive ability of ORF in relation to the Michigan Educational Assessment Program reading test taken two weeks later. The correlation between the reading rate and the state test score was .67.

Hintze and Silberglitt (2005) extended previous research using ORF to predict outcomes on state assessments by comparing different statistical approaches to determine cut scores. Using a longitudinal design, they followed 1,766 students from 1st through 3rd grade. They found that ORF scores, collected in fall, winter, and spring of each year, were accurate and efficient in predicting which students were likely to pass the Minnesota Comprehensive Assessment reading test taken in the spring of 3rd grade. ORF scores collected closer to the 3rd-grade state test were more predictive, but ORF scores from 1st grade also predicted a student's future performance on the 3rd-grade state test.

Consistent with results from the study by McGlinchey and Hixson (2004), Wood (2006) also observed a correlation of .67 between 4th-grade students' ORF scores and their performance on a state assessment in his study of students across grades 3, 4, and 5. In addition, ORF was found to add a unique contribution in predicting performance on the state assessment beyond the predictability of the previous year's score on the state assessment. Wood found that ORF predicted performance on the Colorado Student Assessment Program reading test equally well for students in grades 3, 4, and 5.

While the study by Wood (2006) appears to demonstrate that the relationship between ORF and state tests may be consistent through the intermediate grades, a study by Rasinski et al. (2005) indicates that the predictive ability of ORF to state tests may be lower for students at the secondary level. This study, conducted with 303 students in 9th grade, found a relationship of .53 between the students' ORF score in June and their performance on a state reading test taken earlier in the school year

Yovanoff et al. (2005) studied the contributions of vocabulary and ORF in predicting reading comprehension as measured by a district assessment which was correlated with the statewide reading assessment. They used linear regression models predicting comprehension with fluency and vocabulary estimated independently for each grade 4 through 8. They reported that vocabulary and reading fluency explain between 40% and 50% of the variance in reading comprehension. They found that fluency effects diminish over grades but remain significant.

Silberglitt, Burns, Madyun, and Lah (2006) also found that correlations dropped from strong to moderate as grade level increased. Silberglitt et al. reported data for 5,472 students in grades 3, 5, 7, and 8, finding that the correlations were .68, .65, .60, and .50 respectively. The authors did note the limitation that there were fewer 7th-grade students participating in the study than students in the other grades due to the fact that 7th-grade students did not take a state test until spring of 2000. However, 528, the number of students participating from that grade level, is still substantial. In this study, the ORF measure was administered within 2 months of the state test. While ORF still accounted for substantial variance in performance across later grades, the predictive value diminished significantly.

Adding Accuracy to Rate

The predictive value of ORF to outcome measures, such as state tests, appears to decline as students increase in grades (Silberglitt et al., 2006). Some researchers have proposed that the decline in the correlations between ORF and measures of

comprehension is due to the fact that reading growth decelerates as students get older and progress through grades (Chall, 1983; Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993; Fuchs et al., 2001; Hasbrouck & Tindal, 2006). Since students at the 50th percentile in grades 6, 7, and 8 all have average reading fluency rates of 150 CWPM (Hasbrouck & Tindal, 2006), it may be that adding an index of accuracy in oral reading to the measure of rate of reading would add to the predictive validity of ORF to comprehension for students in upper grades. This is because, while CWPM does not include words omitted, or words said in error, the CWPM score alone may conceal information about a student's accuracy in word identification. For example, two 7th-grade students both reading 150 CWPM may have very different profiles. One student may read 150 CWPM with no errors and the other student may read 161 CWPM with 11 errors. Each student exhibits very different reading abilities, but these differences are masked by their equivalent score on ORF.

When looking at CWPM alone for ORF, available information about accuracy is lost. However, since CWPM is calculated by deducting the errors from TWPM, the accuracy calculation is readily available as it is merely the CWPM divided by the TWPM. The resulting percentage of accuracy score may be useful information about a student's reading ability.

Some experts categorize the impact of the number of decoding errors and how they affect reading comprehension using the terms *independent level*, *instructional level*, and *frustrational level*. These levels are determined by considering the words in a text that a student can accurately decode, and are based on the idea that text-reading accuracy

is related to reading comprehension. The number, or percentage, of errors in decoding determines whether a particular text is appropriate or too challenging for a reader. According to Carnine et al. (2004), text read at the independent level contains only words that students have knowledge to decode, instructional level is text that is challenging but manageable for a student because it can be read with few errors, and frustrational level contains difficult words that a student may find unreadable. Kuhn and Stahl (2003) matched independent decoding level to a rate of 95% accuracy and instructional level at 90% accuracy. Other researchers also report that an accuracy rate of 95% is considered as instructional level (Howell & Nolet, 2000; Rasinski et al., 2005).

Three of the previously mentioned studies examined students' accuracy in word identification. However, while the first NAEP study, by Pinnell et al. (1995), the second NAEP study, by Daane et al. (2005), and the study of 9th-grade students by Rasinski et al. (2005), explicitly calculated levels of accuracy in word reading, none of these studies examined whether the predictive ability of ORF would have improved if both rate and accuracy were considered.

The first NAEP study, by Pinnell et al. (1995), reported extensive information about students' levels of accuracy in word identification, and reported a positive relationship among accuracy, rate, and fluency. In the second NAEP study, by Daane et al. (2005), errors, or deviations from print, were categorized into the following ranges by percentages: 100-98, 97-95, 94-90, and less than 90. Significant differences were found between the average scores of students in each accuracy category. The greater the number of oral reading errors made by students, the lower the average scores on the

comprehension assessment. This finding is evidence that accuracy in decoding affects performance on measures of comprehension. However, the NAEP studies were designed only to examine aspects of oral reading that could not be observed from results of the standardized assessment. These studies did not analyze the predictive ability of rate and/or accuracy of oral reading to the standardized assessment, an analysis which would have been of great utility to those attempting to determine the value of ORF as a screening measure.

Rasinski et al. (2005) assessed each student's level of accuracy in word identification in addition to their rate of oral reading. However, they reported only the average for the group, which had an overall level of accuracy at 97.4%. They did not report the impact of accuracy on a student's score on the state test. This analysis may have resulted in more information about the nature of students' reading abilities and perhaps added information about the predictive ability of oral reading to state assessments.

Information about the predictive value of ORF to the increasingly important standardized tests of reading is a high priority for educators working with adolescent readers. Such information would help inform decisions about which students would benefit from interventions to improve basic reading skills. This information is particularly important for students in junior high or middle schools where there may be time to provide intervention and improve students' reading skills prior to beginning high school.

This study investigates the relationship between ORF (rate and accuracy) and a statewide reading assessment for 7th-grade students. In addition, it proposes to determine

whether adding an index of accuracy to a measure of rate improves the predictive ability

of ORF to a state assessment of reading comprehension, and whether these results vary

for certain groups of students. Finally, this study addresses the question of whether rate

and accuracy predict student performance level on the state assessment of reading while

controlling for gender, ethnicity, economic status, and special education identification

classification.

CHAPTER III

METHODOLOGY

In order to investigate the relationship between ORF and a statewide reading assessment for students in grade 7, the following questions are addressed:

1. How are the measures (TWPM, CWPM, and Accuracy) of 7th-grade students' oral reading correlated with each other?

2. What is the relationship between 7th-grade students' oral reading rate, defined as Total Words per Minute (TWPM) and Correct Words per Minute (CWPM), and accuracy and their reading comprehension as measured by the RIT score on the Oregon Statewide Assessment in reading?

3. How does rate of oral reading fluency (TWPM) predict RIT scores on the Oregon Statewide Assessment in Reading?

   a. Does accuracy significantly add to the prediction of RIT scores over and above rate (TWPM)?

   b. Do these results vary by subgroups?

4. Do rate (TWPM) and accuracy predict student performance level categories (*does not meet* vs. *meets/exceeds*) on the Oregon Statewide Assessment in reading while controlling for gender, ethnicity, economic status, and special education identification classification?

Setting and Participants

This research was conducted in a small, suburban district in the Pacific

Northwest. The district served approximately 5,410 students in kindergarten through 12th

grade. The district has two middle schools and two K-8 schools.

Seventh-grade students who took both the fall ORF test and the spring Oregon

State Assessment (OSA) test in 2005-2006 comprised the sample. ORF scores were

collected for 446 students in the second week of September. Enrollment data from the fall

of 05-06 indicate a total of 451 students district-wide in the 7th-grade. It is likely that the

6 students who were not tested on ORF were not in attendance or were not enrolled on

the days during the ORF data collection. Student participation in the OSA was high. The

district tested 99.1% of all 7th-grade students who were enrolled in the district on May

1st, 2006. Of the 446 students who had fall ORF scores, 423 also had an OSA score. The

extant data set did not include demographic data for the 23 students who did not have an

OSA score.

The 423 students who had both ORF and OSA scores comprised the sample,

which included 195 females (46%) and 228 males (54%). Approximately half of the

students (49%) were categorized as economically disadvantaged on the basis of

qualifying for free or reduced lunch. Only two of the students were English Language

Learners (0.5%). Students identified as receiving special education services (for any of

the eligible categories of disability) numbered 82 (19%). Terminology used to describe

ethnicity reflects the terminology used when families were asked to describe the ethnicity

of their student upon enrollment in the district. Families had the option to decline to

provide this information (see Table 2).

Table 2

Sample Demographics Ethnicity

	N	%
African American	4	.9%
Asian Pacific Islander	6	1.4%
Declined	8	1.8%
Latino	50	11.8%
Multi-Ethnic	15	3.5%
Native American	9	2.1%
White	331	78.3%

Measures

The assessments used in this study were the OSA in Reading/Literature taken

between March and May and the test of ORF (NCS Pearson, Inc. [AIMSweb], 2007)

collected the previous September as part of the district-wide data collection that occurs 3

times per year in September, January, and May. Both of these assessments have

established reliability and have been found to be valid for assessing the components of

reading and the constructs they purport to measure. The OSA is a measure of silent

reading comprehension. ORF is a measure of rate and accuracy of oral reading.

Oregon Statewide Assessment

Seventh-grade students take the OSA Grade 7 Reading/Literature Assessment

each year during the testing window, which is open between March and May. According

to the Oregon Department of Education (ODE) Technical Report on Oregon's Statewide

Assessment System, Reliability and Validity, Volume 4, tests developed by The OSA

System adhere to the guidelines provided in the Standards for Educational and

Psychological Testing (1985, 1999) and the Critical Elements identified by NCLB

(Oregon Department of Education, 2007). According to the technical report, all

components, (including reliability, content validity, construct validity, criterion validity,

comparability of scores, and fairness and accessibility) meet these relevant standards for

validity.

Oregon has conducted several studies of reliability, describing that standard errors

of measurement are similar across the short and long versions of the computer-based,

online Technology Enhanced Student Assessment (TESA) and by subgroup (ethnicity,

Limited English Proficiency, and Special Education). Reliability coefficients (Cronbach's

coefficient alpha) for the 2005-2006 7th-grade reading tests were 0.90 for the Long Form

of the test, 0.84 for the Short Form 1, and 0.84 for the short Form 2.

Content validity is the extent to which an assessment accurately represents the

content domain. As evidence of content validity, the ODE reports that test specifications

"provide a clear link between the test content and the content standards and their

corresponding performance levels" and that ongoing studies "evaluate and increase the

extent that instruction, assessments, and the Academic Content Standards are aligned"

(Oregon Department of Education, Technical Report, 2007, p. 11).

The reading passages on the 7th-grade reading test average between 600 and 700

words and represent literary and informative text as well as other items such as graphs,

tables, and charts. Passages are reported to be of high interest and of appropriate readability for 7th-grade students. The test has 7 score categories (word meaning, literal comprehension, inferential comprehension, evaluative comprehension, locating information, literary forms, and literary elements and devices) and a total score.

Students had up to three opportunities to take alternate versions of the computer-based, fully-adaptive test during the testing window. Students encountered either a short or long version of the electronic test. Shorter tests provide a measure of overall performance while longer versions of the test produce an overall score and strand-level information. The accuracy of a student's responses to initial items in the test's multiple-choice format determines the difficulty of the next passage the student will encounter. The computer scans the student response on the "bubble sheet" on screen, and scanned results are scored against an answer key to produce a raw score. The raw score is converted to a scale score based on the number of accurate answers compared to the total items on the test and accounting for the level of difficulty of the question.

A student's highest level of performance was recorded as their score on the test. Students received a Rasch Unit (RIT) score, reflecting their performance in reference to standards based on performance levels associated with cut scores, as having exceeded, met, or not met the benchmark. A score of 226 met the benchmark and a score of 236 exceeded the benchmark.

ORF

The passages used for the ORF assessment were developed by AIMSweb (NCS Pearson, Inc., 2007). They are all narrative passages. Alternate-form reliability,

readability comparisons, and comparisons of means, standard deviations, and standard errors of measurement were used to produce a high level of alternate form reliability at greater than .70 for the passages. These benchmark passages were selected from a pool of 33 passages that met criteria, and the remaining 27 passages were included in a progress monitoring bank. A copy of one of the passages is included (Appendix A).

The directions suggest that one of the three benchmark passages be used each fall, winter, and spring for administration of ORF. The district in which this study took place prefers to give students three passages each fall, winter, and spring and record the median of the three CWPM scores as the student's score for that benchmark period. Therefore, a sufficient number of the grade-level progress-monitoring passages were reserved to augment the three benchmark passages for use in benchmark data collection.

Test Administration

Test administrators followed the administration and scoring procedures described in detail in the AIMSweb Training Workbook (Appendix B). These administration and scoring procedures are aligned with standard practice for the collection of ORF. Testers read the standardized directions to the student and the student reads aloud from his or her copy while the tester records any errors on a numbered copy of the passage that is held discreetly away from the student's view. At the end of a minute the tester records a bracket after the last word read by the student and thanks the student. When the student has read for a minute from each of the three passages, the student is dismissed and the tester calculates the total words read and the errors on each passage and records the student scores on the recording section.

Scoring

A copy of the scoring procedures is included (Appendix C). In addition to recording the student's median score of CWPM (CWPM = TWPM – Errors), the number of decoding errors on that passage is also recorded. From these two scores, the database can also calculate TWPM (TWPM = CWPM + Errors) as well as the percentage of accuracy (Accuracy = CWPM / CWPM + errors or CWPM / TWPM).

Rate Defined

CWPM is the outcome score commonly used by educators and often in practice is referred to as "rate" of reading. However, for research questions 3 and 4 in this study, rate will be referred to as TWPM. This was deemed to be necessary due to the fact that one of the questions of interest is whether adding an index of accuracy increases the predictive value of rate. As CWPM contains an inherent measure of accuracy by definition (CWPM = TWPM – Errors), TWPM, is the preferable measure. Using TWPM reduces the possibility of a linear dependency between the variables of interest.

Procedures

ORF

In mid September, all 7th-grade students were given an ORF assessment. Students read aloud from 3 grade-level passages from the AIMSweb ORF. Students were individually assessed by teachers and educational assistants who had all received training in administration and scoring procedures. This training was conducted for all test administrators prior to each benchmark data collection period. The training is mandatory

for all test administrators in the district. Procedures, scoring, and data entry are taught by a district coordinator or administrator with extensive training and experience in giving the assessment. In the instance of the data collected for this study, the training was conducted by the researcher of this study. Reliability (the number of agreements on correct words per minute divided by the number of disagreements of correct words per minute) was calculated and was high at .98, further indication that the measure is easy to administer accurately. Inter-rater reliability was not formally conducted for this study.

All 7th-grade students in the district were assessed within a one-week window in September. Each school planned testing locations in a quiet place. Data regarding the student's CWPM and number of errors were entered into the district database as soon as possible by a teacher or educational assistant who had training in entering the data.

Oregon State Assessment Test

Students took this computer-based test up to 3 times between March and May of 2006. For students who took the test more than once, only their highest score was reported. The tests were administered by teachers, administrators, or educational assistants who were trained, and adhered to the standardized administration procedures as delineated by the OSA System. Tests were consistent across all classes and schools. As there were no time limits, students were allowed to work at their own pace.

Design

This study examines the correlations between elements of the ORF assessment (TWPM, CWPM, and accuracy). It also examines the relationship between these aspects of a student's rate of oral reading (TWPM, CWPM, and accuracy) and RIT score on the

OSA. Correlations only report the relationship of these variables to the statewide test; one does not cause or predict the other. The next research questions are related to prediction, asking whether oral reading rate is predictive of performance on the OSA, whether the addition of an index of accuracy added to a measure of rate improves this predictability, and whether this varies by subgroups. The final question is whether rate and accuracy predict performance level categories on the OSA into two groups: those who do not meet the standard, and those who either meet or exceed the standard on the state test.

Data Analysis

Pearson's *r* correlation was employed to answer the first two questions regarding examination of the relationship between 7th-grade students' oral reading rate and oral reading accuracy and their reading comprehension (as measured by OSA). Multiple regression analyses were conducted to determine how rate (TWPM), and rate with the addition of accuracy, predict students' RIT scores on the OSA Reading/Literature test, and whether these results vary for subgroups. Finally, a discriminant analysis was conducted to determine whether rate and accuracy predict student performance level categories (*does not meet* versus *meets/exceeds*) on the OSA while controlling for gender, ethnicity, economic status, and special education identification classification.

CHAPTER IV

RESULTS

This study examined the relationship between rate and accuracy of oral reading for students in grade 7. It also examined how rate and accuracy of oral reading correlate with scores on the Oregon State Assessment (OSA). In addition, this study examined whether oral reading rate predicts students' scores on the OSA, whether the addition of an index of accuracy improves the prediction, and if these findings varied by subgroups. Finally, this study examined how well the rate and accuracy of oral reading predict students' performance levels on the OSA while controlling for gender, ethnicity, economic status and special education identification classification.

In the conduct of these analyses, only students who took the ORF assessment in September and also took the state assessment between March and May of the 2005-06 school year were included in the sample of 423 students. There was only one case of missing data, and this case was excluded from all analyses. Error data was not entered for one of the ORF scores, thus scores for 422 students were included in some of the analyses.

An alpha of .05 was used to determine statistical significance. The magnitude of the correlations and other coefficients were interpreted using the following frame of reference, correlation coefficients (Pearson's r): small = 0.10, medium = 0.30, large = .50 (Cohen, 1988).

Descriptive Statistics

Before undertaking the analyses to answer the research questions, information

describing student performance on the measures was examined. The means and standard

deviations, as well as the minimum and maximum scores for each of the measures for the

entire sample are presented in Table 3. As mentioned earlier, error data was not entered

for one of the passages, thus accuracy and error data columns had one fewer score.

The measures of rate, TWPM, and CWPM had similar means, standard deviations,

minimum, and maximum scores. Accuracy is reported as a percentage and denotes the

percentage of words read correctly. Scores on the OSA are reported as RIT scores. The

mean score for the sample was 231. A score of 226, or above, is needed to meet or

exceed the benchmark.

Table 3

Means and Standard Deviations of Student Performance and Minimum and Maximum Scores for Each Reading Measure

	N	Minimum	Maximum	Mean	Standard deviation
TWPM	422	25	259	134.92	42.12
CWPM	423	20	255	131.59	43.08
Accuracy	422	76%	100%	97%	.038
OSA RIT	423	196	265	230.92	9.56

The means and standard deviations for scores on the ORF measures of TWPM

and accuracy as well as for the OSA scores for each of the subgroups and their

counterparts are presented in Table 4. On all three measures, the scores for most of the

groups were fairly close to the mean for the entire sample. For all three measures, the scores for students in the special education group were markedly below the scores for the other groups.

Table 4

Means and Standard Deviations of Measures for Subgroups

Group	N	TWPM	Accuracy	OSA
Female	194	141 (41.62)	97% (3.9%)	233 (9.27)
Male	228	130 (41.90)	97% (3.8%)	230 (9.62)
Other Ethnicities	78	131 (42.69)	97% (3.6%)	230 (10.71)
White	336	136 (42.01)	97% (3.9%)	231 (9.26)
Special Education	82	92 (36.94)	93% (6.1%)	221 (8.05)
General Education	340	145 (36.38)	98% (2.2%)	233 (8.38)
Economically Disadvantaged	208	129 (43.00)	96% (4.2%)	230 (9.33)
Not economically Disadvantaged	214	141 (40.47)	97% (3.4%)	232 (9.62)

In the following section, analyses and results are presented to address each of the research questions. First, the relationship between measures of rate and accuracy will be described. Second, the relationship of rate and accuracy to the OSA will be presented. Third, the predictive value of rate alone, and rate in combination with accuracy, to the OSA will be reported along with information about whether these predictions vary for subgroups. Finally, the question of whether TWPM and accuracy predict student membership into performance levels on the OSA, while controlling for gender, ethnicity, economic status, and special education classification will be addressed.

Relationship between Rate and Accuracy on ORF

To answer the first question regarding how the measures of 7th-grade students'
oral reading are correlated with each other, correlation coefficients were computed for the
relationships between TWPM, CWPM, and accuracy. All correlation coefficients were
positive and statistically significant ($p < .001$). The correlation between CWPM and
TWPM was nearly perfect ($r = .99$), meaning there was negligible difference between the
two measures. The correlation for accuracy and TWPM was .61 while the correlation for
accuracy and CWPM was .65, indicating a moderately strong relationship between the
measures of rate and the measure of accuracy.

Relationship of Rate and Accuracy to the OSA

To answer the second question regarding the relationship between 7th-grade
students' oral reading rate (TWPM and CWPM) and accuracy and their score on the
OSA, correlation coefficients were computed (see Table 5). All measures of oral reading
had statistically significant relationships with RIT scores on the OSA ($p < .001$). Both
measures of rate of oral reading (TWPM and CWPM) had moderately strong positive
correlations ($r = .62$) with RIT scores. The correlation of accuracy with RIT scores was
somewhat lower ($r = .48$).

Table 5

ORF Measures Correlations to OSA RIT Score

Measure	N	Correlation to OSA	Sig. (2-tailed)
Fall ORF TWPM	422	.617	<.001**
Fall ORF CWPM	423	.623	<.001**
Fall ORF accuracy	422	.480	<.001**

Scatter Plots of Correlations between

Measures of Oral Reading

Scatter plots were created to display the nature of the correlation between variables and allow for visual inspection of the data for outliers or other abnormalities. The first scatter plot for TWPM to CWPM had an almost perfect correlation ($r = .99$). Visual examination revealed no apparent outliers or abnormalities (see Figure 1).
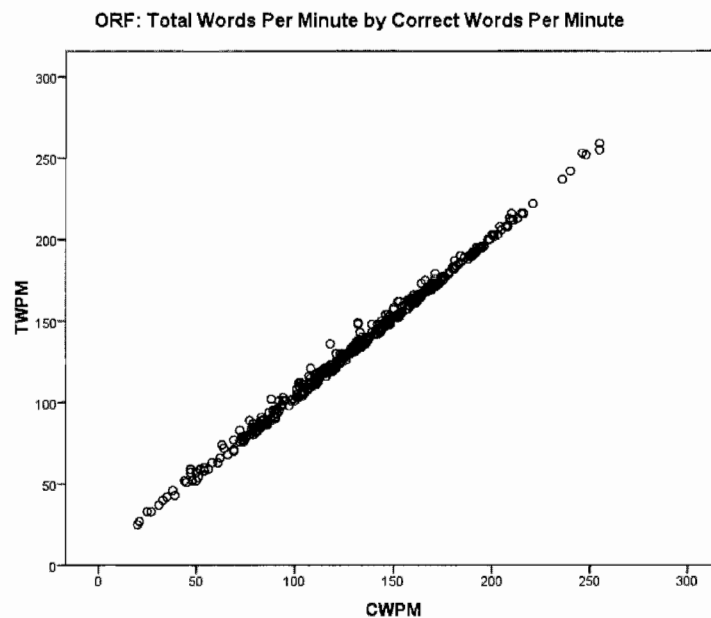
Figure 1. TWPM and CWPM

While the scatter plot for TWPM to CWPM indicated a linear relationship, scatter plots that included accuracy in relation to TWPM and OSA appeared curvilinear in nature (see Figures 2 and 3). The visual display of the correlation between TWPM and accuracy appears to show that less fluent readers generally have lower scores on accuracy and these accuracy scores vary across a range.
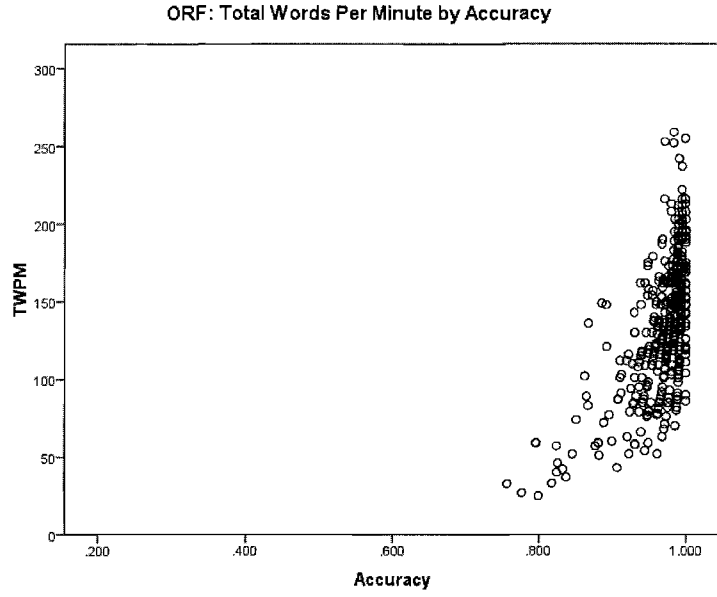
Figure 2. TWPM and Accuracy
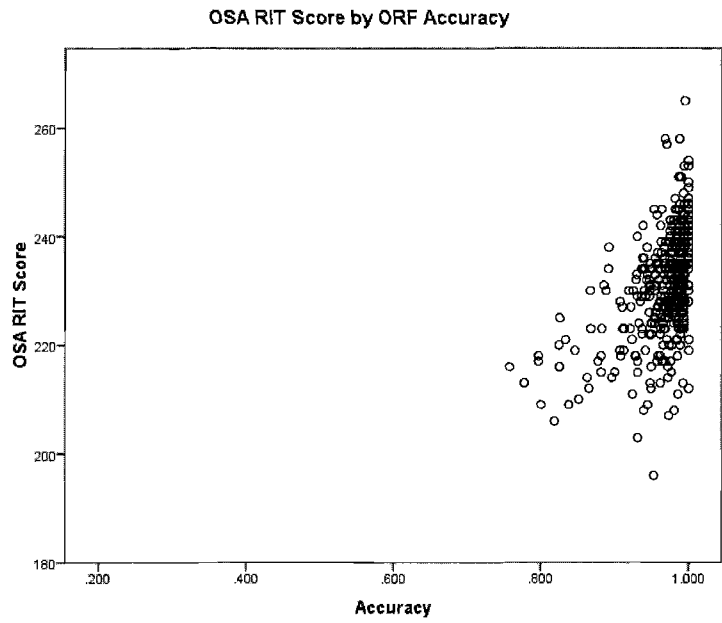
Figure 3. RIT Score and Accuracy

Rate and Accuracy as Predictors of Scores on the OSA

Assumptions and Considerations for

Multiple Regression Analyses

Several multiple regression analyses were conducted to answer the questions regarding whether rate predicts RIT scores on the OSA, whether the addition of accuracy adds to the prediction, and whether these results vary by subgroups. Multiple regression is a data-analytic procedure that determines the linear relationships between a set of predictors and a single criterion (Licht, 2005). Certain methodological concerns and assumptions should be addressed when using multiple regression analysis.

Sample size should be adequate. Gall, Gall, and Borg (2003) suggested that sample size should increase at least 15 subjects for each variable that will be included in the multiple regression analysis. The *n* of 422 for this study is more than adequate for the 8 predictor variables used in this study.

Multicollinearity must also be considered. Predictor variables should not correlate too highly. Licht (2005) proposed that correlations of $r > .80$ between predictor variables may cause technical problems. Correlations between predictor variables in this study do not approach a level of concern. In this study, the predictor variables of TWPM and accuracy indicated a moderately strong correlation ($r = .61$). Also, data should be examined for outliers and abnormalities that may cause difficulties. Scatter plots (Figures 1-3) illustrate that this was not a concern.

Homoscedasticity, or the constant variance of residuals, is another assumption that should be met. This means that variance of residuals around the regression line is

assumed to be constant for all values of X in the population (Cohen, Cohen, West & Aiken, 2003). Inspection of residuals can indicate homoscedasticity. The scatter plot of the standardized residuals by standardized predicted values shows that points appear to be scattered evenly throughout the plot. There was no evidence of heteroscedasticity.

Another assumption that must be met is the independence of residuals. There should be no relationship among the residuals for any subset of cases in the analysis (Cohen et al., 2003). This means that how the model fits for one person does not affect another person. The Durbin-Watson statistic for the overall regression analysis was .889, somewhat below typical rules of thumb and indicating that there was some dependency among residuals. Inspection of the standardized residuals, however, did not show any strong effects of a lack of independence, and it did not appear that this assumption would have a large impact on the reported results.

Finally, outcome variables should be approximately normally distributed and considered for skewness and kurtosis. For the OSA variable, skewness (-.03) and kurtosis (.247) are both within the acceptable range. The assumptions for use of the multiple regression analysis were met.

Potential curvilinear relationships between accuracy and TWPM and accuracy and the OSA required consideration and may have had important implications for the multiple regression analyses required to answer the research questions regarding how rate and accuracy, together and separately, predict scores on the OSA, and whether these results vary by subgroups. Therefore, it was necessary to check for both linear and curvilinear trends in the accuracy variable. A curvilinear effect of accuracy was added as

a predictor to the regression equation. The addition of the curvilinear effect did not appreciably change R Square (.3 of one percent) and was not statistically significant ($p = .125$). As it was determined that the addition of the curvilinear version of accuracy was not statistically significant, only the linear effect of accuracy was used in the remaining multiple regression analysis.

Rate as a Predictor of OSA Score

A multiple regression analysis was applied in two steps. On the first step of the analysis, TWPM was entered as a predictor to the criterion variable, which was the RIT score on the OSA. Results indicate a multiple R value (.617), a moderately strong positive correlation between the predictor variable (TWPM) and the dependent variable (RIT score on OSA). The R Square value (.381) indicates that for the 7th-grade students in this study, approximately 38% of the variance in RIT score is explained by TWPM.

Accuracy Added to Rate as Predictors of OSA Scores

On the second step of the analysis, accuracy was added to the regression model. As shown in Table 6, R^2 increased by .017, which was a statistically significant increase, $F(1, 419) = 11.809, p = .001$. For the students in this sample, approximately 40% of the variance in OSA RIT score is explained by TWPM and accuracy together.

The regression analysis addressed the question of whether rate predicts OSA scores and whether rate and accuracy in combination produce a better prediction of OSA scores. A summary of the results can be seen in Table 7. In this study, accuracy was coded as a proportion because the range for accuracy is 0 to 1. Thus, every 10% increase in accuracy would result in a gain of approximately 4 points on the OSA RIT score.

Table 6

Model Summary

			Change Statistics				
			R				Sig.
		R	square	F			F
Model	R	square	change	change	df1	df2	change
TWPM only	.617[a]	.381	.381	258.360	1	420	<.001
Accuracy (linear) added	.631[b]	.398	.017	11.809	1	419	.001
Accuracy (curvilinear) added	.633[c]	.401	.003	2.359	1	418	.125

Table 7

Coefficients

Model	Unstandardized coefficient B	Unstandardized coefficient std. error	Standardized coefficient beta	t	Significance
Constant	175.523	10.685		16.428	<.001
TWPM	.117	.011	.516	10.781	<.001
Accuracy	40.863	11.891	.165	3.436	.001

Rate and Accuracy as Predictors of

OSA Scores for Subgroups

To determine whether these results differed by subgroup, regression analyses were repeated separately for selected groups of students. Students were grouped by gender, by classification of special education or general education, and by whether or not they were economically disadvantaged (based on their eligibility for free and reduced lunch status).

In addition, a group was included based on ethnicity. The original intention was to examine the subgroups in the required reporting categories designated in the AYP reports for NCLB. However, due to the small size in some of these groups it was not possible to use the exact reporting categories. Therefore it was decided to delineate two groups. White students and the 6 Asian students, who tend to have similar performance, were grouped together, while all other students were designated as other ethnicities. For the number of students in each of these groups, refer to Table 8.

Table 8

Multiple Regression Analysis

| | | TWPM | | Accuracy Added | |
Group	N	F	R square	R square change	F
Gender –Female	194	105.564**	.355	.031	9.618*
Gender – Male	228	141.129**	.384	.010	3.689
Ethnicity – All Other	78	47.430**	.384	.001	.176
Ethnicity – White & Asian	336	201.387**	.376	.025	14.064**
Special Education	82	29.252**	.268	.003	.329
General Education	340	103.858**	.235	.016	7.075*
Economically Disadvantaged	208	135.418**	.397	.006	2.127
Not Economically Disadvantaged	214	113.540**	.349	.038	13.117**

TWPM does predict RIT scores on the OSA and was statistically significant at the $p < .001$ level for all subgroups. The R^2 values range from a low of .268 for students in special education, to a high of .397 for students categorized as economically disadvantaged.

For 4 groups, the addition of accuracy was statistically significant ($p < .05$). These groups were white students, females, students categorized as not being economically disadvantaged, and students in general education. The R^2 change values for all of the groups range from a low of .01 for males to .04 for students not economically disadvantaged. Though the addition of accuracy was statistically significant for the 4 groups mentioned above, the amount of variance for these groups ranged from 1% to 4%. These results indicate that the addition of accuracy for the subgroups is similar to that of the addition of accuracy for all students (2%); it does not add much to the prediction rate established by TWPM.

Rate and Accuracy as Predictors of

Performance Level on OSA

In order to evaluate whether rate and accuracy predict student performance level categories on the OSA in reading, while controlling for gender, ethnicity, economic status, and special education classification, a predictive discriminant function analysis (DFA) was conducted. DFA is a statistical technique that allows one to predict, or classify, the assignment of observations into well-defined groups based on a set of attributes (Licht, 2005). DFA is used when the measure of the criterion variable is

categorical, as in this case, in which the groups are those students who did not meet the standard on the OSA and those who either met or exceeded the standard.

On the basis of their RIT score, the OSA categorizes students into one of three groups: does not meet, meets, or exceeds the state standards. However, for the purpose of this analysis, two groups were formed based on the critical cut score of 226. Students who scored 226 or above were classified into a *meets or exceeds* group, and students scoring below 226 were classified into a *does not meet* group. These were the two performance levels used in the analysis.

The group of six predictors used in the DFA were significantly related to the two proficiency groups, χ^2 (6) = 196.52, p = .001, Λ = .624. The value of Wilk's lambda (1 - Λ) indicates that approximately 38% of the variation in the groups was related to the weighted linear combination of the 6 predictor variables.

The structure matrix and standardized function coefficients (Table 9), produced by the DFA, show the relationship between each of the discriminating variables and the discriminant function. The pattern is similar for both the structure coefficients and the standardized function coefficients so the structure coefficients are discussed. The structure coefficients show that the discriminate function is most related to a student's special education classification (r = -.78), TWPM (r = .77), and accuracy (r = .70). The strong positive correlations for TWPM (r =.77) and accuracy (r =.70) indicate that knowing a student's score on these measures is highly predictive of performance level and that as TWPM and accuracy increase, so does the likelihood of being in the *meets/exceeds* group. The strong negative correlation for special education classification

($r = -.78$) indicates that a student's classification of being in special education is highly predictive of performance level and that there is an increased likelihood of being in the *does not meet* group. Gender, economic status, and ethnicity contributed little in predicting performance level.

Table 9

Structure Matrix and Standardized Function Coefficients

	Structure coefficients*	Standardized function coefficients
Special Education	-.776	-.541
TWPM	.765	.433
Accuracy	.696	.310
Gender	-.178	-.099
Economically Disadvantaged	-.118	-.029
Other Ethnicities	.076	.155

The means and standard deviations for the same OSA proficiency categories used in the DFA were examined for the TWPM and for accuracy. Table 10 shows that the mean rate for the *does not meet* group was 97 TWPM, while the mean rate for the *meets/exceeds* group was 147 TWPM. In addition, the mean score for accuracy for the *does not meet* group was 94%, while the mean score for accuracy for the *meets/exceeds* group was 98%.

The classification results (Table 11) show that, for this sample, the predictor variables correctly classify 82% of the students in regard to performance level on the OSA. Classification of students into the *meets/exceeds* group membership was higher at a

Table 10

Means and Standard Deviations of Measures for Performance Levels

Performance level	TWPM	Accuracy
Does not meet	97 (39.38)	.94 (.057)
Meets/exceeds	147 (35.19)	.98 (.021)

rate of 88%. The group that did not meet the state standard was predicted 62% of the

time. The rates of misclassification were 12% for students in the *meets/exceeds* group and

38% for students in the *does not meet* the standard group. Predictions were more accurate

for students in the *meets/exceeds* performance level.

Table 11

Discriminant Function Analysis – Classification Table

Performance level	Predicted group membership – DNM	Predicted group membership – M/E	Total
% Does not meet	62.1	37.9	100.00
% Meets/exceeds	11.6	88.4	100.00

82% of original grouped cases correctly classified.

CHAPTER V

DISCUSSION

For those working to improve adolescent literacy, one of the challenges is the scarce availability of technically adequate measures designed to inform decisions about when adolescents have acquired adequate basic reading skills. Fuchs et al. (2001) suggested that extending the research base by identifying acceptable oral reading rates by grade and identifying information about performance levels, by grade, that predict success on standardized tests of reading would provide information for determining which adolescents require intervention. Having this data early in the school year would provide educators much needed information for making decisions about which students are likely to require additional support and instruction. Snow, Burns and Griffin (1998) reported that students who do not acquire such basic skills are at risk for developing reading difficulties. Secondary students who lack these skills will continue to struggle in reading unless they are identified so that they can receive intervention. This study examined 7th-grade students' performance on a fall Oral Reading Fluency (ORF) measure in relation to their performance on a spring state test of silent reading comprehension.

In this chapter, the results for each research question are summarized and interpreted. Next, implications for practice are discussed. Finally, study limitations and recommendations for future research are presented.

Summary and Interpretation of the Findings

Measures of Oral Reading – Rate and Accuracy

The relationship between the rate and accuracy resulting from the ORF assessment was analyzed. Rate of oral reading was measured in two ways, as Total Words per Minute (TWPM) and as Correct Words per Minute (CWPM). The standard measure used in the field to report ORF is in CWPM, which is a count of the words a student reads correctly in one minute. CWPM is calculated by deducting the number of errors from the TWPM (which is the total words included in the passage at the point where the student stops reading at the one minute mark). Accuracy is calculated by dividing the CWPM by the TWPM, resulting in a percentage score.

The correlations among the three measures of ORF were positive and statistically significant. The correlation between CWPM and TWPM was very strong ($r = .99$). There was a moderately strong correlation between accuracy and both CWPM ($r = .65$) and TWPM ($r = .61$). As CWPM and TWPM are very similar, it was expected that these correlations would be high. Accuracy had a weaker correlation than the measure of rate. One reason was likely the result of the constricted range of scores for accuracy. Scores ranged from 76% to 100%, while the range of scores for rate had more variation. TWPM scores ranged from a minimum of 25 to a maximum of 259. CWPM scores ranged from 20 to 255. Another reason for the weaker correlation may be that while fluent readers tend to be more accurate readers, this is not always the case. For all students there is a range of accuracy in reading, with the largest variation occurring for less fluent readers.

Though the reliability and validity of the ORF assessment is well established (Fuchs et al., 2001; Shinn, 1998), the correlations between the measures that result from the ORF assessment (rate and accuracy) were examined in this study as a precursor to the multiple regression analysis. This is because rate and accuracy were constructs of interest in this study.

Relationship of Rate and Accuracy to the OSA

Correlations for rate and accuracy to the OSA were computed to determine the nature of the relationship between aspects of oral reading and silent reading comprehension. Both measures of rate (CWPM and TWPM) had statistically significant and moderately strong positive correlations with the OSA measure ($r = .62$). This result is very similar to the finding by Silberglitt et al. (2006), who also studied 7th-grade students ($n = 528$), and reported a moderately strong positive correlation ($r = .60$) between a measure of ORF and scores on a state test. Also, this finding is solidly located within the range of findings from other studies in which correlations found between ORF and state tests for students in grades 2 through 7 were all between .60 and .67 (Crawford et al., 2001; McGlinchey & Hixson, 2004; Silberglitt, et al. 2006; Wood, 2006).

Some researchers have proposed that the correlation between ORF and measures of comprehension declines as students advance in grade level due to the fact that reading growth decelerates (Chall, 1983; Fuchs et al., 1993; Fuchs et al., 2001; Hasbrouck & Tindal, 2006; Yovanoff et al., 2005). Silberglitt et al. (2006) reported a correlation of .50 for 8th-grade students and Rasinski et al. (2005) reported a correlation of .53 for 9th-grade students. The present study indicates that for 7th-grade students, ORF provides

important information for intervention planning and decision making. Therefore, more studies are needed to understand the relationship between ORF measures and measures of comprehension. Specifically, more studies are needed to examine whether ORF provides useful information even if a slight decline in correlation to the OSA is substantiated.

Accuracy was found to have a statistically significant, modest correlation to the OSA ($r = .48$). The more specific question in this study regarding accuracy was whether the addition of an index of accuracy to a measure of rate would increase prediction of students' scores on the OSA.

Rate, Accuracy, and the Prediction of Scores on the OSA

Several multiple regression analyses were conducted to determine the answer to 3 related questions: (a) how rate alone predicts score on the OSA, (b) how rate with the addition of accuracy predicts students' scores on the OSA, and (c) how these results varied by subgroup. The results for each of these questions are discussed below.

Rate as a Predictor of OSA

The first question was whether rate predicts students' scores on the OSA. Results indicate a moderately strong positive correlation ($r = .62$) between the predictor, TWPM, and the score on the OSA. Thus, for 7th-grade students in this study, approximately 38% of the variance in the spring OSA score is explained by performance on the fall ORF. This may be considered a substantial amount in light of the multitude of factors that contribute to a student's performance on a measure of silent reading comprehension and the length of time between the fall ORF and the spring OSA.

For older readers the factors that contribute to reading performance are even more varied than they are for younger readers. Researcher Catherine Snow, in *Reading Next: A Vision for Action and Research in Middle and High Schools* (Biancarosa & Snow, 2004), states that for secondary students, reading skills are "more complex, more embedded in subject matters, and more multiply determined" (p. 2) than for young readers. This description is particularly true of the reading skills assessed on the OSA, which include reading, analyzing, and responding to multiple-choice questions regarding various types of text and subject matter. In addition, rate is deemphasized by the administration procedures, which allow students an unlimited amount of time to complete the test.

Knowing that potentially 38% of the variance in performance on the OSA is predicted by ORF, a one-minute measure of TWPM, is important formative information for educators attempting to improve outcomes for adolescent readers.

Rate and Accuracy as Predictors of the OSA

The second question was whether adding an index of accuracy to a measure of rate improves the predictive value of ORF to the OSA. Results indicate that the addition of accuracy to the regression model produced an R^2 increase of .017 which was statistically significant. Therefore, TWPM and accuracy combined account for 40% of the variance in performance on the OSA. While the increase with the addition of accuracy was statistically significant, it may or may not be educationally or practically significant.

Data regarding accuracy is currently available for each ORF score, as it is part of the calculation to determine CWPM, the standard format for reporting ORF scores. The

contribution of the accuracy data to the prediction of OSA score could be quite modest and still valuable because the information is readily available and therefore does not require additional assessment. However, the additional 2% of variance found for the addition of accuracy to the prediction model, may, or may not, be useful enough to warrant the data entry required to calculate the addition of accuracy for the purpose of prediction alone.

There may, however, be reasons other than the calculation of prediction to enter the number of errors, or the percent of accuracy, into the data base. The number of errors a student makes, or the percentage of accuracy with which they read, has potential implications for planning instruction, and for that reason some educators enter this information into a database along with CWPM score for rate of reading. Thus, since error information is collected as part of the ORF assessment and errors must be counted to determine the calculation of CWPM, the additional step of entering the number of errors or the percentage of accuracy along with the CWPM into the data base is probably worthwhile for instructional planning regardless of whether it is used to calculate prediction.

Rate and Accuracy as Predictors of the OSA by Subgroups

The third part of this question was whether the findings about rate and accuracy in regard to predicting scores on the OSA varied by subgroups of students. Students were grouped by gender, by classification of special education or general education, by whether or not they were economically disadvantaged (based on eligibility for free and reduced lunch status), and by ethnicity. The original intent was to group students by the

categories of ethnicity used to report scores for NCLB. However, due to small group size in some of these categories, students were divided into two groups, (a) white students and Asian students, who were performing similarly, and (b) all other students, who were designated as other ethnicities.

The result of the multiple regression analyses indicated that TWPM does predict OSA scores and was statistically significant ($p < .001$) for every subgroup. The finding that rate predicts OSA scores for all subgroups and is statistically significant for all subgroups provides further evidence that knowing students' ORF scores in the fall can help inform decisions regarding which students may need intervention.

While the addition of the accuracy score to the regression model produced a statistically significant increase of about 2% for the sample as a whole, adding accuracy to the regression model was statistically significant ($p < .05$) for only 4 of the subgroups. These groups were white students, females, students categorized as not being economically disadvantaged, and students in general education. This finding was perplexing, as it had been expected that the addition of accuracy to the regression model might make more of a difference for groups with lower rates of oral reading. Instead, the four groups for which the addition of accuracy to the regression model was statistically significant were the 4 groups with the highest average oral reading rates.

Though the addition of accuracy was statistically significant for these 4 groups, the educational significance of the added amount of variance (1% to 4% overall) may not be of practical importance. One possible reason that the addition of accuracy was not statistically significant for the groups with lower scores for rate may be that for these

students rate alone is a sufficient predictor, so while the addition of accuracy adds to the prediction, that addition is fairly small. Regardless, the addition of accuracy for subgroups is similar to the findings for the entire sample.

Predicting Performance Levels on the OSA

The predictive discriminant function analysis (DFA) was conducted to evaluate whether rate and accuracy of oral reading predict performance categories on the OSA reading test, while controlling for gender, ethnicity, economic status, and special education classification. Students were grouped into one of two groups on the basis of their score on the OSA: (a) students categorized as *does not meet* the state standard and (b) students categorized as *meets/exceeds* the state standard.

The 6 predictors used in the DFA (TWPM, accuracy, gender, ethnicity, special education, and economically disadvantaged) were significantly related to the 2 proficiency groups, χ^2 (6) = 196.52, p = .001, Λ = .624. The value of Wilk's lambda (1 - Λ) indicates that approximately 38% of the variation in the groups was related to the weighted linear combination of the 6 predictor variables. The discriminant function is most strongly related to a student's special education classification (-.77), TWPM (.77) and accuracy (.70). Gender (-.178), economic disadvantage (-.118), and ethnicity (.076) were not strongly related to the discriminate function.

Good News for Closing the Achievement Gap

The finding that gender, economic disadvantage, and ethnicity were not strongly related to the discriminant function and that TWPM and accuracy are strongly related to the discriminant function should be good news for those working to close the

achievement gap between minority and economically disadvantaged students and their peers. This is because there is evidence that the achievement gap results from the "lagging literacy skills" (Snow & Biancarosa, 2003, p. 18) of students of color and poverty, and that "shrinking the gap will require improving literacy instruction for these groups in particular" (p. 31). Therefore, evidence that increasing rate and accuracy in reading may improve performance on a grade-level test of silent reading comprehension has important implications for closing the gap. Educators can provide instruction and intervention for students needing to improve their basic reading skills and help them increase their accuracy in decoding and their rate of reading.

In addition to TWPM and accuracy, identification as a student receiving special education was also found to be significantly related to the discriminant function. As a group, students in special education typically perform substantially below their peers by definition; therefore this result is not surprising; however, the findings are not without potential positive implications. That TWPM and accuracy were equally, or almost equally, as related to the discriminant function as special education classification is promising, in that instruction can be provided to increase students' performance on these two important variables.

Further Information about Rate and Accuracy

The DFA also produced group statistics regarding the average performance on rate and accuracy for students in the two performance levels. Students in the *meets/exceeds* performance level read an average of 147 TWPM and had an average of

98% accuracy. Students in the *does not meet* performance level read an average of 97 TWPM and had an average rate of accuracy of 94%.

In regard to rate of reading, the finding that students in the higher performance level are reading on average 50 TWPM faster than the average student in the lower performance level has important implications. According to Pikulski and Chard (2005), students with lower fluency scores are spending more time on decoding, and this negatively impacts their comprehension. In addition, these students may take 50% more time to complete each assigned reading task. The potential cumulative effect of these two factors may explain why many struggling adolescent readers fall behind in content classes or eventually drop out of school.

It may seem that the average scores of the two performance levels for accuracy in decoding were more similar, at 94% and 98%. However, there is evidence to suggest that these scores may also imply quite different levels of reading skill. Although the percentages cited vary slightly, researchers suggest that small differences in the percentage of accuracy in decoding indicate very different levels of reading comprehension. Students able to read text with 99% - 100% accuracy are said to be at the *independent* level (Carnine et al, 2004; Johnson, Kress, & Pikulski, 1987). Students who read text at 95% - 98% accuracy are said to be appropriately placed in their *instructional* level (Howell & Nolet, 2000; Johnson et al, 1987; Rasinski et al., 2005). Students who read text below 90% accuracy are said to be at the *frustrational* level in that material and are unable to successfully respond even with support from a teacher (Johnson et al., 1987). While there are gaps between the percentages reported by researchers and slight

differences among researchers in percentages defining the ranges, this study supports the notion that students need to be near the independent level of accuracy to pass the OSA. The average scores for percentage of accuracy in decoding for the two performance levels may support the claim that in order for students to comprehend well they need to have a very high degree of accuracy in the material they are reading. Incorrect identification of one or two key words in a passage could greatly alter a student's ability to understand the meaning of the passage as a whole, for example, substituting the word 'tourism' for 'terrorism'.

Classification Result

Finally, the DFA produces a classification result which shows that, for this sample, knowing students' TWPM and accuracy scores correctly classified 82% of the students in regard to their performance level on the OSA. Predictions for students in the *meets/exceeds* group was quite high with 88% of cases correctly classified, while prediction for students in the *does not meet* group was 62% correct. Conversely, only 12% of cases were misclassified in the *meets/exceeds* group, but 38% of cases were misclassified for the *does not meet* group.

The higher rate of classification error for cases in the *does not meet* group may be a result of the fact that ORF is a timed measure but the OSA is not timed. Students are given extended time (as much time as they need) on the state test, and often across multiple days. Additional time is more likely to benefit students for whom rate is a problem. Perhaps one reason that 38% of students predicted to be a member of the *does*

not meet group actually did meet or exceed the standard may be related to the additional time provided on the state test.

Another reason that may have contributed to the higher misclassification rate for the *does not meet* group was the fact that students were provided up to 3 opportunities to take the test. For students who took the OSA more than once, their highest score was recorded as the result. Perhaps if all students' scores on their first OSA attempt had been recorded, the accuracy in classification would be higher for students in the lower performance level group. As it is, if a student failed the first 2 attempts and passed the third, the prediction may have been right 2 out of 3 times. Unfortunately, since only the students' highest score on state test is known, this potential explanation cannot be verified.

Nevertheless, it is noteworthy that overall 82% of the students were correctly classified into the correct performance level based on TWPM and accuracy on the September ORF score. That a 1-minute measure in the fall has the potential to correctly predict 82% of the students' membership into performance level group on the state test in the spring is remarkable. This is especially true in light of the fact that ORF may be thought of as measuring the lower level, or basic reading skills, of accurate decoding and fluency. These are the basic reading skills described by Chall (1983) as levels 1 and 2, reading skills associated with decoding and fluency of reading. The state test, on the other hand, is designed to measure higher level reading skills such as vocabulary and comprehension, which Chall describes as levels 3 and 4.

Educators attempting to intervene for adolescent readers at risk of not meeting or exceeding state tests of reading comprehension are interested in assessment information related to the interface between levels 2 and 3 on Chall's (1983) developmental model of reading. There is a need to determine when adolescents have the sufficient basic skills associated with level 2. Assessment data that can inform decisions in the fall about which adolescents would benefit from a reading intervention to improve basic reading skills would be most welcome. The results of the DFA provide promising evidence that further research may produce this much needed information.

Implications for Practice

Educators in secondary schools across the country see firsthand the challenges faced by the many students who appear to have inadequate reading skills. Earnest in their desire to improve literacy levels, and to provide specific reading instruction to adolescents in need, educators are often daunted by the task and uncertain of how best to address these concerns. The numerous reports that have recently addressed this issue have offered little in the way of specific guidance regarding how to determine which students need which type of instruction. Educators need to know specifically which students need further instruction in decoding and fluency and which students have acquired sufficient basic reading skills to benefit from instruction in the myriad of comprehension and content literacy strategies available.

The results of this study indicate that ORF may still provide valuable information about rate and accuracy of reading for adolescent readers. ORF measures rate and

accuracy and together these two variables account for 40% of the variance on a silent reading test intended to measure comprehension. Licht (2005) reported that drawing a conclusion about the utility of being able to predict a certain percentage of the variance in the criterion, and the corresponding inability to predict the remaining percentage of the variance "is a judgment that must be made on the basis of the consequences of the various outcomes in applied settings" (p. 30). In the context of secondary schools, knowing that 40% of the variance in performance on the state test of reading can be attributed to rate and accuracy, data produced by a brief and easily administered measure, has high utility. There are numerous implications for practice regarding rate and accuracy. First, rate and then accuracy are discussed.

Rate

That rate alone accounts for 38% of the total variance on the state test is remarkable considering that the state test of reading is not intended as a test of basic reading skill, and is even more notable in light of the fact that ORF is a timed 1-minute measure while the state test is much more lengthy and is not timed. In fact, rate is virtually open ended on the state test as students may take all the time they need, over a number of days, to read and respond to the test. In most practical applications of reading skill in school and in the workplace it is seldom that individuals are allowed to take all of the time they need to read, comprehend, and respond. Thus, rate may even be more important in life than on the state test.

In addition to the predictive value of rate, there are other findings in this study with implications for practice. Rate is certainly important for success in secondary

content classes, as students are assigned reading in most classes. The data regarding the differences in the average score for rate between the *does not meet* group, at 97 TWPM, and the *meets/exceeds* group, at 147 TWPM, provides some insight into student performance in secondary classrooms. Students with the average score for rate in the *does not meet* group need to spend 50% more time on every reading assignment than students with the average score of the *meets/exceeds* group. Providing secondary teachers with each student's ORF score at the beginning of the year would potentially allow teachers insight into student behavior on work performance and work completion.

Accuracy

While accuracy contributed only a small amount to the predictive value of rate, results of this study indicate that it still may be useful to record and enter the number of errors a student makes on the ORF passage as well as the percentage of accuracy in decoding. This is because the data regarding the percentage of accuracy for the 2 performance levels on the state test may support the research regarding independent, instructional, and frustrational reading levels. Students in the *meets/exceeds* group read with an average of 97% accuracy, which is consistent with the literature on this topic. Students in the *does not meet* group read with an average of 94% accuracy which may indicate that the frustrational level should be set even higher than the 90% that has sometimes been reported.

Another benefit of the information regarding accuracy provided by the ORF measure is that educators may use the percentage of accuracy as a screener. For students below a certain percentage of accuracy, it may be useful to examine the actual protocol to

review the types of decoding errors a student made. This would provide insight into the types of diagnostic assessments or instructional programming the student may need.

Because accuracy adds something to the predictive value of rate, because the percentage of accuracy in decoding appears to provide insight into a students' reading level, and importantly, because this data is readily available as part of the ORF assessment procedure and calculation already, this data should be entered and recorded in addition to a student's score for rate.

Limitations of the Study

Research Validity

External Validity

A convenience sample of one grade level of students in one district during one school year has limited generalizability. This affects the external validity of the study. Readers are cautioned about generalizing the results of this study to other sample populations.

Internal Validity

One potential threat to internal validity is that, while all students took the ORF assessment within the same 5-day window, the OSA window was wider, and there was a span of several months during which students may have completed the online assessment. Students who took the test at the end of the window may have had more instruction than those taking the test at the beginning of the testing window.

Another potential limitation may be a possible selection bias. Students included in the sample were enrolled in the district for at least the 6 to 8 month period between the 2 tests. Twenty-three students with only the fall ORF score were excluded from the study. Therefore it may be that mobility played a role in the exclusion of these students. Given the nature of the extant data set used in this study, it is not possible to identify specifically reasons students did not have both test scores. However, it is possible that students who move frequently may be underrepresented in this sample.

It may also be problematic that the students were allowed to take the OSA test 3 times and record only their highest score. This limits comparison of the results to states that also use this practice. This practice may also have affected the findings, as it is not known if students failed the test twice before passing. In addition, we are comparing the scores of students who took the OSA only once with those who may have taken it 3 times.

Finally, AIMSweb provides only one ORF passage for September screening, and the score on that passage is recorded. The district in which this study took place administers 3 different passages, from the same grade-level bank of passages, to each student for the September ORF data collection. The median score is reported and entered into the database. While this presents a slight deviation from the practice described by AIMSweb, administering 3 ORF passages and taking the median score is widely accepted practice and may be a better indicator of a student's oral reading ability than taking the score from only one passage.

Construct Validity

Reading is a very complex construct, and the correlation of .62 between ORF and OSA necessitates consideration of issues regarding construct validity. Construct validity is the degree to which inferences from a test score accurately reflect the concept that the test purports to measure (Gall et al., 2003). ORF is a measure of rate and accuracy in oral reading, while the OSA is a silent reading test intended to measure reading comprehension. However, reading comprehension is inextricably linked to the ability to decode words accurately and automatically (Fuchs et al., 2001). Pikulski and Chard (2005) reported that fluency, which is comprised of rate and accuracy, has a reciprocal causal relationship with reading comprehension and "fluency depends on and typically reflects comprehension" (p. 517). The precise relationship and interaction between fluency and comprehension is not yet understood. More research is needed to examine how fluency and comprehension influence each other (Stanovich, 2000; Wood, 2006) and to what extent ORF and OSA measure the same construct. Therefore it is prudent to be cautious in interpreting the relationship between these aspects of reading.

Error Types

The definition of accuracy used in this study was simply CWPM divided by TWPM. Accuracy was reported as a percentage. Errors were determined according to the standardized procedures for collecting ORF data. However, there was no further evaluation of the types of errors students made while reading.

It is possible that evaluation of the types of decoding errors made by students may have shed more light on the relationship between accuracy of decoding and reading

comprehension. This is because some errors may negatively impact comprehension more than others. The two previously mentioned studies that compared 4th-grade students' oral reading to their performance of the NAEP, did examine accuracy in regard to error types (Daane et al. 2005; Pinnell et al. 1995).

These two studies classified errors in two categories: meaning-change errors and non-meaning-change errors. According to Daane et al. (2005), errors were evaluated for their potential to change the meaning of a passage. They give the example that substituting 'pony' for 'horse' may not change the meaning of a passage but substituting 'house' for 'horse' certainly might. Both studies reported that the occurrence of oral reading errors, regardless of effect on meaning, negatively impacted comprehension, and when only meaning-change errors were considered, students with higher average accuracy rates had higher average scores on the NAEP.

According to Archer et al. (2003), adolescent readers often make decoding errors on multisyllabic words that carry meaning. Errors of this type would seemingly affect comprehension more than other types of decoding errors such as omitting an article or changing an ending on a word in a sentence. This study did not evaluate the types of decoding errors made by students. This should be considered a possible limitation of this study because classification of errors into meaning-change and non-meaning change types may have affected results.

Directions for Future Research

Though accuracy added little to the predictive value of rate in knowing a student's score on the state test, accuracy was an important variable in the DFA. The results for this analysis indicated that rate and accuracy correctly classified 82% of the 7th-grade students into either the *does not meet* or the *meets/exceeds* groups on the OSA. For students in the *meet/exceeds* group the classification was even higher at 88%. This is an important finding that warrants further research. If educators can know, based on a brief measure in the fall, which students are likely to pass the state reading test and which are not, they can make far better decisions about instructional programs and planning.

Future research should focus on improving the classification rate for the DFA, particularly for the 38% of students who were predicted to be in the *does not meet* group but ended up in the *meets/exceeds* group. In this instance, false negatives are better than false positives, because more students may receive additional help. Improving the classification result will facilitate better decisions about the use of such resources and benefits. One avenue to be investigated is whether using a spring ORF score from the previous year instead of, or in addition to, the fall ORF score improves the classification. Students' spring ORF scores are thought to be a more stable measure of students' reading ability because the fall ORF score may be deflated for students with lower reading skills due to a lack of reading over the summer months. Future research should also investigate whether the results of the DFA are replicated in other samples of 7th-grade students and if results are similar for adolescents in other grade levels.

Future research should also focus on examination of the results of the scores included in the DFA, particularly the interaction between rate and accuracy and their impact on the probability of passing the OSA. Knowing the probabilities associated with being in the *meets/exceeds* group based on scores for rate and accuracy would provide educators a powerful tool for decision making regarding identification of students in need of intervention. This information would provide an important first step towards closing the achievement gap and helping all students gain the literacy skills they need to be successful in school and in life.

APPENDIX A

SAMPLE READING PASSAGE FOR ORF
GRADE 7

Josh walked out of camp that morning into a forest that was perfect for deer hunting. The air was cool and damp, and the forest floor was quiet for walking. Josh took a deep breath of the late fall air and knew that today was his kind of day. He turned east off the old logging road leading from the cabin and headed towards Big Bay ridge. He noticed quite a few deer tracks in the soft forest floor and discovered a sapling near the trail that had been rubbed by a large buck.

Upon reaching the ridge, he settled down into a spot that promised some good action. The white-tailed deer were apt to move this morning, and the promise of deer activity excited Josh. As he carefully kept watch, his mind raced through previous hunts where the conditions were similar. He had been successful on several of those hunts. Today's conditions, however, spelled trophy. His senses were keen and as sharp as the newly purchased hunting knife that hung from his belt.

Josh waited in complete silence looking for any sign – a flick of an ear, tail, or anything that didn't look just right. Concentrating on the hunt was not always easy because his mind wandered from time to time and small things, like a chipmunk playing in the leaves, distracted him…

(The passage continues for a total of 360 words.)

APPENDIX B

STANDARDIZED DIRECTIONS FOR ORF ASSESSMENT

Directions for a 1-Minute Administration of Oral Reading Fluency Assessment

Materials:

1. Unnumbered copy of the passage (student copy)
2. Numbered copy of the passage (examiner copy)
3. Stopwatch
4. Tape recorder (optional – tape recorders facilitate error analysis)

Directions:

1. Place the unnumbered copy in front of the student.

2. Place the numbered copy in front of you but shielded so the student cannot see you record.

3. Say these specific directions to the student for each passage: "When I say 'begin' start reading aloud at the top of this page. Read across the page (DEMONSTRATE BY POINTING). Try to read each word. If you come to a word you don't know, I'll tell it to you. Be sure to do your *best* reading. Are there any questions?" (Pause)

4. Say "Begin" and start your stopwatch when the student says the first word. If the student fails to say the first word of the passage after 3 seconds, tell him, or her, the word and mark it as incorrect, then start your stopwatch. (On rare occasions the student may "speed read" – read the passage very fast and without expression. If this occurs, tell the student, "This is not a speed reading test. Begin again, and be sure to do your best reading.")

5. Follow along on your copy. Put a slash (/) through the words read incorrectly (see scoring procedures).

6. If a student stops or struggles with a word for 3 seconds, tell the student the word and mark it as incorrect.

7. At the end of 1 minute, place a bracket (}) after the last word and say, "Stop."

APPENDIX C

SCORING RULES FOR ORF

Scored As Correct:

- A word must be pronounced correctly, in accordance with the context of the sentence.

- Repetitions: Words that are repeated or inserted are ignored.

- Self-Corrections: Words misread initially, but corrected within 3 seconds, are scored as correct.

- Dialect/articulation: Variations in pronunciation explainable by local language norms or speed sound production are correct.

Scored as Errors:

- Mispronunciations/word substitutions: Words either pronounced or substituted for other words are errors.

- Omissions: Each word omitted is an error.

- Hesitations: When a student hesitates or fails to correctly pronounce a word within 3 seconds, the student is told the word and an error is recorded.

- Reversals: When a student transposes two or more words, those words not read in correct order are errors.

Special Scoring Examples:

- Numerals: Numbers written as numerals are counted as words and must be read correctly within the context of the passage.

- Hyphenated words: Each morpheme separated by a hyphen(s) is counted as an individual word if it can stand alone.

- Abbreviations: Abbreviations are counted as words and must be read correctly within the context of the sentence.

- Insertions: If a student adds extra words, the words are not counted as correct words or as reading errors.

BIBLIOGRAPHY

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1984). *Becoming a nation of readers: The report of the Commission on Reading.* Washington, DC: National Institute of Education.

Archer, A. L., Gleason, M. M., & Vachon, V.L. (2003). Decoding and fluency: Foundation skills for struggling older readers. *Learning Disability Quarterly, 26*(2), 89-101.

Armbruster, B. B., & Anderson, T. H. (1988). On selecting considerate content textbooks. *Remedial and Special Education, 9*, 47-52.

Barton, P. E. (2000). *What jobs require: Literacy, education, and training, 1940-2006.* Washington, DC: Educational Testing Service. http://www.ets.org/medical/ Research/pdf/PICJOBS.pdf

Behavioral Research and Teaching. (2005). *Oral reading fluency: 90 years of assessment* (Tech. Rep. No. 33). Eugene: University of Oregon.

Berman, I., & Biancarosa, G. (2005). *Reading to achieve: A governor's guide to adolescent literacy.* National Governors Association Center for Best Practices.

Biancarosa, G., & Snow, C. E. (2004). *Reading next: A vision for action and research in middle and high school literacy.* Washington, DC: Alliance for Excellent Education. http://www.all4ed.org/publications/ReadingNext/ReadingNext.pdf

Buly, M. R., & Valencia, S. (2003). *Meeting the needs of failing readers: Cautions and considerations for state policy.* Center for the Study of Teaching and Policy. University of Washington.

Carnine, D., Silbert, J., & Kame'enui, E. J. (1990). *Direct instruction reading* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Carnine, D. W., Silbert, J., Kame'enui, E. J., & Tarver, S. G. (2004). *Direct instruction reading* (4th ed.). Upper Saddle River, NJ: Pearson.

Chall, J. S. (1983). *Stages of reading development.* New York: McGraw-Hill.

Chall, J. S., & Jacobs, V.A. (2003). Poor children's fourth-grade slump. *American Educator, 27,* 14.

Coalition for Juvenile Justice. (2001). *Abandoned in the back row: New lessons in education and delinquency prevention.* (Coalition for Juvenile Justice 2001, Annual Report). Available at http://www.juvjustice.org/media/resources/ resource_122.pdf. Retrieved April 17, 2008.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied Multiple Regression/ Correlation Analysis for the Behavioral Sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Crawford, L., Tindal, G., & Stieber, S. (2001). Using oral reading rate to predict student performance on statewide achievement tests. *Educational Assessment, 7*(4), 303-323.

Cunningham A. E., & Stanovich K. E. (2001). What reading does for the mind. *Journal of Direct Instruction, 1,* 137-149.

Curtis, M. E., & Longo, A. M. (1999). *When adolescents can't read: Methods and materials that work.* Cambridge, MA: Brookline Books.

Daane, M. C., Campbell, W. S., Goodman, M. J., & Orange, A. (2005). *Fourth-grade students reading aloud: NAEP 2002 special study of oral reading* (NCES 2006-469). U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. Washington, D.C.: Government Printing Office.

Davis, D., Spraker, J., & Kushman, J. (2005). Improving adolescent literacy: Findings from research. Northwest Regional Educational Laboratory. Portland, OR.

Edformation. (2004). *AIMSweb progress monitoring and assessment system.* Retrieved February 14, 2006, from http://www.edformation.com

Espin, C. A., & Deno, S. L. (1993). Performance in reading from content area text as an indicator of achievement. *Remedial and Special Education,* 14,(6),

Fuchs, L. S., Fuchs, D., Hamlett, C.L., Waltz, L.,& German, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22,* 27-48

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*(3), 239-256.

Gall, M. D., Gall, J. P., & Borg, W. R., 2003. *Educational research: An introduction* (7th ed.). Boston: Allyn & Bacon.

Good, R. H., III, & Kaminski, R. A. (Eds.). (2002). *Dynamic indicators of early literacy skills* (6th ed.). Eugene: University of Oregon, Institute for the Development of Educational Achievement.

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5*, 257-288.

Greene, J. P., & Forster, G. (2003). *Public high school graduation and college readiness rates in the United States*. New York: Manhattan Institute for Policy Research.

Grigg, W.S., M.C. Daane, Y. Jin, and J.R. Campbell. *The Nation's Report Card: Reading 2002* (NCES 2003-521). U.S. Department of Education, National Center for Education Statistics. Washington, D.C.: U.S. Government Printing Office.

Hamilton, C., & Shinn, M. R. (2003). Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skllls. *School Psychology Review, 32*, 228-240.

Harris, T. L., & Hodges, R. E. (1995). *The literacy dictionary: The vocabulary of reading and writing*. Newark, DE: International Reading Association.

Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher*, 59, 636-644.

Hasbrouck, J. E., & Tindal, G., (1992). Curriculum-based oral reading fluency norms for students in grades 2-5. *Teaching Exceptional Children, 24*(3), 41-44.

Heller, R., & Greenleaf, C. L. (2007). *Literacy instruction in the content areas: Getting to the core of middle and high school improvement*. Washington, DC: Alliance for Excellent Education.

Hintze, J. M., & Silberglitt, B. (2005). A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high-stakes testing. *School Psychology Review, 34*(3), 372-386.

Hosp, M. K., & Fuchs, L.S., (2005). Using CBM as an indicator of decoding, word reading, and comprehension: Do the relations change with grade? *School Psychology Review, 34*, 9-26.

Howe, K. B., & Shinn, M. M. (2002). *Standard reading assessment passages (RAPs) for use in General Outcome Measurement: A manual describing development and technical features.* Eden Prairie. MN: Edformation.

Howell. K. W., & Nolet, V. (2000). *Curriculum-based evaluation: teaching and decision making.* Belmont, CA: Wadsworth/Thomson Learning.

Johnson, M. S., Kress, R. A., & Pikulski, J. J. (1987). *Informal reading inventories* (2nd ed.). Newark, DE: International Reading Association.

Joftus, S. (2002). *Every child a graduate: a framework for an excellent education for all middle and high school students.* Washington, DC: Alliance for Excellent Education. Available at www.All4Ed.org

Kame'enui, E. J., Fuchs, L., Francis, D. J., Good, R. G., III, O'Connor, R. E., Simmons, D. C., Tindal, G., & Torgesen, J. K. (2006). The technical adequacy of tools for assessing reading competence: A framework and review. *Educational Researcher, 35*(4), 3-11.

Kamil, M. I. (2003). *Adolescents and literacy: Reading for the 21st century.* Washington, DC: Alliance for Excellent Education.

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95*(1), 3-21.

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

Lenz, K. (2005). Special Connections. www.specialconnections.ku.edu. 2005.

Licht, Mark H. (2005). Multiple regression and correlation. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 19-54). Washington, DC: American Psychological Association.

Marzano R. J. (2004). *Building background knowledge for academic achievement.* Alexandria, VA: ASCD.

Marzano, R., & Pickering, D. (2005). *Building academic vocabulary: Teacher's manual.* Alexandria, VA: ASCD.

Mastropieri, M. A., Scruggs, T. E., & Graetz, J. E. (2003). Reading comprehension instruction for secondary students: challenges for struggling students and teachers. *Learning Disability Quarterly, 26,* 103-116.

McCombs, J. S., Kirby, S. N., Barney, H., Darilek, H., & Magee, S. (2004). *Achieving state and national literacy goals: a long uphill road.* A report to Carnegie Corporation of New York.

McGlinchey, M. T., & Hixson, M. D. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review 33*(2), 193-203.

Meyer, M. S., & Felton, R. H. (1999). Repeated reading to enhance fluency: Old approaches and new directions. *Annals of Dyslexia, 49,* 283-306.

Moats, L. C., (2001). When older students can't read. *Educational Leadership, 58,* 36-40.

Moore, D. W., Bean, T. W., Birdyshaw, D., & Rysik, J. A. (1999). *Adolescent literacy: A position statement.* International Reading Association.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly 19*(3), 304-30.

National Association of Secondary School Principals. (1996). *Breaking ranks: Changing an American institution.* Reston, VA: Author.

National Association of Secondary School Principals. (2004). *Breaking ranks II: Strategies for leading high school reform.* Reston, VA: Author.

National Association of State Boards of Education. (2006). *Reading at risk: The state response to the crisis in adolescent literacy.* Retrieved April 17, 2008, from http://www.nasbe.org/Adolescent_Literacy/Files/Reading_At_Risk_Full_ Report.pdf

National Center for Education Statistics (NCES). (2003). *Nation's Report Card: Reading 2002.* Washington, DC: U.S. Government Printing Office. Retrieved April 16, 2008, from http://nces.ed.gov/nationsreportcard/pdf/main2002/2003521.pdf

National Center on Student Progress Monitoring. U.S. Office of Special Education. Retrieved April 17, 2008. http://www.studentprogress.org/

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* (NIH Publication No. 00-4769). Washington DC: U.S. Government Printing Office.

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

Oregon Department of Education. (2006). Superintendent's Update. #149. April 21, 2006.

Oregon Department of Education. Technical Report: Volume 4, Reliability and Validity. (2007).

Orfield, G., Losen, D., Wald, J., & Swanson, C. (2004). *Losing our future: How minority youth are being left behind by the graduation rate crisis,* Cambridge, MA: The Civil Rights Project at Harvard University. Available at http://www.civilrightsproject.harvard.edu/research/dropouts/LosingOurFuture.pdf

Partnership for Reading. (2001). *Put reading first: The research building blocks for teaching children to read.* Washington, DC: National Institute for Literacy.

Pellegrino, J. W., Chudowsli, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment.* Washington, DC: National Academy Press.

Perie, M., Grigg, W., & Donahue, P. (2005). *The Nation's Report Card: Reading 2005* (NCES 2006-451). U.S. Department of Education, National Center for Education Statistics. Washington, D.C.: U.S. Government Printing Office.

Pikulski, J. J., & Chard, D. J. (2005). Fluency: Bridge between decoding and comprehension. *The Reading Teacher, 58,* 510-519.

Pinnell, G. S., Pikulski, J. J., Wixson, K. K., Campbell, J. R., Gough, P. B., & Beatty, A. S. (1995). *Listening to children read aloud: Data from NAEP'S Integrated Reading Performance Record (IRPR) at Grade 4,* (NCES 95-726). Washington DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Rasinski, T. V., Padak, N. D., McKeon, C. A., Wilfong, L. G., Friedauer, J. A., & Heim, P. (2005). Is reading fluency a key for successful high school reading? *Journal of Adolescent & Adult Literacy, 49,* 22-27.

Salvia, J., & Ysseldyke, J. E. (1995). *Assessment* (6th ed.). Boston, MA: Houghton Mifflin.

Share, D., & Stanovich, K. (1995). Cognitive processes in early reading development: Accommodating individual differences into a mode of acquisition. *Issues in Education: Contributions from Educational Psychology, 1,* 1-57.

Shinn, M. R. (Ed.). (1998). *Advanced applications of curriculum-based measurement.* New York,: Guiliford.

Shinn, M. R., Good, R. H., Knutson, N., Tilly, W. D., & Collins, V. L. (1992). Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review, 21,* 459-479.

Shinn, M. R., & Shinn, M. M. (2002). *AIMSweb training workbook: Administration and scoring of reading curriculum-based measures (R-CBM) for use in general outcome measurement.* Eden Prairie, MN: Edformation.

Siberglitt, B., Burns, M. K., Madyun, N. H., & Lail, K. E. (2006). Relationship of reading fluency assessment data with state accountability test scores: A longitudinal comparison of grade levels. *Pschology in the Schools, 4*(5), 527-535.

Snow, C. E. (2002) *Reading for understanding: Toward a research and development program in reading comprehension.* Santa Monica, CA: RAND.

Snow, C. E., & Biancarosa, G. (2003). *Adolescent literacy and the achievement gap: What do we know and where do we go from here?* New York: Carnegie Corporation of New York.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Stage, S. A., & Jacobsen, M. D. (2001). Predicting student success on a state-mandated performance-based assessment using oral reading fluency. *School Psychology Review, 30*(3), 407

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 36-407.

Stanovich, Keith, E. 2000. *Progress in understanding reading: Scientific foundations and new frontiers.* New York: Guiliford Press.

Tindal, G. (2002). Large-scale assessments for all students: Issues and options. In G. Tindal & T. M. Haladyna (Eds.), *Large-scale assessment programs for all students* (pp. 1-24). Mahwah, NJ: Lawrence Erlbaum Associates.

Torgesen, J. K., Houston, D. D., Rissman, L. M., Decker, S. M., Roberts, G., Vaughn, S., Wexler, J., Francis, D. J., Rivera, M. O., & Lesaux, N. (2007). *Academic literacy instruction for adolescents: A guidance document from the Center on Instruction.* Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Wood, D. E. (2006). Modeling the relationship between oral reading fluency and performance on a statewide reading test. *Educational Assessment, 11*(2), 85-104.

Yovanoff, P., Duesbery, L., Alonzo, J., & Tindal, G. (2005). Grade-level invariance of a theoretical causal structure predicting reading comprehension with vocabulary and oral reading fluency. *Educational Measurement: Issues and Practice,* 4-12.