USING CURRICULUM-BASED MEASUREMENT TO PREDICT EIGHTH-GRADE STUDENT PERFORMANCE ON A STATEWIDE READING ASSESSMENT

by

MINDY J. LEROUX

A DISSERTATION

Presented to the Department of Educational Methodology,
Policy, and Leadership
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Education

June 2012

DISSERTATION APPROVAL PAGE

Student: Mindy Jean LeRoux

Title: Using Curriculum-Based Measurement to Predict Eighth-Grade Student Performance on a Statewide Reading Assessment

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

Keith Zvoch	Chair
Keith Hollenbeck	Member
Gina Biancarosa	Member
Beth Harn	Outside Member

and

Kimberly Andrews Espy	Vice President for Research & Innovation/Dean of the Graduate School

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

Degree awarded June 2012

DISSERTATION ABSTRACT

Mindy J. LeRoux

Doctor of Education

Department of Educational Methodology, Policy, and Leadership

June 2012

Title: Using Curriculum-Based Measurement to Predict Eighth-Grade Student Performance on a Statewide Reading Assessment

The purpose of this study was to analyze the relationship between oral reading fluency (ORF) and Maze, two common Curriculum-Based Measures (CBMs), and the statewide large-scale assessment of reading in Oregon, the Oregon Assessment of Knowledge and Skills- Reading (OAKS-R). A sample of three cohorts of eighth-grade students in an Oregon school district was used to examine concurrent validity, predictive validity, and the relation between demographic characteristics, disability status, and socio-economic status and the ability to predict performance on the OAKS-R.

Findings of the concurrent validity analysis revealed a moderately strong positive correlation between the OAKS-R and both ORF and Maze measures, with ORF demonstrating a slightly stronger correlation with the OAKS-R. Multiple regression analyses were used to analyze the predictive relationship between ORF and Maze and scores on the eighth-grade OAKS-R. Both ORF and Maze were statistically significant predictors of OAKS-R, demonstrating moderately positive relationships with scores on the state reading test. Although no interaction effect was found between disability classification or eligibility for free or reduced-price meals and the different CBMs, in relation to the OAKS-R, student disability status was negatively related to performance

on the OAKS-R. The relationship between OAKS-R performance and low socio-economic status, as measured by eligibility for free or reduced-price lunch status, was not consistent across the cohorts. This finding is promising, as it indicates that there may be factors that schools can take advantage of to ameliorate the relationship between poverty and reading outcome measures for eighth-grade students.

In light of study results, suggestions for future research, as well as implications for the field, are discussed. This study adds to the research literature documenting that ORF and Maze assessments provide schools with valuable information to predict student performance on statewide large-scale assessments of reading. With CBM data available early in the school year, schools can provide additional intervention as needed, potentially leading to improved end-of-year student performance on the OAKS-R.

CURRICULUM VITAE

NAME OF AUTHOR: Mindy J. LeRoux

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon

DEGREES AWARDED:

Doctor of Education, Educational Methodology, Policy, and Leadership, 2012, University of Oregon
Master of Science, Communication Disorders and Sciences, 1999, University of Oregon
Bachelor of Science, Psychology, 1994, University of Oregon

AREAS OF SPECIAL INTEREST:

Special Education and Support Services
Academic and Behavioral Interventions
Response to Intervention Systems
Speech-Language Pathology
Assistive Technology

PROFESSIONAL EXPERIENCE:

Assistant Director, Special Services, Bethel School District, 2006-present
Oregon Initial Administrator License, 2006-present
Principal, Bethel School District Summer School, 2010-11
Speech Language Pathologist, Bethel School District, 1999-2006
Oregon Continuing Teaching License, 1999-present
ASHA Certificate of Clinical Competence, 2000-present
Speech Language Pathologist, Riverpark Rehabilitation, 2000-2006
District Staff Trainer, IDEA ’04, Bethel School District, 2005-present
Clinic Supervisor, University of Oregon, 2002
Oregon Initial Administrator License, 2006-present

ACKNOWLEDGMENTS

I wish to express my appreciation to my advisor, Professor Keith Zvoch for his guidance as I learned and grew from the process of obtaining a doctoral degree. My committee members, Keith Hollenbeck, Gina Biancarosa, and Beth Harn contributed to my successful completion of my doctoral degree requirements by supporting my ideas while also pushing me to expand my knowledge and grow as an educational leader. Keith Hollenbeck far exceeded my expectations as a committee member. I sincerely appreciate his trust and belief in me, not only as a doctoral candidate, but also as a practicing special educational leader. I give many thanks to Gina Biancarosa for her instruction and guidance. It was through her instruction and patient guidance that it all came together for me. I appreciate Beth Harn, as she was instrumental in helping me to get the process started. Beth supported me in my writing and, in general, was flexible and supportive throughout the process. I was fortunate to have such an amazing committee. I will fondly remember each member for his or her unique contributions to my product.

DEDICATION

Completing the requirements for a doctoral degree is not an easy feat. It takes a team of support and genuine understanding from those closest to you. For me the process started with an internal desire to improve my personal achievement in a way that impacted others meaningfully. It slowly dissolved into a marathon in which those around me provided positive encouragement and made numerous exceptions that allowed me to continue to propel forward in the process. Those closest to me include my mother, Barbara McEntee; my grandpa, Willard Burke; my husband, Doug; my closest friends, Nancy McNamara and Dr. Elise Guest; and my coworkers, Kim Breeding, Monique Siemerink, Lori Smith, and Drew Braun.

My mom is an amazing human being and my greatest cheerleader. She was a special educator for many years, working with students with profound disabilities. She understands my passion for special education and the need to continue learning so that we can best serve our neediest students. She has taught me to live life with an open mind, listen carefully, and treat others as I would like to be treated.

My grandpa Dr. Willard Burke was also a passionate educator. He instilled in me the importance of a good education as the cornerstone of my career. He supported me through undergraduate college and sadly passed away on June 27, 1993. He was an exceptional educational leader who modeled fairness, honesty and perseverance as a person and professional. Thoughts of my grandfather propel me through every finish line.

My husband Doug supported me emotionally as I encountered the inevitable ups and downs of the doctoral process. His actions and support during the four-year process

allowed me to keep my focus, by supporting difficult choices I had to make to ensure I stayed the course. My sincere thanks and undying love for Doug as he also endured the marathon, right by my side.

My closest friends are my family of choice. Nancy Mcnamara supported me and understood when I disappeared and didn't seem to be the friend I was prior to entering the program. Throughout the process, she was always there for me. Nancy is a true friend and her support allowed me the freedom to take this road less traveled, a road that in the end led right back to our friendship. Elise endured the doctoral program with me and was the first in our cohort to graduate. She led the way with grace, modeling perseverance and dedication, and ultimately writing the longest dissertation I ever hope to read. Elise listened to me with a special understanding like no one else. She always ended our conversations with a positive affirmation to guide me toward my goal.

My coworkers, bless their hearts, for even mentioning that I should pursue a Doctoral degree in Education. Kim Breeding, Monique Siemerink, Lori Smith, and Drew Braun all understand the intense dedication and focus it takes to complete the program. They believed in me and supported me in numerous ways to ensure that I continued making forward progress, even when it seemed the process was barely moving. I deeply appreciate all their support and I humbly acknowledge that without it, I might not have made it.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

Adolescents need strong reading skills to affect positive academic and life outcomes (Biancarosa & Snow, 2006; Carnegie Council on Advancing Adolescent Literacy, 2010; Deshler & Hock, 2007). Strong literacy skills are the foundation for completing a K-12 education (Deshler & Hock, 2007; ODE, 2009). The specific reading skills adolescents need in middle and high school include the ability to decode print, read fluently, understand increasing levels of vocabulary, have adequate background knowledge, and ultimately, to think critically about the text they read (Deshler & Hock, 2007). For students who do not have strong literacy skills, attaining a high school diploma is a challenge that may be too great to overcome (ODE, 2009).

For students who do not complete high school, negative outcomes affecting social participation, civil participation, and career opportunities are more likely (Biancarosa & Snow, 2006; Duncan, 2009). The prosperity of our society and individual quality of life are dependent upon the skills learned in school (Strickland & Alvermann, 2004). The meager earnings of individuals who have not achieved a high school diploma negatively affect the value of 401K plans, Social Security, the stock market, and in general, the health of the American economic system (Strickland & Alvermann, 2004). These negative outcomes, for both the individual and society, highlight the urgency and need for improvements in adolescent literacy outcomes.

The Adolescent Literacy Crisis

An adolescent literacy crisis was first described in *A Nation at Risk*: *The Imperative for Educational Reform* (National Commission on Excellence in Education, 1983). *A Nation at Risk* reported that 23 million adult Americans and about 13% of all 17-year-olds in the United States were functionally illiterate, SAT scores were dropping, and more students were entering college in need of remedial courses. These concerns were not alleviated in the more recent update *A Nation Accountable: 25 Years After a Nation at Risk* (U.S. Department of Education, 2008). *A Nation Accountable* reported that today Americans are at even greater risk than in 1983, due to the rising demands of the global economy and demographic shifts in the United States population. The Reading First initiative (Office of Elementary and Secondary Education, 2002) provided states financial support to implement scientifically-based instruction and assessment practices with the goal of ensuring all students read at grade level by the end of third grade. The Reading First initiative did not specifically address the importance of ensuring students are developing reading comprehension skills that will be needed in middle and high school (Biancarosa & Snow, 2006). The adolescent struggling reader encounters different challenges than the child struggling reader. These differences necessitate educators understand the unique nature of the adolescent struggling reader, in order to assess and intervene appropriately (Biancarosa & Snow, 2006; Kamil, 2003; Mastropieri, M.A., Scruggs, T.E., Graetz, J.E., 2003).

Adolescent Struggling Readers

Adolescent students who are struggling to read are typically able to read the words, but are unable to comprehend what they read for a variety of reasons (Biancarosa

& Snow, 2006; Kamil, 2003; Kamil et al., 2008). Lack of reading comprehension skills may be due the fact that they have not developed enough fluency for comprehension to occur, or they may lack the specific strategies to facilitate comprehension and the ability to adjust strategies according to the purpose of reading (Biancarosa & Snow, 2006). For example, middle and high school students encounter multiple subjects in a single day, and their reading strategy must shift accordingly. In science class, students must learn to read skeptically and question the material read, while in math class, students must read and follow the directions as stated. Adolescent students are expected to learn new words, facts, and concepts from reading as well as to interpret, critique and summarize what they read and all with an eye toward the concepts and logic privileged within science rather than literature (Carnegie Council on Advancing Adolescent Literacy, 2010).

As students advance through middle and high school, many changes occur in their textbooks: (a) texts become longer, (b) words, sentences, and text structures become more complex, (c) graphic representations are more important, (d) conceptual challenges increase, and (e) texts begin to vary widely across content areas (Carnegie Council on Advancing Adolescent Literacy, 2010). These changes in learning demands, which may be less of an obstacle for some, prove daunting to the struggling adolescent reader. For example, middle and high school textbooks are frequently written at levels that are far above the grade level in which the text is being used (Mastropieri, Scruggs, & Graetz, 2003). For students whose reading skills are below their current grade level, it is not a reasonable expectation that they will learn as expected from reading a textbook that is written with language and vocabulary that exceeds their current grade level.

In addition to the increased literacy demands, adolescents simultaneously face

developmental, cognitive, and emotional challenges (Carnegie Council on Advancing Adolescent Literacy, 2010). The average adolescent is simultaneously engaged in trying to forge a sense of self, preparing goals for his or her adult life, and navigating complex social and emotional relationships. Adolescents may also be learning English, experiencing poverty, resolving gender identity and sexual orientation issues, dealing with minority and/or immigrant status, or have special needs (Spencer, 1999). Moreover, in the later grades, students are less motivated to read, and schools may not be engaging to students (Kamil, 2003; National Institute of Child Health and Human Develompent, 2000). These additional factors contribute to the challenge of improving adolescent literacy skills.

Eighth grade, in particular, is a critical year for students. A student's performance during eighth grade is deemed the best predictor of college and career readiness by high school graduation (American College Test, 2008). Eighth-grade students who have achieved strong literacy skills, especially reading comprehension, are more likely to be on track for lifelong access to learning (Durkin & Allington, 2004). Ensuring that all eighth-grade students have attained the knowledge and skills that will earn them a high school diploma is the single most important step that can be taken to improve their college and career readiness (American College Test, 2008).

Thus, it is important to consider the literacy skills students need as they leave eighth grade to enter high school, particularly because high school diploma requirements have become increasingly stringent. A closer look at the high school diploma requirements underscores the importance of ensuring eighth-grade students are on track when they enter high school to achieve the high school diploma by the time they are

slated to graduate.

Increased Diploma Expectations

As an important step toward high school graduation, students should end their eighth-grade year with a passing score on the statewide reading assessment. This achievement serves as an important demonstration of their literacy skills and ability to learn academic content in high school, preparing them for post-secondary education and careers. To meet the rapidly-changing expectations in the workforce and to improve students' post high school outcomes, several changes were recently made to the requirements of the Oregon high school diploma, the Oregon Diploma (Oregon Department of Education, 2007b). These changes were a product of Oregon's involvement in the American Diploma Project.

Oregon joined 26 other states in the American Diploma Project to identify the English and Mathematics benchmarks high school graduates need to be successful in post high school education and careers (Achieve, 2007). An important outcome of Oregon's participation in the American Diploma Project was the addition of a provision that students must demonstrate proficiency in *essential skills* prior to being awarded an Oregon Diploma. The essential skills are the skills deemed critical for future success. The essential skills requirements will be implemented in Oregon on the following timeline: (a) By 2012, to read and comprehend a variety of texts; (b) by 2013, to write clearly and accurately; and (c) by 2014, to apply mathematics in a variety of settings. Of importance to this proposal is that beginning in 2012, students will need to demonstrate the essential skill of reading by passing the Oregon Assessment of Knowledge and Skills- Reading (OAKS-R) in high school. Students who cannot pass the OAKS will not earn an Oregon

Diploma (ODE, 2007b).

Students who do not earn an Oregon Diploma may have the option to earn an alternative diploma and/or certificate. Students who do not meet the requirements of an Oregon Diploma may earn a modified diploma, extended diploma, or School District certificate of attendance and/or completion. However, any diploma or certificate other than the Oregon Diploma limits the potential post-secondary options for students (ODE, 2007b). For example, most four-year colleges do not accept students who have earned a modified diploma. Additionally, some branches of the military do not accept the modified diploma, and the modified diploma does not meet the eligibility requirements for obtaining financial aid at a post-secondary institution (ODE, 2010).

Recognizing the importance of ensuring students achieve a high school diploma, the challenge is to find an efficient and accurate assessment, or combination of assessments, to inform instruction. Assessments are needed to identify students in need of additional reading instruction prior to high school. Providing the necessary reading instruction during the eighth grade year may improve students' performance on the statewide reading assessment. This study investigated the value and utility of linking eighth-grade students' curriculum-based reading performance to OAKS-R performance. Through early identification of eighth-grade students who are at risk for not passing the spring-administered OAKS-R, schools have the opportunity to provide additional reading instruction prior to administering the OAKS-R. Passing the eighth-grade OAKS-R is an important indicator that students are meeting grade level expectations and as such, are on track for meeting the reading essential skill requirement of the Oregon Diploma (ODE, 2009).

Increasing Workplace Demands

Achieving a high school diploma is the foundational step toward post-high school opportunities. The labor market has shifted dramatically in the last few decades, to the current state where a post-secondary education is necessary for nearly 90% of all new, high growth and high wage jobs (Amos, 2008). The shift in the labor market is the result of drastic reductions in manufacturing jobs that were previously an attractive option for high school dropouts. Of the manufacturing jobs that remain today, increasingly the positions require higher levels of education and training (Amos, 2008). Unfortunately, improvements in the literacy skills of adolescent students have not kept pace with the increasing demands for literacy in the workplace (Amos, 2008; Guensburg, 2006). Now, more than ever before, the literacy demands of post-secondary education and the workplace are greater than the skills students have acquired by the time they leave high school (Lee, Grigg, & Donahue, 2007). Over the last several years, many educators, researchers, advocacy groups and others have lobbied policy makers to better prepare students for the careers of today (Heller & Greenleaf, 2007). The resulting policy has focused on improving methods for identification of students with learning disabilities, implementing appropriate assessment practices, and mandating the use of research-based curriculum in our schools (IDEA, 2004; NCLB, 2001).

The current study examined the use of two curriculum-based measures (CBM), Oral Reading Fluency (ORF) and Maze, with eighth-grade students as concurrent and predictive measures of the OAKS-R. Specifically, this study explored whether ORF and Maze were related to student performance on a large-scale reading assessment, the Oregon statewide reading assessment (OAKS-R).

Assessments of Reading

Assessments are an important part of the modern educational environment in the United States. Educators must be familiar with large-scale as well as more formative assessments.

Large-Scale Assessments of Reading

Large-scale assessments have been used to evaluate students' reading ability over the last several decades and provide longitudinal information regarding students' reading performance. There are both international and US-based assessments. An international assessment, the Program for International Student Assessment (PISA), sponsored by the Organization for Economic Cooperation and Development (OECD), is designed to evaluate student achievement in three key content areas across 29 countries (Institute of Education Sciences, 2009). The PISA measures 15-year-old students' reading, mathematics, and science literacy every three years and permits international comparisons of student performance. The particular focus of the PISA reading assessment is on cognitive skills and the application of reading to problems within a real-life context.

In PISA 2003, U.S. 15-year-old students' average literacy score was 495, which was very close to the international average of 494. This score placed U.S. 15 year-olds in the middle third of participating OECD nations. Only 9% of U.S. 15-year-old students achieved the highest level of proficiency on the PISA, corresponding to a literacy score at or above 625 (Institute of Education Sciences, 2009). The PISA results indicate that on an international level, the United States' eighth-grade students are not leading performers with respect to the cognitive, problem solving with real-life contexts that the PISA

measures. These results are consistent with another large scale, national assessment, the National Assessment of Educational Progress (NAEP) (U.S. Department of Education, 2009).

The NAEP (U.S. Department of Education, 2009) serves as a national measure of student academic achievement in the areas of reading and mathematics. The purpose of the NAEP is to provide a current, common comparison among states as well as to track U.S. students' performance over time. Results from the NAEP indicate that eighth-grade students' reading performance has been mostly flat, with only a slight gain since 1992 (U.S. Department of Education, 2009). In 1992, students' average score was 260, and in 2009, students' average score was 264. In 2007, 69% of the nation's eighth-grade students performed below the *proficient* level in reading at their grade level (NAGB, 2009). In 2009, the number of students performing at or above *basic* and at or above *proficient* increased 1 percentage point from 2007 to 2009, with a higher number in 2009 than in 1992. U.S. students' performance at the *advanced* level has remained stagnant; since 1992, only 3% of students have performed at the *advanced* level. The NAEP results portray a general lack of progress in literacy achievement in the United States for eighth-grade students.

Oregon Statewide Summative Reading Assessment

The *No Child Left Behind Act of 2001* (NCLB, 2001) mandates all students in grades three through eight, and one time in high school, must demonstrate adequate yearly progress as measured by high-stakes, summative assessments. In Oregon, the high stakes, summative reading assessment for eighth-grade students is the OAKS-R. The OAKS-R is not designed, nor intended to provide, curricular or instructional assistance to

educators. The OAKS-R is designed to capture a single snapshot of student and district performance that may be used to paint a larger picture of the school, district, state, and national performance. Schools, districts, and states are held accountable for their respective performance on these assessments as per the *No Child Left Behind* (NCLB, 2001) legislation.

The OAKS-R is a criterion-referenced assessment based on the Oregon Content Standards. The purposes of OAKS are to: (a) provide information on individual student achievement, (b) provide information for federal NCLB requirements, (c) provide state and local policy makers information on which to base decisions, (d) support instructional program improvements, and (e) inform the public about student achievement in Oregon schools (ODE, 2007a).

Although such federally mandated summative assessments meet a societal need to measure and report student achievement and school accountability, they do not inform teachers about individual student progress prior to 3rd grade. Nor do summative assessments provide teachers with current student-level data that can be used to inform instructional decisions in response to student performance *during* the school year.

To ensure timely and adequate development of reading skills, more frequent measurement of skills is necessary to inform instruction and improve student literacy and performance on statewide reading assessments (Good, Simmons, & Kame'enui, 2001). With valid and reliable data on student learning, educators can adjust instruction to meet individual student needs. For eighth-grade students, time is of the essence; quick, valid, and reliable assessments to measure progress, inform instruction, and predict outcomes on the statewide assessment are needed.

Formative Assessment of Reading

Given the nature of high stakes assessments and the need to ensure all students are learning to read, school districts have begun to utilize screening methods of assessment to identify students who are not on track to meeting reading benchmarks (Good et al., 2001) and progress monitoring assessments to evaluate the degree to which students identified as needing more help are "catching up" to grade-level peers (Deno, 2003). Sometimes referred to as formative assessments, such tests are quick assessments providing educators frequent feedback on student learning and may also be used for school wide screening (Deno, 2003). More frequent assessment of students' reading skill development provides educators valuable information to inform instruction and ensure students' timely development of proficient reading skills (Shinn & Bamonto, 1998; Shinn & Good, 1992). Teachers use many different formative assessments to evaluate student learning (e.g., short classroom quizzes, reflective journals, and end of chapter comprehension questions). Without simple and quick measurement systems that provide information on achievement and learning, schools and districts cannot effectively serve the diverse needs of the population they serve (Megert, 2010).

Curriculum-Based Measurement

Curriculum-based measurement (CBM) is an approach to formative assessment that offers reliable and valid measures to efficiently assess student skills (Deno, 2003; Shinn & Bamonto, 1998). In general, CBM are a set of quick, standardized procedures, for assessing and monitoring student skill level in reading, math, spelling, and writing (Hasbrouck & Tindal, 2006; Shinn & Bamonto, 1998). Shinn and Bamonto (1998) describe CBM as evaluation procedures designed to function as "academic

thermometers" (p. 1), measuring students' growth in the skills relevant to school outcomes. CBM typically take less than five minutes to complete and may be administered as frequently as two times a week to measure a student's progress. CBM may also be used in a school-wide screening process to track large groups of students by collecting individual student information two or three times in a school year. With a school wide screening protocol, reading progress can be measured at the individual, classroom, and school level, providing data for comparisons with grade level benchmarks. School wide screening may identify students needing additional instruction earlier rather than relying entirely on individual teachers to identify struggling students. In the area of reading, two common CBM reading measures are oral reading fluency (ORF) and Maze.

Curriculum-Based Measure: Oral Reading Fluency. Oral reading fluency can be considered as the oral translation of text with speed and accuracy (Fuchs, Fuchs, Hosp, & Jenkins, 2001). The CBM Oral Reading Fluency (ORF) task requires students to read aloud for one minute, from either an informal passage selected from the student's curriculum, or a formal passage from a standardized assessment, such as AIMSweb (Shinn & Shinn, 2002b). The student's total words read per minute (TWPM) minus the words read incorrectly equal the words correct per minute (WCPM).

Curriculum-Based Measure: Maze. The CBM Maze is a timed, silently read, multiple-choice, cloze task that is presented within the context of a story. The first sentence of the passage is left intact. Thereafter, every n^{th} (e.g., fifth, seventh, or ninth) word is deleted and replaced with three words inside parenthesis. One of the words is the exact one from the passage while the other two words are distracters, representing a near

distracter and a far distracter. Students must read through the passage and choose the correct word that makes sense in the context of the sentence and story. The Maze requires a student to read fluently, accurately, and with comprehension to earn a high score. The Maze score is the number of correct word choices within the given time limit (e.g., two minutes, three minutes, or five minutes).

Assessment to Improve Adolescent Literacy Outcomes

Strong evidence supports the validity of curriculum-based measurement systems for use with elementary students (Crawford, Tindal, & Stieber, 2001; Deno, 2003; Marcotte & Hintze, 2009; Shinn & Bamonto, 1998; Shinn & Good, 1992; Wiley & Deno, 2005). However, evidence regarding the utility of curriculum-based measures for use with students in the middle school grades, and eighth grade in particular, is less abundant, and the results are not consistent.

What follows is a literature review highlighting the need for quick and valid assessments of eighth-grade students' reading skills to guide instructional practices to improve adolescent literacy outcomes. The resulting study evaluated curriculum-based measures, ORF and Maze, as valid concurrent and predictive measures of eighth-grade student performance on the statewide reading assessment, the Oregon Assessment of Knowledge and Skills- Reading/Literature (OAKS-R).

CHAPTER II

LITERATURE REVIEW

In this chapter, relevant research and theories of reading development are presented, followed by research linking fluency and comprehension. What follows is a review of current assessment practices and needs through the available research on the use of curriculum-based measures, oral reading fluency and Maze CBM for eighth-grade students. The purpose of the review is to highlight the relevance and need for efficient and valid measures that can form the basis of a formative assessment system for identifying eighth-grade students at risk of performing poorly on the Oregon statewide reading assessment, the Oregon Assessment of Knowledge and Skills- Reading/Literature (OAKS-R).

Developing Reading Skills

Learning to read is a complex process requiring a combination of skills that ultimately culminate in the ability to comprehend text (Carnine, Silbert, Kame'enui, & Tarver, 2004; Chall, 1983; Ehri, 2005; Hoover & Gough, 1990; NICHD, 2000; Smith, 2008). Reading comprehension is the result of an individual's mastery of the component skills of reading (Carnine, et al., 2004; Chall, 1983; Ehri, 2005; NICHD, 2000). The component skills, as outlined by the National Reading Panel (NRP), are phonemic awareness, phonics, fluency, vocabulary, and text comprehension. If students can learn

and become proficient in these skills, they will have a foundation on which future academic learning can occur (Heller & Greenleaf, 2007; NICHD, 2000).

One theory on how reading comprehension develops is the Simple View of Reading (Hoover & Gough, 1990). The Simple View of Reading describes reading as consisting of two components— decoding and linguistic comprehension. In this explanation of the process and development of reading comprehension, both components are equally important to developing reading skills; that is, individuals must have skills in both decoding and linguistic comprehension. The Simple View essentially compartmentalizes the components of reading to being either decoding or linguistic in nature (Hoover, et al. 1990).

In a study of the role of fluency and whether it should be added to the Simple View of Reading, oral reading fluency was measured to determine the unique contribution fluency has on concurrent and predictive models of reading comprehension after controlling for word recognition accuracy and listening comprehension abilities (Adlof, Catts, & Little, 2006). The study sample contained 2nd, 4th, and 8th grade students, a total of 604 students. Approximately half of the students in the sample had been previously identified as having either language and/or nonverbal cognitive impairments in kindergarten. To reduce the potential sampling bias given the sample characteristics, weighting procedures were used.

In a grade level analysis of the sample, 2nd, 4th, and 8th grades, the results did not support adding a component of fluency to the Simple View of Reading. The additional component of fluency did not account for any unique variance in either concurrent or predictive models in any of the three grade levels in the study. Although

this study diminishes the importance of fluency as an additional factor to the Simple View of Reading, Adlof et al. (2006) were careful to point out that practitioners should not eliminate fluency components to interventions for students with reading disabilities. The findings support intervention programs that are focused on developing the fundamental skills involved in word recognition and listening comprehension, which include fluency.

Another prevalent theory on how reading develops is the Stages of Reading Development (Chall, 1983) which are based on the cognitive, developmental stages described by Jean Piaget. The Stages of Reading Development offer a theoretical perspective of how reading skills develop, ultimately providing the foundation for the reader's ability to read with comprehension. In contrast to The Simple View of Reading (Hoover & Gough, 1990), the Stages of Reading Development provide a particularly detailed account of reading comprehension including the length of time, and necessary instruction, to acquire advanced reading comprehension skills. Of the two theories presented, the Stages of Reading Development is most germane to this study as it focuses on the importance and development of fluency and reading comprehension skills as it relates to eighth-grade students.

According to the Stages of Reading Development, readers pass through six distinct stages in their development of reading skills. The six stages of reading development describe learning to read as a lengthy process, beginning at birth and continuing (potentially) through the college experience (Chall, 1983). Eighth-grade students should be developing their reading skills in Stage Three. The average learner progresses through the six stages of reading development as follows:

(1) Stage 0: Pre reading: Birth to Age 6. Language skill awareness and overall development sets the stage for initial reading or decoding.

(2) Stage 1: Initial Reading or Decoding: Ages 6-7 (grades 1-2). The reader learns sound-symbol correspondences.

(3) Stage 2: Confirmation, Fluency, Ungluing from Print: Ages 7-8 (grades 2-3). The reader learns high proficiency words, gaining courage and fluency with practice.

(4) Stage 3: Reading for Learning: The Next Step (grades 4-8). The reader expands vocabulary, comprehends diverse texts and contexts, the learner's background experience enhances the learner's reading abilities.

(5) Stage 4: Multiple Viewpoints: Ages 14-18 (high school). In this stage, readers are exposed to a greater variety and depth including multiple viewpoints in school texts and related assignments.

(6) Stage 5: Construction and Reconstruction- A World View (Age 18 and above). In this final stage, readers analyze, synthesize, and form judgments. The reader has learned what to read in order to gain the necessary knowledge. This level of reading is typically acquired by attendance in higher education.

Stages of Reading Development: Stages Zero Through Two

Prior to and during elementary school, typically-developing readers learn as described in the first three stages of the Stages of Reading Development (Chall, 1983). In Stage Zero, ages birth to six years, children are learning language. They are learning vocabulary, developing syntax, and gaining control over words. A reader's growth in Stage Zero may be evidenced by an increase in the understanding of word-sound structures with activities such as rhyming, alliteration, and a beginning understanding that

sounds come together to form words. In Stage One, grades 1 and 2, children are learning the letters of the alphabet and how those letters correspond to the sounds that they represent. In Stage Two, grades 2 and 3, children grow in confidence in their skills learned in Stage One and can read words and stories with increasingly complex words. The reader begins to sound more fluent while reading.

The first three stages of the Stages of Reading Development (Chall, 1983) provide the foundation on which future reading development occurs. This research proposal, however, focuses on eighth-grade students who should be developing skills in Stage Three of the Stages of Reading Development.

Stages of Reading Development: Stage Three

Beginning in Stage Three, readers should be on track toward the next step in their educational career—reading to learn. *Reading to learn* means a student can read to learn new information, in a variety of subject areas. For example, in a middle school science class students are expected to read the textbook to gain detailed information that may not be discussed in class. If a student cannot read the textbook, he/she will miss the material, and not learn as expected. Stage three has two phases, the first phase typically ends at the end of sixth grade and includes the ability to read with comprehension, both orally and silently, up to beginning seventh-grade level material. At this point, readers can read and understand easier parts of adult newspapers and magazines, independently use a dictionary, and they generally possess strong vocabulary comprehension. The second phase of stage three typically concludes at the end of eighth grade, when students can comprehend what is read and are able to learn new information in a variety of formats such as adult literature, popular fiction, science, social studies texts, adult encyclopedias,

and other reference materials. Students who have reached the end of stage three can use print materials and the library efficiently, which may be reasonably extrapolated to mean in the current day of technological advances, students can efficiently use computer methods to conduct research, such as internet exploration. By the end of stage three, students possess the general and technical vocabulary skills that facilitate the ability to read and comprehend adult newspapers and magazines.

Stages of Reading Development: Stages Four and Five

Stage Four and Stage Five address the breadth, depth, and importance of developing advanced reading comprehension to secure future learning and achievement (Chall, 1983). Advanced reading comprehension is a high level skill that requires cultivation; the type of experience and exposure to text that requires time and is commonly acquired through attendance in higher education (e.g., college).

Stage Four spans the high school education experience. In this stage, students demonstrate the ability to read analytically and critically from a broad range of texts including fiction and nonfiction. Academically, students are reading books, textbooks, newspapers, and magazines, and developing extensive vocabulary comprehension skills. Students in stage four read for pleasure and for school academic requirements, demonstrate efficient study strategies, and are able to locate material independently for written reports and other academic purposes. At the conclusion of Stage Four, students will have acquired the skills necessary to either be employed or to attend college where they will enter into the final stage of reading development (Chall, 1983).

The Stages of Reading Development portray the process of learning to read as a linear process, but Adams (1990) points out that reading requires many different systems

and skills to work in simultaneous coordination. The Stages are overlapping, and yet there is a necessary foundational level of skill proficiency in each stage prior to continuing on to the next stage.

Attention and Fluency

In their seminal article, LaBerge and Samuels (1974) asserted a cognitive capacity theory to describe the influence of attention and cognitive resources on learning to read. They explained that as students become more fluent with their word-level reading skills, their cognitive resources are freed up to attend to the higher-level cognitive skills (e.g., comprehension). Another similar approach to describing a cognitive capacity theory of reading development, refers to *automaticity* as the referent to skilled reading which is performed with minimal effort and attention, enabling the reader to focus on comprehension of the read text (Samuels & Flor, 1997). Similar to the cognitive capacity theory, as students are able to rapidly decode words, they are focusing less on the task of decoding and more on the text being read, resulting in greater reading comprehension.

In comparison to the Stages of Reading Development, the cognitive capacity and automaticity theories provide little detail about the various learning processes students must progress through while learning to read, such as language development and sound-symbol correspondence. Rather, learning to read is portrayed as an issue of attention and cognitive capacity to explain fluency's role in developing comprehension. Although the aforementioned theories of reading are not synonymous, the Stages of Reading Development, Cognitive Capacity and Automaticity theories all recognize the important role of fluency in developing reading comprehension. Recognizing reading as a complex system of skills and knowledge working in simultaneous coordination, theorists from all

three approaches regard oral reading fluency as the most salient feature of skillful reading in developing reading comprehension (Adams, 1990).

The Fluency and Comprehension Connection

Researchers have developed reading theories describing the role and relative importance of fluency in developing reading comprehension. The Simple View of Reading (Hoover & Gough, 1990) does not specify fluency as a component providing a unique contribution to the construct of reading comprehension (Adlof, Catts, & Little, 2006), while Automaticity theory (Samuels & Flor, 1997) and Cognitive Capacity theory (LaBerge & Samuels, 1974) incorporate fluency as a critical feature to explain the development of reading comprehension. Reading developmental theories have described learning to read as a process that begins with learning decoding skills that will ultimately lead to reading comprehension (Chall, 1983; Pikulski & Chard, 2005). As students become more fluent, they begin to understand more of what they read, they are learning more words, gaining fluency, and increasing their comprehension. In this fashion, what develops is a "reciprocal, causal relationship" (p. 511) in terms of continued reading development and comprehension of text (Pikulski & Chard, 2005).

Researchers have offered many definitions that begin to describe the role of fluency in developing competent readers who can decode and comprehend varied and complex text. Reading fluency has been compared to speaking; the ease and effortless production of fast, fluent reading sounds like connected speech, whereas the non-fluent reader sounds choppy and plodding while trying to read (Carnine et al., 2004). Fluency has been described as an important indicator of overall reading competence, estimating the readers' ability to comprehend at the sentence level and to "infer the macrostructure

of a passage" (p. 242), which enables the reader to gain meaning from text by incorporating prior learning and make necessary inferences to supply missing information (Fuchs, Fuchs, Hosp, & Jenkins, 2001). In this way, fluency allows readers to focus on comprehension and to think critically about what they read, by incorporating prior knowledge and learning. The fluency with which a reader translates text into spoken words should function as an indicator of both word recognition skill and the individual's comprehension of that text (Fuchs et al., 2001).

To further explore and understand the connection between fluency and comprehension, researchers have used the CBM oral reading fluency (ORF) to compare with other standardized, direct measures of comprehension. Research has examined the concurrent and predictive validity of curriculum-based measurement ORF to both standardized measures of comprehension and high-stakes, large-scale statewide assessments of reading comprehension (Chard et al., 2008; Crawford, Tindal, & Stieber, 2001; Fuchs, Fuchs, & Maxwell, 1988; Marcotte & Hintze, 2009; Megert, 2010; Shinn & Good, 1992; Smith, 2008; Wanzek et al., 2010; Wiley & Deno, 2005; Wood, 2006). The research consistently reports that at the elementary level, ORF has strong, positive concurrent and predictive validity correlation coefficients to both standardized measures and statewide reading assessments (Good, Simmons, & Kame'enui, 2001; Shinn & Good, 1992; Wood, 2006).

To evaluate ORF as a measure of comprehension, Wood (2006) examined the concurrent relationship between third, fourth, and fifth grade students' ORF score (WCPM) and their performance on the Colorado statewide assessment of reading, the Colorado Student Assessment Program (CSAP). The ORF assessment was administered

two months prior to the statewide assessment. Results demonstrated a strong, positive relation between students' ORF and the CSAP across grades: third grade $r = .70$, fourth grade $r = .67$, and fifth grade $r = .75$ ($p < .0001$). This study revealed that for the study sample of third, fourth, and fifth grade students, ORF showed a strong, concurrent relationship with the Colorado Statewide Reading assessment.

In Oregon, Good et al. (2001) reporting on a different study of third grade ORF benchmarks, concluded that 96% of the sample students who met the third grade oral reading fluency benchmark also met or exceeded the Oregon Statewide Assessment of reading grade level benchmark. This finding provided evidence supporting the use of ORF benchmarks for judging whether third grade students are on-track for passing the Oregon statewide reading assessment.

To counter an argument that ORF is merely a measure of decoding automaticity, Shinn and Good (1992) used confirmatory factor analysis to investigate the contribution of ORF (to comprehension) with respect to theoretical reading models. Using a sample of third ($n = 114$) and fifth grade ($n = 124$) students, the investigators examined whether ORF played a significant role in a single-factor model of reading, or whether it should be defined as: (a) a decoding construct, (b) a comprehension construct, or (c) as a separate construct. The outcome measure was the Comprehension subtest of the Stanford Diagnostic Reading Test (Karlsen, Madden, & Gardner, 1975).

For third grade students, a single factor model of reading, which included decoding, fluency and comprehension, could not be rejected. For fifth grade students, a single factor model of reading was rejected, however. For fifth grade students, fluency and decoding represented one factor while comprehension represented a second factor.

The authors argued that ORF fits into many developmental theories, including the Stages of Reading Development (Chall, 1983) and "can be validated as a measure of general reading achievement, including comprehension…the face validity arguments regarding CBM oral reading measures should be put to rest" (p. 471). The authors claimed the results from their study provided evidence of ORF as a measure of reading comprehension, despite claims of validity issues related to the perception that a measure of oral reading fluency should be considered to be only a measure of decoding (Shinn & Good, 1992).

Research has validated the use of ORF as an indicator of basic reading competence, including comprehension, for elementary students, especially in the third through fifth grades (Fuchs et al., 1988; Good et al., 2001; Kame'enui & Simmons, 2001; Shinn & Good, 1992; Silberglitt & Hintze, 2005; Wanzek et al., 2010; Wood, 2006). What remains unclear is whether oral reading fluency is a valid CBM to predict reading comprehension, as measured by a statewide reading assessment, with adolescent students.

Oral Reading Fluency and Maze Measures with Middle School Students

The question of whether ORF is a valid CBM of reading comprehension for adolescent readers in the middle school years has been explored to a lesser degree than with elementary students (Espin, Wallace, Lembke, Campbell, & Long, 2010; Fore III, Boon, Burke, & Martin, 2009; Fuchs et al., 1988; Ticha, Espin, & Wayman, 2009). Studies including eighth-grade students have examined both the concurrent validity of ORF and Maze with standardized tests as well as their predictive validity with statewide reading assessments. Table 1 provides a summary of concurrent and predictive validity research studies utilizing curriculum-based measures, ORF and Maze, with eighth-grade

students. The studies reported moderate to strong positive correlation coefficients for ORF and Maze with standardized measures and statewide reading assessments.

Oral Reading Fluency and Standardized Tests

Concurrent validity studies of ORF to standardized tests have yielded results that support the use of ORF with the middle school population (Fore III et al., 2009; Fuchs et al., 1988). In a study of 70 junior high and middle school students with reading disabilities, the correlation of ORF and three other, direct measures of comprehension to the Reading Comprehension subtest of the Standard Achievement Test were examined (Fuchs et al., 1988). The three direct measures of comprehension were: question answering, passage recall, and cloze. Criterion validity coefficients for the question answering, the recall, and the cloze measures were .82, .70, and .72, respectively. The correlation coefficient for oral reading fluency was .91. All three direct measures of comprehension were moderately correlated with the criterion measure, but the ORF measure was the most strongly associated with the capacity to read passages and answer questions about the passages on the Reading Comprehension subtest of the Standard Achievement Test (Fuchs et al., 1988). This study supports the use of ORF with lower performing middle school readers, students identified with a reading disability.

For middle school students with identified emotional and behavioral disorders, Fore III et al. (2009) found less impressive results for the use of ORF. In this study, ORF was correlated with two subtests from the Woodcock Johnson III Tests of Achievement (Woodcock, McGrew, & Mather, 2001): the Passage Comprehension (WJ-PC) subtest and the Reading Fluency (WJ-RF) subtest (Fore III et al., 2009). A statistically significant, positive correlation between ORF and the WJ-PC subtest was reported,

although the strength of the relationship was moderate ($r = .452$, $p < .05$). The weaker correlation between ORF and the two standardized measures, WJ-PC and WJ-RF, may be due to the variability in behavior and academic performance of the population in the sample, or it could be an indicator of reduced validity of ORF as a measure of reading comprehension and fluency for older students.

In a small study ($n = 35$) of eighth-grade students Ticha et al. (2009), reported the concurrent validity correlation coefficient of ORF and the Woodcock Johnson III Pretest was strong and statistically significant ($r = .87$, $p < .008$). The results of concurrent validity studies with these eighth-grade students indicated moderate to strong, positive relationships between ORF and other, standardized reading outcome measures. However, additional research with a more representative sample of eighth-grade students including students with and without disabilities, are needed to further investigate the relationship between ORF and other standardized measures of reading comprehension. With additional research and evidence, the validity of using ORF at the eighth-grade level can be clarified. Recent studies have also examined the predictive validity of ORF to performance on statewide reading assessments for middle school students, including eighth grade (Espin et al., 2010; Silberglitt, Burns, Madyun, & Lail, 2006; Ticha et al., 2009). The next section explores the research in this area.

ORF and Statewide Reading Assessments

Oral reading fluency assessments provide educators meaningful assessment data to inform instruction for elementary students (Fuchs & Fuchs, 2004; Hasbrouck & Tindal, 2006; Hosp & Fuchs, 2005). Although research has validated the use of ORF with elementary students, its utility at the middle grades remains unclear (Espin et al., 2010;

Fuchs et al., 1988; Megert, 2010; Silberglitt et al., 2006; Smith, 2008; Ticha et al., 2009). To date, the predictive studies of *eighth-grade* students using ORF and statewide reading assessments are from one state, Minnesota (Espin et al., 2010; Fore III et al., 2009; Silberglitt et al., 2006; Ticha et al., 2009). Additional research is needed to validate the utility of ORF with diverse samples of eighth-grade students in predicting other statewide reading assessments.

Espin et al. (2010) evaluated the technical adequacy of ORF as a predictive measure of performance on the Minnesota statewide reading assessment, using performance data from 238 academically and ethnically diverse eighth-grade students. In the study, ORF demonstrated a strong, statistically significant predictive validity coefficient ($r = .78$, $p = <.01$) related to the Minnesota statewide reading assessment, the Minnesota Basic Standards Test (MBST). Results of the study support the use of ORF as a predictor of performance on the MBST, for a population similar to the sample population of eighth-grade students taking the MBST. The researchers acknowledged the use of the MBST as a limitation. The MBST is a state test that represents "survival skill levels of reading performance" (Espin et al., 2010, p. 73) and has reported technical inadequacies in the areas of validity and reliability. A strength of the study was the ethnic diversity of the sample, with 58% English Speakers, 42% non-English speakers, 34% Caucasian, 24% Asian American, 20% African American, 19% Hispanic, and 3% Native American.

In another study of eighth-grade students Ticha et al. (2009) examined the predictive validity of ORF with the MBST. In this study, the predictive validity correlation coefficient of ORF-MBST was statistically significant and demonstrated a

strong, positive relationship ($r = .77$, $p > .008$). The study results support the use of ORF as an indicator of future performance on the MBST, for an eighth-grade population demographically similar to the study sample.

Although the results of the previous studies provided evidence to support the predictive validity of ORF with one statewide reading assessment for eighth-grade students, another study warned against the use of ORF to make predictions about performance on a statewide assessment for eighth-grade students (Silberglitt et al., 2006). A longitudinal study including 5,472 students across grades 3, 5, 7, and 8 demonstrated with advancing grade levels, the magnitude of the relationship between ORF and the statewide reading assessment declined (Silberglitt et al., 2006). The eighth-grade sample in this study included 843 students. The criterion assessment used for comparison was the Basic Standards Test- Reading (BST-R), the Minnesota statewide assessment designed to measure a minimum benchmark of performance that students must meet or exceed to receive a high school diploma.

For eighth-grade students, the correlation between the ORF and BST-R was moderately strong and statistically significant ($r = .60$, $p < .001$). The correlation coefficients reported in this study were lower than those identified in the Espin et al. (2010) and the Ticha et al. (2009) studies. These findings are somewhat counterintuitive because the standardized passages used in the Silberglitt et al. (2006) study were reported to have excellent reliability and validity evidence (Shinn & Shinn, 2002b). As such, it would seem that if ORF were a valid assessment of reading comprehension for eighth-grade students, this study would have produced stronger correlations. Plausible explanations for the lower correlations include: (a) the different criterion measures used

in each study, (b) a difference in the sampling techniques, and (c) sample demographics. The Silberglitt et al. (2006) study had much larger sample sizes in the lower grades than in the upper grades, including the eighth grade, because fewer years of data were considered for the upper grades than for students in the lower grades. However, with respect to the available research for eighth-grade students, the sample size in this study was relatively large. The larger sample size also included more students who were not identified as having a disability, and the lower correlation reported may be an indicator that ORF is a less effective measure for the *general* eighth-grade population. Previous research (Yovanoff, Duesbery, Alonzo, & Tindal, 2005) has suggested that for students who are progressing as expected, ORF plateaus in late middle school, becoming a less informative measure of reading progress. However, for students who are still struggling to read and need to make progress, whether they have a disability, or are an English Learner or a student from a lower SES family, ORF appears to continue to be a sensitive measure (Fore III et al., 2009; Fuchs et al., 1988; Wiley & Deno, 2005; Yovanoff et al., 2005).

The predictive validity studies of ORF and statewide reading assessment demonstrated moderate to strong, positive correlation coefficients in the eighth-grade samples in one state. Additional predictive validity studies of ORF and other statewide reading assessments are needed to validate its use with the general eighth-grade population in the United States.

Maze and Standardized Tests

The use of CBM Maze assessment may also offer information that is informative and yet easier and less expensive to collect than ORF data (group vs. individual

administration). The concurrent and predictive validity research on the use of Maze with eighth-grade students follows.

The Maze has been suggested to have face validity because the task of choosing the correct word in the context of a passage emulates what it claims to measure: reading comprehension (Parker, Hasbrouck, & Tindal, 1992). For this reason, secondary educators may more readily accept the Maze as a meaningful measure over ORF. To determine the validity of using CBM Maze as an indicator of comprehension for eighth-grade students, studies have examined the concurrent validity of Maze with standardized tests of reading comprehension (Fore III et al., 2009; Ticha et al., 2009). In a study of eighth-grade students with behavioral disorders Fore III et al. (2009) reported a strong, positive correlation between Maze and WJ-PC ($r = .89$, $p < .05$). For this sample, the strength of the relation between Maze and WJ-PC was much stronger than the relation between ORF and WJ-PC ($r = .45$).

With a more ethnically diverse, but smaller sample of eighth-grade students ($n = 35$), Ticha et al. (2009) found the concurrent validity correlation coefficient of Maze with the WJ-III Pretest was statistically significant, strong and positive ($r = .88$, $p < .008$). This study provided more evidence supporting the use of Maze with eighth-grade students as an indicator of reading ability, and reading comprehension in particular.

More evidence about the concurrent validity of Maze with standardized tests of reading comprehension for eighth-grade students is needed to generalize to the greater population of eighth-grade students. However, another approach to analyzing the potential of Maze for use with eighth-grade students, researchers have examined concurrent and predictive relations between the Maze and statewide reading assessments.

Maze and Statewide Reading Assessments

To explore the utility of Maze in predicting eighth-grade student performance on a statewide reading assessment, researchers have examined predictive validity relations between the Maze and statewide reading assessments (Espin et al., 2010; Silberglitt et al., 2006; Ticha et al., 2009). Table 1 provides a summary of these findings. Espin et al. (2010) evaluated the technical adequacy of the Maze as a predictive measure of performance on the MBST. As a predictive measure, the Maze demonstrated a statistically significant, strong, positive predictive validity coefficient when related to the MBST ($r = .82$, $p < .01$). The results of this study provided evidence to support the use of both CBM, ORF and Maze, with eighth-grade students to predict performance on the MBST (Espin et al., 2010).

To further evaluate the potential of Maze to predict performance on statewide assessments, Ticha et al. (2009) analyzed the coefficients for predictive validity of Maze with the MBST. The result was a statistically significant, strong, and positive correlation coefficient, Maze-MBST ($r = .82$, $p > .008$). The study sample was diverse with 49% white, 46% African American, 3% Asian, and 3% Hispanic. A limitation of this study includes the questionable reliability and validity of the MBST, which the authors openly cite as having limited reliability and validity (Ticha et al., 2009). Otherwise, the study provides evidence to support the use of Maze with an eighth-grade population similar to the study sample.

In a longitudinal study of how the relation between Maze and a statewide assessment changes with advancing grade levels, Silberglitt et al. (2006) found that for eighth-grade students, the concurrent validity relationship of the Maze to the Basic

Standards Test- Reading (BST-R) was moderately, positively correlated (r = .48). The sample included 1,028 eighth-grade students, primarily Caucasian (94%). This finding provided contrary evidence with a larger sample of eighth-grade students, including students in a variety of performance ranges, indicating the Maze was not as strongly correlated with a statewide reading assessment.

Table 1

Concurrent and Predictive Validity for ORF and Maze Among Eighth-Grade Students

Study	*n*	Grade	CBM	State Assess	WJII-Pretest	WJII-PC	WJIII-RF
Fore III, Boon, Burke, & Martin (2009)	55	6, 7, 8	ORF			.45	.47
			Maze			.89	.22
Espin, Wallace, Lembke, Campbell, & Long (2010)	236	8	ORF	.78			
			Maze	.78			
Ticha, Espin, & Wayman (2009)	35	8	ORF	.77	.87		
			Maze	.82	.88		
Silberglitt, Burns, Badyun, & Lail (2006)	843	8	ORF	.50			
	1028		Maze	.48			

With respect to sub-populations, research studies have provided some evidence about the use of CBM ORF and Maze with various middle school-aged samples of students with disabilities (Fore III et al., 2009; Fuchs et al., 1988). Although the relationship between income and the overall reading development for middle school students has not specifically been studied, there is a national trend indicating a negative relationship between low family-income and reading development for eighth-grade students. An explanation of this trend follows in the next section.

The Relation between Socio Economic Status and Reading Development

The 2003-09 NAEP (U.S. Department of Education, 2009) results revealed an established pattern for lower-income eighth-grade students. Students who are *eligible for free lunch* consistently scored the lowest, with an average score of 247, the *reduced-price lunch* students scored slightly better with an average score of 256, while the students *not eligible for free lunch* scored the highest, with an average score of 273. These results suggest that family income, as demonstrated by the free and reduced-price lunch data, is related to eighth-grade student reading performance. Students who qualify for free or reduced-price lunch read less well, as measured by NAEP, than do students who are not eligible for free or reduced-price lunch (U.S. Department of Education, 2009).

The available data and research findings are suggestive, but not conclusive, regarding the potential for using CBM ORF and Maze with eighth-grade students. Research with eighth-grade students including students with the following characteristics is needed: (a) with disabilities, (b) without disabilities, and (c) from a variety of income levels. The following research study incorporated these features to further define and clarify the validity of using CBM, ORF and Maze, with eighth-grade students.

Purpose of the Study

The rationale for using CBM, ORF and Maze, to predict performance on the statewide assessment is to inform teachers in an efficient and timely manner which students are at risk of failing the summative, high-stakes, statewide reading assessment. With data available early in the school year, teachers can adjust instruction as needed to maximize student learning and improve outcomes on the statewide assessment. Researchers have recently begun to examine the utility of CBM in the eighth-grade, but

the non-representative samples are not generalizable to the greater population of students (Espin et al., 2010; Fore III et al., 2009; Silberglitt et al., 2006; Silberglitt & Hintze, 2005; Ticha et al., 2009; Yovanoff et al., 2005). Schools need more evidence of valid and efficient measures to predict, early in the eighth-grade school year, end of year outcomes on a high-stakes, statewide reading assessment.

The purpose of this study was to examine the concurrent and predictive validity of the CBMs, ORF and Maze, in three eighth-grade student samples. Fall and spring ORF and Maze scores were correlated to students' scores on the Oregon statewide reading assessment, the OAKS-R. The results of the study will reveal for six different samples of eighth-grade students during three years (06-07, 07-08, 08-09) the strength of the concurrent and predictive relations between CBMs, ORF and Maze, and the OAKS-R.

If the Maze is more strongly related to the OAKS-R as a measure of reading comprehension (concurrent validity) and as a predictor of performance on the statewide assessment, there are a few potential benefits for a district to consider in using the Maze over ORF as a measure of progress. These potential benefits include: (a) the group-administered approach saves instructional time, (b) the greater face validity of the Maze measure with teachers, and (c) potential district cost savings over a system that requires additional staff time to capture individual student oral reading fluency data.

If ORF is more strongly related to the eighth-grade OAKS-R, then districts can consider the utility of ORF as a measure of comprehension and weigh the benefit of information over the cost of administration. Districts will want to consider other research studies and the particular samples utilized, to determine what makes sense for their particular district. If ORF and Maze seem to be nearly equally related to the eighth-grade

OAKS-R, districts may want to prioritize the cost savings and face validity of the Maze over the ORF.

Research Questions

1. What is the concurrent relation between CBMs, Oral Reading Fluency and Maze, and the eighth-grade Oregon Assessment of Knowledge and Skills- Reading/Literature?
2. What is the unique relation between the beginning of the academic year performance on CBMs, Oral Reading Fluency and Maze, and performance on the eighth-grade Oregon Assessment of Knowledge and Skills- Reading/Literature?
3. Is there an interaction between the demographic characteristics of students, disability status and socio-economic status as measured by free or reduced-price meal status, and each of the CBMs, Oral Reading Fluency and Maze, in the prediction of performance on the eighth-grade Oregon Assessment of Knowledge and Skills- Reading/Literature?

This study was conducted with the hypothesis that current results would replicate previous findings with a different population of eighth-grade students and a different statewide assessment. That is, scores on the two CBMs, ORF and Maze, would demonstrate statistically significant, moderate to strong concurrent and predictive validity with student scores on the eighth-grade Oregon statewide assessment of reading, the OAKS-R. Additionally, the disability and poverty status of students would interact with ORF and Maze scores, yielding a relatively stronger relation between the ORF and Maze measures and OAKS-R score for low income students and students with disabilities.

CHAPTER III

METHODOLOGY

The current study included analysis of extant data from three years (06-07, 07-08, 08-09) of district-wide eighth-grade reading assessments, oral reading fluency and Maze measures, and the Oregon statewide reading assessment for eighth-grade students, the Oregon Assessment of Knowledge and Skills/Reading Literature (OAKS-R).

Research Design

The study utilized a non-experimental, descriptive research design that used correlation and regression analyses to examine the concurrent and predictive validity of CBM in multiple samples of eighth-grade students. Specifically, the bivariate and unique relation between fall and spring ORF and Maze scores and scores on the OAKS-R were estimated. An alpha value of .05 was used as the cutoff criteria for all statistical significance tests.

To answer the first research question, what is the concurrent validity of the CBMs, ORF and Maze, and the OAKS-R, the correlation between spring ORF and spring OAKS-R, and spring Maze and spring OAKS-R, was estimated. To answer the second research question, what is the predictive validity of the CBMs, ORF and Maze, OAKS-R was regressed on fall ORF and fall Maze. To answer the third research question, a multiple regression analysis was conducted to estimate the independent and joint

relations between disability and socio-economic status and each fall-administered CBM, ORF and Maze, in the prediction of eighth-grade spring OAKS-R assessment scores.

Study Participants

Utilizing three years of eighth-grade extant data, six samples were created: three to analyze concurrent validity (Spring CBM- Spring OAKS-R), and three to analyze predictive validity (Fall CBM-Spring OAKS-R). The participants in this study who met the inclusion criteria were included in the appropriate samples. Participants in the samples met the following criteria: (a) enrolled as an eighth-grade student during one of the following school years: 06-07, 07-08, 08-09, (b) had an eighth-grade fall ORF *and* fall Maze *and/or* had an eighth-grade spring ORF *and* spring Maze, and (c) took the eighth-grade Oregon Assessment of Knowledge and Skills- Reading/Literature. Table 2 presents the demographic characteristics of the sample.

Table 2

Demographic Characteristic Percentages for Samples in Study

	F 06	Sp 07	F 07	Sp 08	F 08	Sp 09
Sex						
Girls	47	47	47	47	53	54
Boys	53	53	53	53	48	46
Race/Ethnicity						
White	77	77	71	71	77	69
Hispanic	12	13	19	18	16	15
Disability Status						
Special Education	15	15	14	12	15	16
Economic Disadvantage	48	47	49	50	51	52
Total Number of Students	377	375	350	338	379	379

Setting

The data were obtained from a school district located in a moderately-sized city in the Pacific Northwest. The city had an approximate population of 155,000 residents and contained two school districts. The school district where the data were collected was the smaller of the two districts and was located in a suburban setting. The school district was considered a medium-sized school district with approximately 5,700 students. Within the district there were five elementary schools, two middle schools, two K-8 schools, one comprehensive high school, and an alternative high school. The school district primarily served white students, but with an increasing Latino population, was becoming progressively diverse. The sample included students who met federal and state criteria as students with a disability in accordance with the Individuals with Disabilities Education Act 2004 (IDEA, 2004). The percent of the sample with an identified disability ranged from 11.8% to 15.6% during the years of the study. The socio-economic status of the students was measured by identifying students who qualified for free or reduced-price meals. The range of students qualifying for free or reduced meals ranged from 48.3% to 52% during the years of the study.

Measures

The extant data analyzed in this study included the predictors AIMSweb Oral Reading Fluency and AIMSweb Maze student performance data, and the criterion variable was the RIT score on the eighth-grade Oregon Assessment of Knowledge and Skills- Reading/Literature (OAKS-R). During the time frame of the data collection in the study, the school district routinely collected ORF and Maze performance data on all

middle school students during a school-wide screening, three times a year, in September, January, and May.

The National Center on Response to Intervention's Technical Review Committee indicated that both AIMSweb measures, ORF and Maze, have established convincing, direct evidence of technical adequacy as progress monitoring tools (National Center for Response to Intervention, 2009). The Technical Review Committee (TRC) evaluates the scientific rigor of progress monitoring tools against an independently established set of criteria. Oral Reading Fluency is a one-minute, oral reading assessment of rate and accuracy. Maze is a three-minute, silent reading test of fluency and comprehension.

Although the testing window for the OAKS-R was October to May, the District practice was to administer between February and May during the years, 2007-09. Students had up to three opportunities to take the OAKS-R. The test was different each time the student took the test, reducing the potential for a practice effect. Students' best score was recorded as their performance on the assessment for the year. The OAKS-R reported student performance in several categories: (a) Vocabulary, (b) Read to Perform a Task, (c) Demonstrate General Understanding, (d) Develop an Interpretation, (e) Examine Content and Structure: Informational Text, and (f) Examine Content and Structure: Literary Text. The score reporting categories describe the types of questions that students must answer.

Oral Reading Fluency Measures

Standardized AIMSweb passages were selected and administered as detailed in the AIMSweb Training Workbook (Shinn & Shinn, 2004). See Appendix A for a sample eighth-grade Oral Reading Fluency passage, Appendix B for Oral Reading Fluency

administration directions, and Appendix C for Oral Reading Fluency scoring rules. AIMSweb Oral Reading Fluency passages have fully met the seven standards derived by the Standards for Educational and Psychological Testing developed by the joint committee appointed by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement Used in Education and the Individuals with Disabilities Education Act (Shinn & Shinn, 2002b). The seven standards are: (a) sufficient number of alternate forms with evidence of equal difficulty, (b) rates of improvement specified, (c) benchmarks specified, (d) evidence of improved student learning or teacher planning, (e) sensitivity to student improvement, (f) reliability, and (g) validity. The AIMSweb eighth-grade passages have established alternate form reliability (r = .90), a mean of 147.3 words read correct, a standard deviation of 42, and a standard error of measurement of 13.3 (Shinn & Shinn, 2002b). AIMSweb reported extensive criterion-related validity studies with strong correlations (e.g., Woodcock Reading Mastery Test r = .91; Iowa Test of Basic Skills r = .83) (Howe & Shinn, 2002).

Administration and Scoring: Oral Reading Fluency Test

The oral reading fluency assessment was individually administered to all eighth-grade students within a one-week window in the months of September, January, and May. Each school conducted the testing in a quiet location. Prior to each collection of ORF data, the district ensured all test administrators had been trained on standard administration and scoring procedures as outlined in the AIMSweb Training Workbook (Shinn & Shinn, 2002a). Trained teachers, administrators, and educational assistants collected, scored, and entered student performance data, including total words per minute (TWPM), correct words per minute (CWPM), and numbers of errors, into the district

database (see Appendix C for scoring rules). The scoring procedures require the test administrator to: (a) encourage the student to do his/her best reading, (b) cue the student to "start", (c) follow along with the student, crossing out words read incorrectly, (d) at the end of the minute, the assessor should say "stop" and, (e) indicate on the passage, the last word read. Should the student stop or struggle with a word for more than three seconds, the administrator should read the word and then mark the word as incorrect.

Maze Measures

Standardized AIMSweb Maze passages were selected and administered to students. The passages were approximately 400 words in length (see Appendix D). AIMSweb Maze passages have also fully met the seven standards derived by the Standards for Educational and Psychological Testing (Shinn & Shinn, 2002b). Although AIMSweb Maze passages fully met the aforementioned criteria, the reliability and validity studies were conducted at the lower grades, and there were no studies that included a general population of eighth-grade students, which was the focus of this dissertation study (Shinn & Shinn, 2002a).

Administration and Scoring: Maze Test

The AIMSweb Maze measures were administered by teachers following a standardized format for "Older Students and Students Familiar with Maze Directions" (p.16) as outlined in the AIMSweb Training Workbook (Shinn & Shinn, 2002b), see Appendix E. The Maze assessment was administered in a large group setting, typically during reading class. Scoring the Maze includes determining the Number of Words (items) Correct (NWC). An answer is correct if the student circles the word that matches the correct word on the scoring template. An answer is incorrect if the student either

circles an incorrect word or omits word selections (skips) other than those the student was unable to complete before the 3 minutes expired. The scorer puts a slash through incorrect words. To get the NWC score, the scorer subtracts the number of incorrect word choices from the correct word choices. The classroom teacher scored the Maze measures, and then the results were sent to the district office for database entry. Appendix F outlines the Maze scoring rules.

Oregon Assessment of Knowledge and Skills- Reading/Literature

The Oregon Assessment of Knowledge and Skills- Reading/Literature (OAKS-R) was the statewide reading assessment for eighth-grade students. To meet adequate yearly progress targets in reading during the school year 2010-11, all districts in Oregon needed to ensure that at least 95% of their students took the OAKS-R assessment and 70% of all students needed to pass the assessment. By 2014, 100% of all students will need to pass the OAKS-R. To pass the OAKS-R, beginning in third grade, students need to demonstrate grade level reading competence.

The eighth-grade OAKS-R uses three types of reading selections to gain a variety of assessment information. These passages include: (a) reading for literary experience, (b) reading to gain information, and (c) reading to perform a task. Each reading selection is followed by at least three, and no more than ten, knowledge and skills test items. Each test item includes four answer choices.

Reliability and Validity

The Oregon Assessments of Knowledge and Skills are "rigorously examined" (ODE, 2007a, p. 2) in reference to the guidelines provided in the Standards for Educational and Psychological Testing (American Educational Research Association,

American Psychological Association, & National Council on Measurement in Education, 1985, 1999). The Standards for Educational and Psychological Testing are a set of standards developed jointly by the American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education. Reliability of classification accuracy analyses indicated alternate forms of the OAKS-R including the paper-pencil, TESA, and the short form, were similar in classification accuracy across the range of ability levels (ODE, 2007a). Concurrent validity studies of the eighth-grade OAKS-R to other measures of similar constructs yielded statistically significant correlations ($p < .001$) with: (a) California Achievement Test $r = .77$, (b) Iowa Test of Basic Skills, $r = .82$, (c) Northwest Evaluation Association Subject Test, $r = .80$, and (d) Lexile Scale, $r = .73$ (McCall, 2005; ODE, 2001, 2007a).

The Oregon Department of Education (ODE, 2007a) reported evidence of content validity including: (a) rigorous content standards reviewed by Oregon educators, parents and other concerned citizens, (b) clear links between test content and the content standards and their corresponding performance levels, (c) a consensus-driven test item development process, and (d) ongoing studies to evaluate and increase the alignment of instruction, assessment, and content standards.

Administration and Scoring: OAKS-R

The OAKS-R is primarily administered via a computer-based format called the Technology Enhanced Student Assessment (TESA). In this format, students are provided up to three opportunities to participate in an adaptive assessment delivered by computer. The adaptive assessment format allows for the accuracy of student responses in one passage to determine the next passage and the next set of questions with which the

student will be presented. Although the assessment window for the computer-based assessment format stretched from October through May each school year included in this study, the district practice was to administer the computer-based format of the assessment in the spring. Students took the OAKS-R between February and May, which allowed for maximum learning time during the school year prior to the evaluation. A paper/pencil version of the test was also available, but this version had several limitations: (a) the district must order the paper/pencil version months in advance, (b) students were only given one attempt with the paper/pencil version, (c) the paper/pencil version was not adaptive, (d) the student did not receive a score immediately, and (e) the paper/pencil testing window was significantly shorter, about one month in length, March to April. All schools in the district chose to administer all assessments via the computer-based format, with exceptions made for the paper-pencil version, on a case-by-case basis.

Proctors, mostly administrators, teachers, and educational assistants, were trained yearly prior to administration of the statewide assessment. Proctors were trained to strictly adhere to the standardized administration protocols. All students were allowed several accommodations, one of which was to have unlimited time to complete the assessment. With this accommodation, the testing session may have spanned several days. While the testing session takes only 60-75 minutes to complete, some students may not have been able to attend for that long. For students who needed to take the assessment in shorter time frames, such an accommodation is allowable, which may have resulted in the assessment administration spanning several days.

The OAKS-R yields a Rasch Unit (RIT) score that falls into one of three reporting categories, based on cut scores that reflect student performance on the benchmark

assessment. These performance level categories are: *exceeds*, *meets*, and *does not meet* the benchmark. Scoring of the OAKS-R computer-based format is instant; yielding a RIT score that is available to the student, school, and district. With the immediate feedback, schools can determine whether students will need to be re-assessed or if they have completed their statewide reading assessment for the school year. Scoring of the paper/pencil version of the OAKS-R follows a prescribed protocol; the assessment must be mailed to the Oregon Department of Education where it will be scored. In this format, schools do not receive the immediate feedback on student performance.

Data Analysis

A correlational analysis was conducted to estimate the concurrent relations between the spring-administered CBMs, ORF and Maze, and the spring-administered eighth-grade OAKS-R. A multiple regression analysis was conducted to estimate the independent relation between each of the fall-administered CBMs, ORF and Maze, and the spring-administered eighth-grade OAKS-R. Additional multiple regression analyses were conducted to estimate the independent and joint relations between disability and socio-economic status and each fall-administered CBM, ORF and Maze, in the prediction of spring OAKS-R scores.

CHAPTER IV

RESULTS

In this chapter, the results associated with each of the study research questions are provided. The first research question was designed to facilitate the investigation of the concurrent relation between the spring-administered CBMs of Oral Reading Fluency (ORF) and Maze, with the OAKS-R. To answer Question 1, a correlational analysis was conducted to index the relation between each spring-administered CBM and the OAKS-R. The second question was designed to examine the extent to which ORF and Maze uniquely predicted performance on the OAKS-R. To answer Question 2, linear regression analyses were conducted on CBM and OAKS-R data collected from three cohorts of eighth-grade students.

The third research question was designed to evaluate the independent and joint relations between demographic characteristics of disability status (SpEd) and socio-economic status (FRL), and ORF and Maze scores, in the prediction of student performance on the OAKS-R. To answer Question 3, multiple regression analyses were conducted using data from students in each of the three eighth-grade cohorts with OAKS-R as the dependent variable and ORF, Maze, socio-economic status (FRL) and disability status (SpEd) as the predictors. Based on the results of the individual cohort analyses, the cohorts were combined, and one final multiple regression analysis was conducted.

Prior to presenting the results of the first research question, decision rules that

were used to identify the analytic samples and descriptive statistics associated with each cohort are provided. The OAKS-R was administered during the months of January through May while the CBMs were administered in the second and third week of May. A decision was made to exclude OAKS-R data prior to April 1 of each year in order to facilitate the analysis of the concurrent relations between study assessment measures. The decision resulted in a 6-week window in which both the OAKS-R and CBMs were included for the concurrent validity analysis. However, it should be noted that because of statewide technological difficulties that resulted in a change to mandatory paper-pencil assessments in the middle of the assessment year, Cohort One was excluded from the concurrent validity analysis. See Table 3 for descriptive statistics on each of the assessment measures examined within the Cohort Two and Cohort Three analytic samples.

Table 3

Descriptive Statistics of Concurrent Measures

	Cohort 2 (n = 126)			Cohort 3 (n = 160)		
	ORF	Maze	OAKS-R	ORF	Maze	OAKS-R
Mean	159.55	25.43	231.83	154.83	28.54	233.16
Std. Dev.	47.47	9.34	8.36	35.93	8.62	7.92
Minimum	51.00	3.00	211.00	52.00	4.00	211.00
Maximum	255.00	48.00	258.00	251.00	47.00	256.00

Analytic Samples

Concurrent validity was evaluated in two student cohorts: Cohort Two (Spring 2008) and Cohort Three (Spring 2009). Chi-square tests were first conducted to determine if there were statistically significant differences between cases that were excluded and students included in the analytic samples. Excluded cases were those whose OAKS-R score fell outside the six-week window in which both the CBM and ORF were administered. Comparisons were conducted on the following demographic variables: (a) gender, (b) disability status, as measured by special education identification (SpEd), (c) socio-economic status, as measured by free and reduced-price meal status (FRL), and (d) ethnicity, which was coded as Caucasian or Other Ethnicity (not Caucasian). All minority students were analyzed as a group (Other Ethnicity) against Caucasian students, as the sample sizes for other ethnic groups were quite small. Cohort Two included the following number of students as Other Ethnicity: American Indian/Alaskan Native- 2, Asian- 7, Black- 3, Hispanic- 22. Cohort Three included as Other Ethnicity: American Indian/Alaskan Native- 5, Asian- 2, Black- 4, Hispanic- 30.

In Cohort Two, no statistically significant differences were found between the excluded cases and the cases included in the analytic sample. However, in Cohort Three, there were statistically significant differences with respect to ethnicity ($p < .01$) and FRL ($p < .05$). The Cohort Three analytic sample was not completely representative of the student population. The Cohort Three analytic sample was more ethnically and economically diverse sample than the larger eighth-grade student population and included a greater percentage of students who were not Caucasian and who qualified for free and

reduced-price meal status. See Table 4 for the concurrent validity sample descriptive statistics.

Table 4

Concurrent Validity Sample Descriptive Statistics

	Cohort 2				Cohort 3			
	Included (N = 126)		Excluded (N = 210)		Included (N = 160)		Excluded (N = 216)	
	Freq	%	Freq	%	Freq	%	Freq	%
Boys	63	50	112	54	77	48	97	45
White	87	69	160	76	111	69	180	83
Sp Ed	20	16	18	9	30	19	26	12
FRL	65	52	97	46	95	59	102	47

Research Question 1: Concurrent Validity of Measures

The concurrent relation between spring administered CBM, ORF and Maze, with the OAKS-R was indexed by Pearson product moment correlation coefficients. The results of the analysis identified positive, moderately-sized correlations between ORF, Maze, and OAKS-R. All correlations were statistically significant ($p < .05$) indicating that CBM, ORF and Maze, were concurrent measures with the eighth-grade OAKS-R (See Table 5).

Table 5

Concurrent Correlations: ORF, Maze, with OAKS-R

	Cohort 2	Cohort 3
Oral Reading Fluency	.64*	.58*
Maze	.58*	.54*

*$p < .01$

Research Question 2: Do ORF and Maze Predict Performance on the OAKS-R?

The second research question utilized ORF and Maze as predictors of performance on the spring-administered, eighth-grade OAKS-R. Three cohorts of eighth-grade student data were analyzed.

Sample Description

Included in the samples were students with fall ORF and Maze scores who also took the standard administration of the OAKS-R, with and without accommodations. Excluded from the analyses were students who did not take the standard administration of the OAKS-R. Excluded students included students whose parents excluded them from the statewide assessment, students with significant disabilities who took an alternate assessment, as well as students who did not take all three assessments in the cohort year. For Cohort One, there were 49 students excluded from the sample, in Cohort Two, 45 students were excluded, and in Cohort Three, 46 students were excluded. Table 6 presents descriptive statistics on each of the measures included in the predictive analysis.

Table 6

Descriptive Statistics of Predictive Analysis Measures

	Cohort 1 (n = 374)			Cohort 2 (n = 348)			Cohort 3 (n = 374)		
	ORF	Maze	OAKS	ORF	Maze	OAKS	ORF	Maze	OAKS
Mean	158.66	21.90	233.62	160.70	26.40	234.03	149.75	27.62	234.82
SD	41.95	7.57	9.54	46.56	9.68	8.30	39.12	8.10	8.10
Min.	43.00	2.00	208.00	34.00	6.00	215.00	48.00	8.00	212.00
Max.	255.00	45.00	257.00	255.00	57.00	266.00	255.00	48.00	280.00

ORF and Maze as Predictors of Student Scores on the Eighth-Grade OAKS-R

To answer the second research question, linear regression analyses were performed with OAKS-R as the dependent variable and ORF and Maze as predictor variables.

The Cohort One analysis revealed that the overall regression of ORF and Maze scores on OAKS-R scores was statistically significant, $F(2, 371) = 58.78$, $p < .01$, $R^2 = .36$. The regression weight associated with ORF indicated that for every one-word increase in ORF, OAKS-R increased .05 RIT points. For a positive difference of 40 correct-words-per-minute (CWPM) in ORF (approximately one standard deviation), the OAKS score is predicted to be about two points higher, or a quarter of a standard deviation in RIT units. The regression weight associated with Maze indicated that for every one-word increase in Maze, OAKS-R increased by .51 RIT points. For a positive difference of 8 number-words-correct (NWC) in Maze (approximately one standard deviation), the OAKS score is predicted to be about four points higher, or a half of a

standard deviation in RIT units. The Cohort One squared semi-partial correlations revealed that ORF and Maze uniquely accounted for 3% and 8% of the variance in OAKS-R, respectively.

The Cohort Two analysis revealed that the overall regression of ORF and Maze scores on OAKS-R scores was statistically significant, $F(2, 346) = 41.45$, $p < .01$, $R^2 = .40$. The regression weight associated with ORF indicated that for every one-word increase in ORF, OAKS-R increased .06 RIT points. For a positive difference of 46 CWPM in ORF (approximately one standard deviation), the OAKS score is predicted to be about four points higher, or a half of a standard deviation in RIT units. The regression weight associated with Maze indicated that for every one-word increase in Maze, OAKS-R increased by .27 RIT points. For a positive difference of 10 NWC in Maze (approximately one standard deviation), the OAKS score is predicted to be about three points higher, or a half of a standard deviation in RIT units. The Cohort Two squared semi-partial correlations revealed ORF and Maze uniquely accounted for 5% and 4% of the variance in OAKS-R, respectively.

The Cohort Three analysis revealed that the overall regression of ORF and Maze scores on OAKS-R scores was statistically significant, $F(2, 371) = 41.732$, $p < .01$, $R^2 = .37$. The regression weight associated with ORF indicated that for every one-word increase in ORF, OAKS-R increased .09 RIT points. For a positive difference of 39 CWPM in ORF (approximately one standard deviation), the OAKS score is predicted to be about four points higher, or a half of a standard deviation in RIT units. The regression weight associated with Maze indicated that for every one-word increase in Maze, OAKS-R increased by .19 RIT points. For a positive difference of 8 NWC in Maze

(approximately one standard deviation), the OAKS score is predicted to be about two points higher, or a quarter of a standard deviation in RIT units. The Cohort Three squared semi-partial correlations showed ORF and Maze uniquely accounted for 10% and 2% of the variance in OAKS-R, respectively. Table 7 presents the coefficients associated with the linear regression.

Table 7

Linear Regression Analysis Summary: ORF and Maze Predicting OAKS-R Performance

	Cohort 1			Cohort 2			Cohort 3		
	b	SE	β	*b*	SE	β	*b*	SE	β
Constant	213.86*	1.56		216.73*	1.24		215.50*	1.36	
ORF	0.05*	0.01	.24	0.06*	0.01	.35	0.09*	0.01	.45
Maze	0.51*	0.08	.41	0.27*	0.06	.32	0.19*	0.06	.19

*$p < .01$

These results demonstrate that over the three cohorts, the unique variance explained by Maze decreased, while the unique variance explained by ORF increased in terms of accounting for student performance on the eighth-grade OAKS-R. The Maze went from 8% to 2% unique variance explained, while ORF increased from 3% to 10% unique variance explained across the three sequential cohorts.

Research Question 3: Analyzing the Relation between Demographic Characteristics

The third research question was designed to examine the independent and joint relations between demographic characteristics of disability status (SpEd) and socio-economic status (FRL), and ORF and Maze scores, in the prediction of scores obtained on the OAKS-R. The specifics of each demographic characteristic are described in the Methodology section (p. 38). This research question was initially investigated by

analyzing data from each cohort individually, including examination of interactions between demographic characteristics and each CBM measure. Following the individual cohort analyses, the cohorts were then combined for an analysis that included dummy codes to capture potential relations between cohort membership and OAKS-R outcomes.

Cohort One: Analyzing CBM and Demographic Variables

To analyze the predictive relations of the CBM measures, ORF and Maze, and demographic variables, socio-economic status (FRL) and disability status (SpEd), all variables were included in a multiple regression analysis with OAKS-R as the dependent variable. Overall, this model explained 38% of the variance in predicting scores on the OAKS-R and was statistically significant, $F(4, 370) = 55.56$, $p < .05$, $R^2 = .38$. The addition of interaction terms: ORF-FRL, Maze-FRL, ORF-SpEd, Maze-SpEd, did not statistically improve the prediction of performance on the OAKS-R, $p < .05$.

Table 8 shows the results from the multiple regression analysis with ORF, Maze, SpEd, and FRL as the predictors and OAKS-R as the dependent variable. The ORF, Maze, and SpEd predictors were statistically significant, $p < .05$. The demographic variable FRL ($p= .297$) was not statistically significant in this model. The standardized coefficients indicated that Maze (β= .388) was relatively more predictive than ORF (β= .182) of performance on the OAKS-R. The regression weight associated with ORF indicated that for every one-word increase in ORF, OAKS-R increased .04 RIT points. The regression weight associated with Maze indicated that for every one-word increase in Maze, OAKS-R increased by .49 RIT points. The squared semi-partial correlations showed Maze and ORF uniquely accounted for 7% and 2% of the variance in OAKS-R, respectively. Disability status was negatively related to performance on the OAKS-R, $p <$

.05. Students with disabilities scored 3.46 RIT points lower on the OAKS-R, relative to students without disabilities. The squared semi-partial correlation indicated that special education status uniquely accounted for 1% variance in OAKS-R.

Table 8

Cohort One Multiple Regression Model

Model	Std. Error	*b*	Beta	*t*	Squared Semi-Partial
(Constant)	217.31	1.95		111.40*	
ORF	0.04	0.01	.18	2.95*	.02
Maze	0.49	0.07	.39	6.56*	.07
SpEd	-3.46	1.25	-.13	-2.76*	.01
FRL	-0.83	7.95	-.04	-1.05	.00

*$p < .05$

Cohort Two: Analyzing CBM and Demographic Variables

In Cohort Two, a multiple regression analysis examined the predictive relations of ORF, Maze, SpEd and FRL with scores obtained on the OAKS-R. The overall regression explained 42% of the variance in predicting scores on the OAKS-R and was statistically significant, $F(4, 344) = 61.76$, $p < .05$, $R^2 = .42$. The addition of interaction terms, ORF-FRL, Maze-FRL, ORF-SpEd, and Maze-SpEd did not statistically improve the prediction of performance on the OAKS-R ($p < .05$).

Table 9 shows the results from the multiple regression analysis with ORF, Maze, SpEd, and FRL as the predictors and OAKS-R as the dependent variable. The ORF, Maze, and SpEd predictors were significant, $p < .05$. The FRL predictor was not statistically significant ($p = .065$). The regression weight associated with ORF indicated

that for every one-word increase in ORF, OAKS-R increased .05 RIT points. The regression weight associated with Maze indicated that for every one-word increase in Maze, OAKS-R increased by .26 RIT points.

The standardized coefficients indicated that Maze (β= .304) and ORF (β= .296) were similarly predictive of performance on the OAKS-R. The squared semi-partial correlations showed that Maze and ORF uniquely accounted for 4% and 3% of the variance in OAKS-R, respectively. Disability status was negatively related to performance on the OAKS-R; students with disabilities scored 3.02 RIT points lower on the OAKS-R, relative to students without disabilities. The squared semi-partial correlation indicated that special education status uniquely accounted for 1% of the variance in OAKS-R.

Table 9

Cohort Two Multiple Regression Model

	b	Std. Error	Beta	*t*	Squared Semi Partial
(Constant)	219.64	1.54		142.32*	
ORF	0.05	0.01	.30	4.31*	.03
Maze	0.26	0.06	.30	4.58*	.04
SpEd	-3.02	1.10	-.13	-2.75*	.01
FRL	-1.08	0.70	-.07	-1.55	.00

*$p < .05$

Cohort Three: Analyzing ORF, Maze, and Demographic Variables

The same multiple regression analysis model was applied to Cohort Three student scores to evaluate the predictive relations between CBM and demographic variables with performance on the eighth-grade OAKS-R. The overall regression explained 40% of the variance in predicting scores on the OAKS-R and was statistically significant, $F(4, 367) = 60.14$, $p < .05$, $R^2 = .40$. The addition of the following interaction terms: ORF-FRL, Maze-FRL, ORF-SpEd, and Maze-SpEd, did not statistically improve the prediction of performance on the OAKS-R.

Table 10 shows the results of the multiple regression analysis with OAKS-R as the dependent variable and the predictors ORF, Maze, SpEd, and FRL. In this model all the predictors were statistically significant ($p < .05$). The standardized coefficients indicated that ORF ($\beta = .373$) was a relatively better predictor of performance on the OAKS-R than Maze ($\beta = .174$). The regression weight associated with ORF indicated that for every one-word increase in ORF, OAKS-R increased .08 RIT points. The regression weight associated with Maze indicated that for every one-word increase in Maze, OAKS-R increased by .17 RIT points. The squared semi-partial correlations indicated that ORF and Maze uniquely accounted for 6% and 1% of the variance on OAKS-R, respectively. The demographic variables, special education and free or reduced lunch status, were negatively related to performance on the OAKS-R. Students with disabilities scored 3.08 RIT points lower on the OAKS-R, relative to students without disabilities. Students identified as low socio-economic status (FRL) scored 1.96 RIT points lower on the OAKS-R relative to students not identified as low socio-economic status. The Cohort

Three squared semi-partial correlations indicated special education and FRL status each uniquely accounted for 1% of the variance in OAKS-R.

Table 10

Cohort Three Multiple Regression Model

Model	*b*	Std. Error	Beta	*t*	Squared Semi Partial
(Constant)	219.89	1.69		130.01*	
ORF	0.08	0.01	.37	6.01*	.06
Maze	0.17	0.06	.17	2.96*	.01
SpEd	-3.08	1.10	-.13	-2.79*	.01
FRL	-1.96	0.69	-.12	-2.85*	.01

*$p < .05$

Combined Cohort Analysis

In a final analysis, data from the three student cohorts were combined and the relations between predictors in the previous cohort models, ORF, Maze, FRL, and SpEd, with RIT-scores were evaluated within the context of a single, combined analysis. Dummy codes were added to the combined model to capture potential cohort-to-cohort differences. Table 11 displays the results from the multiple regression analysis with OAKS-R as the dependent variable and ORF, Maze, SpEd, FRL, and dummy codes representing Cohort One and Cohort Two as predictors. The overall regression explained 39% of the variance in scores on the OAKS-R and was statistically significant, $F(6, 1088) = 113.71$, $p < .05$, $R^2 = .39$. The ORF, Maze, SpEd, FRL, and Cohort Two predictors were statistically significant ($p < .05$). The squared semi-partials indicated that the performance measures ORF and Maze explained the greatest amount of unique

variance in prediction of performance on the eighth-grade OAKS-R. The squared semi-partial correlations of all significant variables in the model were as follows: Maze (4%), ORF (3%), SpEd (1%), and FRL (0.5%), and Cohort 2 (0.2%). Socio-economic status (FRL), disability status (SpEd) and Cohort Two (Cohort 0708) had negative predictive relationships with performance on the OAKS-R.

The regression weight associated with ORF indicated that for every one-word increase in ORF, OAKS-R increased .06 RIT points. The regression weight associated with Maze indicated that for every one-word increase in Maze, OAKS-R increased by .29 RIT points. Students with disabilities scored 3.38 RIT points lower on the OAKS-R, relative to students without disabilities. Students identified as low socio-economic status (FRL) scored 1.29 RIT points lower on the OAKS-R relative to students not identified as low socio-economic status. Lastly, students in Cohort Two scored 1.07 RIT points lower than students in Cohort Three on the OAKS-R (See Table 11).

Table 11

Combined Cohort Multiple Regression Model

Model	*b*	Std. Error	Beta	*t*	Squared Semi-Partial
(Constant)	219.37	1.03		213.39**	
ORF	0.06	0.01	.28	7.50**	.03
Maze	0.29	0.04	.26	7.95**	.04
SpEd	-3.38	0.67	-.14	-5.06**	.01
FRL	-1.29	0.42	-.07	-3.05*	.01
Cohort One	-0.02	0.56	.00	-0.03	.00
Cohort Two	-1.07	0.52	-.06	-2.05*	.00

*$p < .05$, ** $p < .001$

Summary of Results

The first and second research hypotheses were that results associated with the current student samples would replicate previous findings obtained with different populations of eighth-grade students and with a different statewide assessment. In previous research, conducted in different states, student ORF and Maze outcomes demonstrated moderate to strong concurrent and predictive validity with scores on statewide reading and standardized assessments (Campbell & Long, 2010; Espin et al., 2010; Silberglitt et al., 2006; Ticha et al., 2009). The concurrent validity analysis in the current study corroborated these previous research findings, but with less strong relations noted with the different dependent measure, the OAKS-R. The predictors ORF and Maze had positive, moderate correlations with OAKS-R in the concurrent validity analysis. The second analysis also corroborated previous research findings, in which ORF and Maze

were found to be positive, moderate predictors of scores obtained on the OAKS-R (Megert, 2010; Smith, 2008).

The third research hypothesis was that disability and poverty status would interact with ORF and Maze scores, yielding a relatively stronger prediction of performance for low-income students and for students with disabilities, on the OAKS-R. This study demonstrated that the interaction of disability and socio-economic status (poverty) with ORF and Maze was not statistically significant. However, the variables disability status (SpEd) and socio-economic status (FRL) were found to have a negative, statistically significant predictive relationship with student performance on the eighth-grade OAKS-R, $p < .05$. Analyzing the trend of Maze and ORF across the cohorts, Maze was a stronger predictor than ORF in Cohort One but lost relative predictive strength across the cohorts. The opposite was true for ORF. Initially, ORF was less predictive in Cohort One, but gained relative strength as a predictor across the cohorts, ending as the strongest predictor in Cohort Three.

An extension of the third research hypothesis was that when the cohorts were combined for an analysis, Maze and ORF would emerge as the strongest predictors of performance on the eighth-grade OAKS-R, in this analysis. While both ORF and Maze were the strongest predictors in the model, Maze was the strongest predictor of performance on the OAKS-R. The SpEd and FRL demographic variables, while statistically significant ($p < .05$), were negative predictors of performance on the OAKS-R. In the Combined Cohort analysis, the presence of a disability and low family income, as measured by free and reduced-price lunch status, were negatively associated with performance on the eighth-grade OAKS-R. In the Combined Cohort analysis, and in all

the individual cohort analyses, the prediction strength of demographic variables, SpEd and FRL, were not as strong as the prediction strength of CBM predictors, ORF and Maze. The results of the current study corroborated previous research findings (Campbell & Long, 2010; Espin et al., 2010; Silberglitt et al., 2006; Ticha et al., 2009) in that ORF and Maze were predictors of performance on a statewide reading assessment for eighth-grade students in Oregon.

CHAPTER V

DISCUSSION

In this chapter, I will (a) summarize the analyses presented in the previous chapter, (b) connect the present study with previous research, (c) review limitations of the study, (d) discuss the practical implications of the findings, and (e) provide suggestions for future research.

Summary of Results

This study examined the concurrent and predictive relationships of CBM, ORF and Maze, on eighth-grade student performance on the OAKS-R. With accurate prediction, teachers can use CBMs to identify early in the school year, which students are at risk of failing the summative, high stakes, statewide reading assessment. As high school diploma requirements have become increasingly stringent (ODE, 2007b), districts need to ensure eighth-grade students finish the year with the literacy skills necessary to participate in their future high school's college-and-career ready curriculum. The need for high school literacy readiness is consistent with the Stages of Reading Development: Stage Three (Chall, 1983) which asserts that by the end of eighth grade, students should be able to comprehend what is read and be able to learn new information in a variety of formats such as adult literature, popular fiction, science, social studies next, adult encyclopedias and other reference materials. To ensure students are on track and ready for their high school curriculum, one suggestion is for districts to identify early in the eighth-grade school year, which students are at risk of not passing the OAKS-R. With

this information available earlier in the school year, educators have the opportunity to make instructional adjustments to target student needs prior to the students taking the statewide reading assessment. This concept of systematic assessment and follow-up instructional adjustments based on students' current knowledge is one of the most basic tenets of effective instruction to maximize student learning (Jenkins & Jewell, 2010).

Concurrent Validity of CBM to OAKS-R

The concurrent validity analysis of spring administered CBM and the OAKS-R revealed moderately strong relationships between the CBM and OAKS-R. Scores from two cohorts of eighth-grade students were analyzed, and in both cohorts ORF had the strongest correlation with the OAKS-R ($r = .64$, $r = .58$), whereas the Maze measure was slightly weaker ($r = .58$, $r = .54$). These results indicate that for eighth-grade students, ORF and Maze, were valid, concurrent measures of reading skills when administered within a six-week timeframe. However, the results of this study were not as strong as those reported in Ticha et al. (2009). In the Ticha et al. study, the ORF ($r = .87$, $p < .008$) and Maze ($r = .88$, $p < .008$) performance measures were strongly correlated with scores on the WJ-II Pretest. The outcome measure in the current study was a statewide assessment as compared to the Ticha et al. study, which utilized a common standardized assessment to measure concurrent validity. Additionally, the sample sizes and characteristics were different between the two studies. The Ticha et al. study utilized a small, ethnically diverse sample, whereas the current study sample sizes were much larger and less ethnically diverse.

Interestingly, the results of this study's concurrent analysis are more consistent with Silberglitt et al. (2006), in which both ORF ($r = .60$, $p < .001$) and Maze ($r = .48$, p

< .001) were found to have moderate, positive correlations, with scores on the Minnesota statewide reading assessment, the Minnesota Basic Standards Test- Reading (MBST). However, Silberglitt et al. utilized a much larger sample size, including 843 eighth-grade students for the ORF measure and 1,028 students for the Maze measure. Like the Silberglitt et al. sample, the District in the current study mirrored their lack of ethnically diverse students. This study had approximately 73 percent of its students identified as Caucasian and 13 percent of its students identified as Hispanic. Similarly, the Silberglitt et al. study included primarily Caucasian students (94%) with a variety of student performance ranges. The similarities and differences between this study and the two previous concurrent validity studies with eighth-grade students suggested that Maze and ORF tended to be more strongly related to standardized measures of reading comprehension with more ethnically diverse populations (Espin et al., 2010; Ticha, 2009).

Using Curriculum-Based Measures to Predict Performance on OAKS-R

The second question utilized linear regression analysis to understand the unique contribution of ORF and Maze in the prediction of performance on the OAKS-R, for three cohorts of eighth-grade students. For Cohort One, the squared semi-partial correlations showed that Maze and ORF uniquely accounted for 8 and 3 percent of the variance in student performance on the OAKS-R, respectively. For Cohort Two, the squared semi-partial correlations showed that ORF uniquely accounted for 5 percent while the Maze uniquely accounted for 4 percent of the variance in student performance on the OAKS-R. For Cohort Three, the squared semi-partial correlations showed that ORF uniquely accounted for 10 percent while the Maze uniquely accounted for 2 percent

of the variance in student performance on the OAKS-R. All results were statistically significant, $p < .01$.

A benefit of examining data from multiple samples of students was to gain more information about the stability of the CBM measures as predictors of performance on the OAKS-R. An important finding of this study was that the results across cohorts were not consistent. With respect to the measures, there were three distinct patterns noted across the cohorts including: (a) ORF gained in predictive power every year, while Maze lost predictive power with each successive cohort, (b) the mean score and the minimum score on the Maze increased with each successive cohort, and (c) the variability of OAKS-R went down with each successive cohort. In summary of the multiple regression analyses, Maze was the strongest predictor of performance in Cohort One, while in Cohort Two, ORF and Maze were virtually equal predictors of performance, and in Cohort Three, ORF was the strongest predictor of performance on the eighth-grade OAKS-R. The results are not conclusive as to which CBM was the most predictive of performance on the eighth-grade OAKS-R, as the results fluctuated and ultimately were split across the cohorts. In fact, the results suggest that the use of *both* ORF and Maze is important for screening students.

One possible explanation for the differences noted across the cohorts is different sample sizes and demographic characteristics between the samples. Interestingly, Cohort One and Cohort Three had nearly the opposite results with predictors ORF and Maze, and yet the two cohorts had identical sample sizes and the demographic characteristics of the cohorts were very similar. Conversely, Cohort Two had a smaller sample size ($n = 348$) and was slightly less ethnically diverse than the other two samples, approximately 77

percent Caucasian, compared to approximately 70 percent Caucasian in Cohort One and Cohort Three. In Cohort Two, ORF and Maze similarly predicted student performance on the eighth-grade OAKS-R while Cohort One and Cohort Two produced nearly opposite results with respect to each measures' unique contribution of prediction of performance on the OAKS-R.

Another possible explanation for the differences noted across the cohorts is sampling error. Inherent in using a sample, or samples, rather than testing the whole population, is the risk of not getting a representative sample of the entire population of eighth-grade students. The differences of the individual students who comprised each of the three cohorts of students may have contributed to the varied findings across the cohorts. While there is no clear indication for the different findings across the cohorts, the differences may be ascribed to typical variations found between and within cohorts that school personnel often experience. For example, Cleary and Chen (2009) found differences not only across middle school grade cohorts, but also within cohorts with respect to math achievement. Examples of across cohort differences found in their study included students' use of self-regulatory strategies and the amount of exhibited maladaptive behaviors. An example of a within-cohort difference noted was gender difference; girls exhibited slightly greater self-adaptive skills and use of positive strategies over their male counterparts. In summary, teachers often report differences within and across cohorts of students; it could be these, or other differences across and within cohorts, such as those identified in Cleary and Chen study that accounted for the ORF and Maze prediction strength fluctuations across the three cohorts in the study.

Another possible explanation for the cohort differences may be a shifting instructional focus across the years. Lepper, Corpus, and Iyengar (2005) stated that students "who are particularly focused on the extrinsic consequences [from their teachers] of their behaviors do particularly well on objective indicators of performance" (p. 186). With this possibility, the cohorts may have engaged in extrinsic behaviors that mirrored the teacher's instructional focus, as reflected in the varied prediction strength of ORF and Maze across the cohorts. For example, during the Cohort One school year, perhaps teachers were emphasizing reading comprehension tasks (similar to Maze) rather than fluency, which resulted in the increasing mean of the Maze measure noted across the three years. In the Cohort Two school year, the teachers could have been more balanced in their instructional approach including both comprehension and reading fluency (similar to ORF), resulting in similar prediction strengths for both the ORF and Maze measures. In the Cohort Three school year, a greater instructional focus on improving reading fluency and less focus on comprehension could explain the rise in ORF prediction strength while the Maze measure lost predictive power. Variance in instructional focus such as these could have impacted cohorts differently, resulting in the variance in scores across the cohorts.

The Relationship between Disability and Family Income and OAKS-R Scores

The third research question was designed to investigate the possibility of a differential relationship between the demographic characteristics, disability status (SpEd) and socio-economic status (FRL) and ORF and Maze scores, in the prediction of performance on the eighth-grade OAKS-R as well as to estimate the main effects of all predictors in the model. To answer this question, interaction terms were created (ORF-

SpEd, ORF-FRL, Maze-SpEd, Maze-FRL) and included in a multiple regression analysis with OAKS-R as the dependent variable. In this model, the interaction terms were not statistically significant, so they were dropped from further analysis. However, the demographic variables FRL and SpEd were retained for further analysis by cohort and then in a final, combined cohort analysis.

The results of the current study indicated the demographic variables disability status (SpEd) and socio-economic status (FRL) were found to have a negative, statistically significant predictive relationship with student performance on the eighth-grade OAKS-R. For students identified with a disability, there was a negative predictive relationship with performance on the eighth-grade OAKS-R across all analyses. For low socio-economic status (FRL) students there was a negative predictive relationship with student performance on the OAKS-R, but the results were not consistent across all analyses. Only in Cohort Three and in the Combined Cohort analyses was FRL a statistically significant, negative predictor of student performance on the OAKS-R. The lack of consistency noted with the FRL variable is promising; it suggests that while there were some statistically significant negative associations with poverty, it was not always true that poverty status predicted performance on the eighth-grade OAKS-R. This finding provides evidence that some eighth-grade Oregon students have been able to overcome the potential negative effects of poverty such that it did not always negatively predict performance on the OAKS-R.

Comparisons to Prior Research

In this section, a comparison between the findings of the current study will be contrasted to studies within the literature review along three specific points. First, the

predictive nature of CBM on statewide assessment for eighth-grade students will be discussed. Next, ORF and Maze with middle school students with disabilities will be examined. Finally, family income and reading development for eighth-grade students with respect to the OAKS-R, ORF, and Maze measures will be reviewed.

Eighth-Grade Students

As demonstrated in the literature review, few studies have investigated the predictive nature of CBM on statewide testing for eighth-grade students. However, with respect to the studies that have been conducted, results of the current study partially corroborated previous findings regarding outcomes associated with eighth-grade students from ethnically and academically diverse populations. Silberglitt et al. (2006) warned against the use of ORF and Maze to make predictions about performance on a statewide assessment for eighth-grade students citing a decrease in magnitude of the relationship of ORF and Maze and the MBST for eighth-grade students. This claim was partially supported by the results of the current study. Although the correlation coefficients in the current study were slightly stronger than Silberglitt et al. (2006), they were not as strong as those in Espin et al. (2010) and Ticha et al. (2009). For example, Espin et al. showed ORF and Maze both had the same, strong relationship (r = .78) with the Minnesota statewide reading assessment. Similarly, Ticha et al. (2009) reported strong relationships for ORF (r = .77) and Maze (r = .82) with the Minnesota statewide reading assessment. The current study results indicated the overall regression of ORF and Maze were positive, moderate predictors of student performance on the OAKS-R across cohorts: Cohort One (r = .60), Cohort Two (r = .63), and Cohort Three (r = .61). The current study included larger samples of eighth-grade students that were much less ethnically diverse than those

utilized in previous eighth-grade studies suggesting that ORF and Maze are less predictive when administered to less ethnically diverse populations (Espin et al., 2010; Ticha et al., 2009). The differences between the current and previous studies could also be due to the use of different predictor and criterion measures. The predictor measures used in the current study were AIMSweb measures, while the predictor measures used in the previous studies, both the fluency and maze measures were created from articles selected from the newspaper (Espin et al., 2010 and the Ticha et al., 2009). The criterion measure in the current study was the Oregon statewide reading assessment, while the criterion measure used in the previous studies was the Minnesota statewide reading test. There continues to be a lack of research that focuses on *eighth-grade* students with respect to utilizing CBM to predict performance on statewide reading assessments.

Students with Disabilities

Two studies have provided evidence regarding the use of ORF and Maze with middle school students with disabilities (Fore III et al., 2009; Fuchs et al., 1988). Fuchs et al. reported that the capacity to read passages and answer questions about the passages on the Reading Comprehension subtest of the Standard Achievement Test was strongly associated with the ORF (r = .91). Fuchs et al. included a small sample of students (n = 70) with disabilities from Grade 4 to Grade 8. The study included students identified with a reading disability (n = 50), emotionally disturbed (n = 16), and intellectually disabled (n = 4). Of these students, 69% were Caucasian and 31% were identified as minority. The study findings corroborated the concept that improvement on oral reading performance over time is consistent with improvement noted on global measures of reading achievement (Fuchs et al., 1984).

Fore III et al. (2009) reported less impressive results with ORF ($r = .45$, $p < .05$), than with Maze ($r = .89$, $p < .05$) when correlated with the Passage Comprehension subtest of the Woodcock Johnson III Tests of Achievement (Woodcock, McGrew & Mather, 2001) with an ethnically diverse sample of middle school students identified with emotional and behavioral disorders. The Fore III et al. study included 45% Caucasian students ($n = 25$), 38% African American students ($n = 21$), and no ethnicity information was available for the remaining 17% of the students in their sample ($n = 9$). The study results indicated that for a small, ethnically diverse sample of students with emotional and behavioral disorders, the Maze was a stronger concurrent measure of reading comprehension than ORF. Generalizing only to a similarly diverse population of students with emotional and behavioral disorders, the results of this study suggested the Maze as the better CBM to gain quick information about eighth-grade students' reading comprehension skills.

The current study of eighth-grade students included all students with disabilities in one category (SpEd). In all multiple regression analyses (e.g., individual cohort analyses and the combined analysis), the presence of a disability was negatively related to performance on the OAKS-R. Students with disabilities in the current study scored lower on the eighth-grade OAKS-R than students without disabilities. Inherent in the identification of a disability, under the IDEA'04 guidelines, is the notion that the presence of a disability affects educational progress (IDEA, 2004). The current study provided further evidence that special education identification is, in and of itself, a negative predictor of performance on the OAKS-R. The Fore III et al. (2009) and Fuchs et al. (1988) studies provided evidence that CBMs, Maze and ORF, can be utilized as

quick measures of reading comprehension for use with middle school students with disabilities. For these students, the measures may be used more frequently to monitor reading progress across the school year to both, inform instruction and improve predictions of eighth-grade student performance on the OAKS-R.

Family Income and Reading Performance

Although the relationship between family income and reading development for eighth-grade students has not been directly studied with respect to the OAKS-R, ORF, and Maze measures, the 2003-09 NAEP (U.S. Department of Education, 2009) results revealed a national trend indicating a negative relationship between family income and eighth-grade student reading performance. The results of this study do not conclusively corroborate these findings. The current study found that low socio-economic status, as defined by students receiving free or reduced lunch (FRL), was statistically significant only in the Cohort Three model analysis (b = -1.96) and the Combined Cohort model analysis (b = -1.07). The semi-partial correlations showed that FRL explained 1 percent (Cohort Three) and 0.5 percent (Combined Cohort) of the unique variance in the models. In the two analyses, low socio-economic status had a statistically significant, negative, weak predictive relationship with performance on the OAKS-R. The current study results partially corroborated a previous study conducted with a *seventh-grade* sample (n = 422), in which Smith (2008) found that for low socio-economic status students, ORF was a statistically significant predictor of performance on the seventh-grade Oregon Statewide Assessment of reading.

The results of the current study are somewhat encouraging as they indicated poverty had a relatively small negative relationship with eighth-grade performance on the

OAKS-R. In the current study, it is possible that the potential negative relationship between poverty and reading fluency was mediated by the instructional policy and practice of the District. Because of the District's increased focus on literacy instruction in grades six through eight, specifically in the area of reading comprehension, this finding may bode well for identifying a mechanism for ameliorating certain factors, like poverty on student reading outcomes. In the middle school grades, a time when many districts discontinue reading instruction and move towards reading to learn, the District in the study continued to provide reading instruction and interventions to students in all middle school grades. The instructional focus was on improving reading skills in order to achieve independent reading comprehension across the disciplines. This practice may explain the weaker relationship between poverty and reading performance noted in the current study as compared to the national trend noted in the 2003-09 NAEP (U.S. Department of Education, 2009) results.

Limitations

The three limitations in this study were largely associated with (a) the sampled population (convenience sampling and/or external validity), (b) the lack of consideration of classroom, teacher and instructional factors, and (c) reliability and validity of the measures. These factors will be discussed in the next three sections.

Convenience Sample

The current study utilized a convenience sample of eighth-grade students enrolled in one school district. Inherent in convenience sampling are limitations associated with the external validity of the results. Using only one grade level provided information that is grade level specific, thereby prohibiting the generalizability of findings to additional

grade levels. Additionally, the results of this study can only be generalized to other demographically similar populations of eighth-grade students. In the current study, the students were mostly Caucasian and the samples cannot be considered representative of the greater population of eighth-grade students in the United States.

Teacher, Curriculum and Instructional Characteristics

This study utilized extant data and does not account for district, school, and classroom differences in terms of instructional and curriculum practices, and teacher-to-teacher differences in experience, professional development, and overall skill. Teacher and school characteristic variables were not measured. It is likely that following the fall administration of the CBMs students were provided instruction to improve the specific skills measured by the CBM and, thus, reduced the predictive nature of the CBM. For example, instructional emphasis within a school on comprehension over reading fluency may have been a contributing factor to the increase in Maze scores across the three years. Focus on vocabulary (e.g., important words, words that carry meaning) and contextual cues may have improved performance on the Maze because it emulates the skills necessary to complete the Maze rather than orally reading connected text.

Similarly, the poor economy in recent years has created teacher movement between grade levels because of layoffs. If that movement created either a strong teacher grouping or a weaker teacher grouping, then teacher variability could come into play. Thus, the results of this study may not generalize as expected to demographically similar populations of students, given the potential impact of the classroom variables of instruction, curriculum, and teacher variability.

Reliability and Validity of the Maze

This study utilized AIMSweb ORF and Maze measures. Both AIMSweb measures have fully met the seven standards derived by the Standards for Educational and Psychological Testing (Shinn & Shinn, 2002b). The AIMSweb eighth-grade ORF passages have established high reliability and extensive criterion related validity with reported strong correlations (see Methodology chapter, p. 40). Although the Maze passages, have met the aforementioned seven criteria (according to AIMSweb), the reliability and validity studies associated with the Maze were conducted at lower grades (grades 4-6). Also, AIMSweb included no studies that sampled the general population of eighth-grade students, which was the focus of this dissertation. Thus, this lack of matched grade level reliability and validity studies of the AIMSweb Maze measure to the dissertation sample must be viewed as a limitation of this study. Nonetheless, the validity data provided in this study contributes to the current knowledge base about the use of the AIMSweb Maze measure with eighth-grade students. Furthermore, though not a research goal of this study, the findings provide positive empirical evidence that the AIMSweb Maze can be used as a viable formative assessment at the eighth-grade level to predict large-scale assessment outcomes.

Practical Implications

Data-based Instructional Decision Making

The findings from this study and previous studies (Campbell and Long, 2010; Espin et al., 2010; Silberglitt et al., 2006; Ticha et al., 2009) have provided evidence to support the use of ORF and Maze to predict early in the school year, eighth-grade student performance on the spring administered statewide reading assessment. With this

information, district and school teams can create systems to collect and analyze data (i.e., universal screening) to guide instructional decisions and improve student reading performance. Districts can create data systems (i.e., benchmarks) using ORF or Maze at different points in the eighth grade (e.g., fall, winter, spring), creating the opportunity for schools to respond when a student does not meet the expected reading level, by adjusting instruction in an effort to improve the student's reading skills (Fuchs & Fuchs, 1992). With developed systems in place, districts can track progress and provide students the necessary, explicit literacy interventions to improve reading skills and ultimately student performance on high-stakes assessments such as the OAKS-R.

District and School Resource Allocation

Districts provide schools a level of staffing and resources. Building administrators have flexibility to adjust staffing within their building to meet the unique literacy needs of their students. Building administrators may adjust the following components to affect student outcomes, based on their identified priorities and needs, (a) building schedules, (b) staffing, and (c) curriculum choice. Within a given district, schools may vary greatly with respect to demographic characteristics, such as students identified with a disability and/or low socio-economic status. This study provided evidence that these characteristics may be negatively related to reading performance. Building administrators may match student needs and resources by adjusting and allocating building resources to meet these needs. Because of the lack of empirical literature on the topic, principals do not have a blueprint to follow for allocating resources. However, building administrators can create priorities for their school, based on the unique student needs of their population. For example, improving reading skills may be a school priority. They may train all staff in

their building to administer to the CBM and then also train all staff to implement reading interventions for students identified as needing additional instruction. As another option, within their building schedule, the administrator may create a reading period, a time when all students receive reading instruction. By creating building level priorities and leveraging resources, building administrators can meet the unique literacy needs of their students in the area of measurement and instruction.

Costs Versus Benefits of Administration and Scoring

Based on the results of this study, districts could choose to use either ORF or Maze to predict performance on the eighth-grade OAKS-R. Prior to making a decision, districts will want to weigh the costs and benefits associated with each assessment.

Costs and benefits associated with oral reading fluency. A cost associated with using the ORF measure with eighth-grade students is its face validity (Fuchs et al., 1988). Although research clearly indicates that ORF is not only a measure of fluency, but also a measure of comprehension, teachers may not be convinced that a fluency measure is a proxy for reading comprehension when used with eighth-grade students (Chall, 1983; Jager-Adams, 1990; LaBerge & Samuels, 1974; Pikulski & Chard, 2005; Samuels & Flor, 1997; Shinn & Good, 1992). The ORF is a measure that districts commonly use with younger students, prior to middle school, but its use with eighth-grade students is less common, despite a small, but growing research literature that validates its use at the middle school level (Espin et al. 2010; Fore III et al. 2009; Megert, 2010; Ticha et al. 2009; Silberglitt et al. 2006; Smith, 2008).

Additional costs of ORF to be considered include administration and scoring as well as minimal student instructional time loss to consider. To administer and score the

ORF, temporary increased staffing is necessary to coordinate the assessment schedule and then to individually administer and score the assessment. The scoring of the ORF can be very efficient, such that the assessment is scored while the assessor is waiting for the next student to arrive for their assessment. With such efficiency, any costs associated with scoring of the ORF are minimal. The 1:1 assessment format results in minimal loss of student instructional time as they are pulled out of class, one-by-one, to complete the assessment, approximately 3- 5 minutes per student. Thus, both additional staffing costs and instructional time loss should be considered.

A benefit of using ORF is the option to determine the *percent accuracy of total words read*. For example, if a student read 150 total, but made 25 errors, the student's reading was 83 percent accurate. When student's reading is less accurate, their reading comprehension is compromised (Archer, et al. 2003; Smith, 2008). Research has shown that including an additional measure of accuracy to the ORF score can improve the prediction of student performance on the seventh-grade OAKS-R (Smith, 2008). Future studies could evaluate the effect of *accuracy* on the prediction of performance on the eighth-grade OAKS-R. With more evidence about the utility of accuracy at different grade levels, districts can consider whether ORF or Maze best meets their needs for measuring student progress and predicting performance on the OAKS-R.

Costs and benefits associated with Maze. A benefit to considering the Maze measure is it has face validity. The Maze is a task that appears age appropriate and complex, measuring the higher-level construct of reading comprehension. The Maze is group administered; it takes three minutes to complete the task, no more than 10 minutes from setup to completion. There is little instructional time loss and no additional staffing

required to administer the Maze. The costs of the Maze are related to scoring. There are three scoring options available to districts. The first option is districts can hand score student responses, which is staff intensive. The second option is they can use an electronic scan system to score student responses, which requires fewer staff but the scan system has associated costs. These costs include buying the scanning machine, staffing costs to create the scanning sheets, and then staffing costs to feed the scanning sheets into the machine. The final option is for districts to use a computer-administered Maze test. This option may, or may not, have an added technology cost of paying for computers- either classroom sets of computers or school computer labs. The additional cost depends on what the district already has for computer resources.

Which Measure Should Districts Use: ORF or Maze?

Depending on the costs and benefits of each measure and how they will utilize the data, districts will need to decide which measure, ORF or Maze, they will use with their eighth-grade students. For example, the District in the study discontinued the use of Maze and continued the administration of ORF with eighth-grade students, utilizing the additional measure of *accuracy* to further aid in identification of students that need reading intervention in their eighth-grade year. Alternatively, districts may choose the Maze if they are concerned about face validity. This choice would result in a small amount of instructional time loss, an overall decrease in staffing costs needed to administer and score the assessment, but a potential cash outlay for a Maze scoring system. Or, they might consider if a computer-based Maze assessment could be accommodated in the district with existing resources. A clear recommendation of *which* CBM, ORF or Maze, to use with eighth-grade students cannot be provided, given the

nature of the results across the three cohorts. In fact, based on the results of this study, districts should use both measures, ORF *and* Maze, to predict eighth-grade student performance on the OAKS-R. Using both measures will provide a more stable prediction of performance accounting for cohort variance that naturally occurs from year to year. The rationalization for using both is that independently the unique variance explained by each measure fluctuated across the cohorts. However, the combination of ORF and Maze together consistently explained more variance than separately, suggesting that use of both is important and necessary for screening students. In summary, districts will want to make a decision to use the measures in a way that meets the needs of their district and setting, while providing the most explained variance in predicting scores on the OAKS-R. The use of both measures will provide the greatest prediction of student performance on the eighth-grade OAKS-R.

Future Research

Analyze Sixth and Seventh Grades

The importance of ensuring students can read has been made clear by the previous research and publications reviewed during the course of this study (Biancarosa & Snow, 2006; Carnegie Council on Advancing Adolescent Literacy, 2010, Deschler & Hock, 2007; ODE, 2009; Strickland & Alvermann, 2004). Extending research to the earlier middle school grades, sixth grade and seventh grade, could continue to inform the profession about the effectiveness of using CBM in the middle grades to predict performance on statewide reading assessments. Studies have largely focused on earlier, primary grade level identification of reading difficulties as it provides greater opportunity for intervention and instruction to improve reading skills and comprehension and

ultimately to be successful in high school and beyond (Chard et al., 2008, Crawford, Tindal, & Steiber, 2001; Fuchs, Fuchs, & Maxwell, 1988; Marcotte & Hintze, 2009, Megert, 2010; Shinn & Good, 1992; Smith, 2008; Wanzek et al., 2010; Wiley & Deno, 2005; Wood, 2006). Continuing to inform the field about the efficacy of using ORF and Maze to predict performance on statewide reading measures in middle school offers educators timely and meaningful information. With this information, they have the opportunity to intervene with instruction and improve student performance on the statewide assessment of reading.

Fidelity of Administration

Given the simplicity of the Maze administration, the District considered building level oversight for fidelity of Maze administration to be adequate. In hindsight, the District should have conducted fidelity studies to corroborate this assumption. Because of the lack of fidelity checks for the Maze and the OAKS-R, internal validity may have been compromised if standardized administration procedures were not followed. During future research studies, fidelity checks should be conducted to ensure the standardized format is being followed. Reliability checks during every assessment period in a universal screening scenario (three times a year) would be ideal. Following standardized administration directions ensures that assessment scores are not compromised as a result of the administration procedures. Future studies should include reliability checks of all performance measures, including the Maze.

Classroom, Instruction, and Teacher Variance

Without examining relationships with instructional, curriculum, and teacher variables, the present study could not account for the variance associated with those three

factors. In effect, if students who scored lower on the fall CBM were then provided instruction, it may have improved the specific skills measured by the CBM and thus reduced the predictive nature of the CBM. Future studies could increase instructional consistency within and across middle schools by utilizing the same curriculum and conducting fidelity checks for consistent implementation of curriculum across classroom and school settings. For example, the use of principal walk-throughs, paying specific attention to the curriculum and instructional strategies being used in the classroom may provide additional consistency around the curriculum and instructional variables.

Conclusion

As high school diploma requirements have become increasingly stringent (ODE, 2007b) districts need to ensure eighth-grade students finish the year with the literacy skills necessary to participate in their future high school's curriculum. The results from the current study provide a promising approach to predicting eighth-grade performance on the OAKS-R. By utilizing CBM, ORF and/or Maze, districts may implement a universal screening procedure in the fall to identify students who may be at risk for not passing the OAKS-R. Following the screening, the students should receive appropriate instruction to improve their reading comprehension skills, and their performance on the eighth-grade OAKS-R.

The results from this study indicate that both ORF and Maze were valid, concurrent measures of reading comprehension with the OAKS-R for three eighth-grade cohorts. It is important to remember that the OAKS-R is primarily a reading comprehension assessment. The ORF is a more distal measure of reading comprehension. It is thought of as distal because it is not actually a comprehension measure, but measures

an automaticity factor necessary in the reading comprehension process. Conversely, the Maze task is a proximal measure of reading comprehension. The Maze assessment, a task in which selecting the correct word in the context of a passage, emulates reading comprehension (Parker, Hasbrouck, & Tindal, 1992). The Maze has face validity, as the task associated with the assessment appears to measure reading comprehension, the same construct measured in the OAKS-R.

In this study, ORF and Maze predicted performance on the eighth-grade OAKS-R. When demographic variables, disability status (SpEd) and family income (FRL) were added to the prediction model, disability status consistently had a negative relationship with performance on the OAKS-R. Family income also had a negative relationship with performance on the OAKS-R, but did not statistically improve the prediction across all cohort analyses. In all the analyses, ORF and Maze emerged as the strongest predictors of performance on the eighth-grade OAKS-R. These results support the use of ORF and/or Maze to predict performance on the eighth-grade OAKS-R. To maximize eighth-grade student outcomes, students can be assessed with either ORF or Maze, early in the fall of eighth grade to determine whether they are on track for passing the eighth-grade OAKS-R benchmark. With fall assessments that provide identification of students not on track for passing the OAKS-R, schools can take proactive steps by providing students reading interventions to students that focus on literacy and comprehension strategies. In this way, schools and districts can address the adolescent literacy crisis (Jacobs, 2008) and improve the readiness of their eighth-grade students for the high school curriculum.

APPENDIX A

SAMPLE ORAL READING FLUENCY PASSAGE: GRADE 8

From the birch-bark canoe, Joseph Anthony scanned the banks of the Serpentine River. His voyageurs paddled hard and fast through the water, so he had to make his observations quickly. Finding a location for his wintering post was an important decision, one that could mean success or failure for his fur-trading mission.

Prior to his mission, Anthony met and stayed with Chippewa leaders for two days. The native leaders assured him the region was excellent.

"The land is rich with game and food," they had said. "The furs are also very plentiful."

Judging by the looks of the area, Anthony believed he had been told the truth. He spied a steep bank through the towering white pines. He commanded his voyageurs, "Put the canoe in there, by that big rock."

Anthony scrambled out of the canoe and climbed a tall, rocky ridge. He eagerly surveyed the land and river from the vantage point. Once the trees were cleared, the men would have a good view of the river and an excellent location for fishing. (The passage continues for 370 words.)

APPENDIX B

ADMINISTRATION DIRECTIONS: ORAL READING FLUENCY

Directions for a 1-minute administration of ORF.

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: ORAL READING FLUENCY

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 productions are correct.

Scored as incorrect:

- 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.

APPENDIX D

SAMPLE MAZE PASSAGE: GRADE 8

Mr. Mooney is an expert at his occupation. In fact, he is one of (has, and, the) few experts left in an occupation (this, that, right) is slowly dwindling and lacking well-(finished, trained, can) professionals.

Mr. Mooney shampoos animals. He (were, finished, is) known to boast that there is (in, one, no) animal too filthy, too big or (too, is, all) ferocious (to, an, for) him. He can get them all (polished, top, clean) and contrite. Mr. Mooney has the (right, bristly, week) to boast.

Once, when Mr. Mooney (top, is, was) younger, he was called upon to (shampoo, polished, beauty) the walruses at the local zoo. (That, When, If) Mr. Mooney was finished with those (heads, walruses, next), their bristly coats were gleaming, their (remained, roof, tusks) were polished to perfection, and the (next, gleaming, few) hairs they possessed on the top (of, the, for) their heads were fit for a (assistant, younger, beauty) pageant.

(The passage continues for a total of 54 maze choice selections.)

APPENDIX E

ADMINISTRATION DIRECTIONS: FOR OLDER STUDENTS

1. After the students have put their name on the cover sheet, start the test by saying… "*When I say 'Begin' turn to the first story and start reading silently. When you come to a group of three words, circle the one word that makes the most sense. Work as quickly as you can without making mistakes. If you finish a/ the page/first side, turn the page and keep working until I say 'Stop' or you are all done. Do you have any questions?"*
2. Then say, "*Begin."* Start your stop watch.
3. Monitor students to make sure they understand that they are to circle only one word.
4. If a student finishes before the time limit, collect the student's maze task and record the time on the student's test booklet.
5. At the end of the 3 minutes say: "*Stop. Put your pencils down. Please close your booklet.*"
6. Collect the maze tasks.

APPENDIX F

MAZE SCORING RULES

Immediate scoring following completion of the Maze is recommended. The most important task is to determine the Number of Words (Items) Correct. The number of errors are important, but less so. Determining Words Correct is easy. Using the answer key put a slash through incorrect words.

What is correct?

An answer is correct if the student circles the word that matches the correct word on the scoring template.

What is Incorrect?

An answer is considered an error if the student:

a. Circles an incorrect word.
b. Omits word selections other than those the student was unable to complete before the 3 minutes expired.

Making Scoring Efficient

1. Count the total number of items up to the last circled word,
2. Compare the student answers to the correct answers on the scoring template. Mark a slash (/) through incorrect responses.
3. Subtract the number of incorrect answers from the total number of items attempted.
4. Record the total number of correct answers on the cover sheet followed by the total number of errors (e.g., 35/2, 45/0).

Prorating

Some students may finish all the items before the 3 minutes is up. To be able to make the most accurate judgment about their progress, the student's score can be prorated to what they would have scored if there were enough items for 3 minutes of student reading. To prorate:

1. When the student finished must be recorded and the number correct is counted. For example, the student may have finished in 2 minutes and correctly answered 40 items.
2. Covert the time taken to seconds. 2 minutes = 120 seconds
3. Divide the number of seconds by the number correct. 120/40 = 3
4. Calculate the number of seconds in the full 3 minutes. 3 minutes = 180 seconds
5. Divide the number of full seconds by the calculated value from step 3. 180/3 = 60

REFERENCES CITED

Achieve. (2007). *ADP benchmarks: End of high school*. Retrieved from http://www.achieve.org/ADPBenchmarks.

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

Adlof, S., Catts, H., & Little, T. (2006). Should the simple view of reading include a fluency component? *Reading and Writing, 19*, 933-958.

American College Test. (2008). *The forgotten middle: Ensuring that all students are on target for college and career readiness before high school*. Iowa City, IA: Author.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Eduction. (1985). *Standards for educational and psychological testing.* Washington, DC: Author.

American Educational Research Association, American Psychological Associaiton, & National Council on Measurement in Eduction. (1999). *Standards for educational and pscyhological testing.* Washington, DC: Author.

Amos, J. (2008). *Dropouts, diplomas and dollars: U.S. high schools and the nation's economy.* Washington, DC: Alliance for Excellent Education.

Biancarosa, G., & Snow, C. E. (2006). *Reading next- A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York* (2nd ed.). Washington, DC: Alliance for Excellent Education.

Carnegie Council on Advancing Adolescent Literacy. (2010). *Time to act: An agenda for advancing adolescent literacy for college and career success.* New York, NY: Carnegie Corporation of New York.

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

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

Chard, D. J., Stoolmiller, M., Harn, B. A., Wanzek, J., Vaughn, S., Linan-Thompson, S., & Kame'enui, E. J. (2008). Predicting reading success in a multilevel schoolwide reading model: A retrospective analysis. *Journal of Learning Disabilities, 41*, 174-188. doi: 10.1177/0022219407313588

Cleary, T. J., & Chen, P. P. (2009). Self-regulation, motivation, and math achievement in middle school: Variations across grade level and math context. *Journal of School Psychology*, *47*, 291-314.

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

Deshler, D. D., & Hock, M. F. (Eds.). (2007). *Adolescent literacy: Where we are- Where we need to go*. New York, NY: Guilford.

Duncan, A. (2009). Turning around the bottom five percent. Retrieved from http://www.ed.gov/news/speeches/2009/06/06222009.html

Durkin, D., & Allington, D. (2004). *Teaching them to read*. Boston, MA: Pearson/Allyn & Bacon.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9*, 167-188.

Espin, C., Wallace, T., Lembke, E., Campbell, H., & Long, J. D. (2010). Creating a progress monitoring system in reading for middle-school students: Tracking progress toward meeting high-stakes standards. *Learning Disabilities Research, 25*(2), 60-75.

Fore III, C., Boon, R. T., Burke, M. D., & Martin, C. (2009). Validating curriculum-based measurement for students with emotional and behavioral disorders in middle school. *Assessment for Effective Intervention, 34*(2), 67-73.

Fuchs, L. S., & Fuchs, D. (2004). Determining Adequate Yearly Progress from kindergarten through grade 6 with curriculum-based measurement. *Assessment for Effective Intervention, 29*(4), 25-37. doi: 10.1177/073724770402900405

Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal measures of reading comprehension. *Remedial and Special Education, 9*(2), 20-28.

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*, 239-256.

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.

Guensburg, C. (2006). Why Johnny (still) can't read. Retrieved from www.edutopia.org/adolescent-literacy-programs

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

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

Hoover, W., & Gough, P. (1990). The simple view of reading. *Reading and Writing, 2*, 127-160.

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. & 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: Shinn & Shinn.

Individuals with Disabilities Education Improvement Act of 2004, 20 U.S.C. 1400 et seq. C.F.R. (2004).

Institute of Education Sciences. (2009). *Special analysis 2009: International assessments* Retrieved from http://nces.ed.gov/programs/coe/2009/ analysis/index.asp

Jacobs, V. (2008). Adolescent literacy: Putting the crisis in context. *Harvard Educational Review*, *78*, 7-39.

Jenkins, J. R., & Jewell, M. (1993). Examining the validity of two measures for formative teaching: Reading Aloud and Maze. *Exceptional Children, 59,* 421-432.

Kame'enui, E. J., & Simmons, D. C. (2001). Introduction to this special issue: The DNA of reading fluency. *Scientific Studies of Reading, 5*, 203-210.

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

Kamil, M., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., & Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices. IES Practice Guide (NCEE 2008-4027).* Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Science, U.S. Department of Education.

Karlsen, B., Madden, R., & Gardner, E. F. (1975). *Stanford Diagnostic Reading Test.* San Antonio, TX: Psychological Corporation.

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

Lee, C.D., Sprateley, A. (2010). *Reading in the disciplines: The challenges of adolescent literacy*. New York, NY: Carnegie Corportation of New York.

Lee, J., Grigg, W., & Donahue, P. (2007). *The nation's report card: Reading 2007. (NCES 2007-496).* National Center for Education Statistics, Institute of Education Sciences. Washington, DC: U.S. Department of Education.

Lepper, M. R., Corpus, J. H., Iyengar, S. S. (2005). Instrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates. *Journal of Educational Pysychology*, *97,* 184-196.

Marcotte, A., & Hintze, J. (2009). Incremental and predictive utility of formative assessment methods of reading comprehension. *Journal of School Psychology, 47*, 315-335.

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

McCall, M. S. (2005). *Technical review of the Oregon Department of Education assessment system.* Pottsboro, TX: Northwest Evaluation Association.

Megert, B. (2010). *Establishing predictive validity for oral passage reading fluency and vocabulary curriculum based measures (CBMS) for sixth grade students*. Dissertation. University of Oregon. Retrieved from http://hdl.handle.net/1794/10872

No Child Left Behind Act of 2001 § 115 Stat.1425 (2001).

National Center for Response to Intervention. (2009). *Progress monitoring tools chart.* Retrieved from http://www.rti4success.org/chart/progressMonitoring/progressmonitoringtoolschart.htm

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.* Washington, DC: Author.

Oregon Department of Education. (2001). *Technical digest- Validity- Oregon statewide assessment program.* Salem, OR: Author.

Oregon Department of Education. (2007a). *2006-2007 technical report: Oregon's statwide assessment system reliability and validity*. Salem, OR: Author.

Oregon Department of Education. (2007b). *Diploma requirements: The Oregon diploma.* Retrieved from http://www.ode.state.or.us/search/page/ ?id=2861

Oregon Department of Education. (2009). *Oregon K-12 literacy framework.* Retrieved from http://www.ode.state.or.us/search/page/?id=2568

Oregon Department of Education. (2010). *The Oregon modified diploma: Frequently asked questions (FAQ)*. Retrieved from www.ode.state.or.us.

Office of Elementary and Secondary Education. (2002). *Guidance for the Reading First program.* Washington, DC: U.S. Department of Education.

Parker, R., Hasbrouck, J., & Tindal, G. (1992). The Maze as a classroom-based reading measure: Construction methods, reliability and validity. *Journal of Special Education, 26*, 195-218.

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

Samuels, S. J., & Flor, R. F. (1997). The importance of automaticity for developing expertise in reading. *Reading and Writing Quarterly, 13*, 107-121.

Shinn, M. R., & Bamonto, S. (1998). Advanced applications of curriculum based measurement: "Big ideas" and avoiding confusion. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement*. New York, NY: Guilford.

Shinn, M. R., & Good, R. H. (1992). Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review, 21*, 138-154.

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

Shinn, M.R. & Shinn, M.M. (2002b). *Reading maze for use in general outcome measurement.* Eden Prairie: MN: Edformation.

Silberglitt, B., Burns, M., Madyun, N., & Lail, K. (2006). Relationship of reading fluency assessment data with state accountability test scores: A longitudinal comparison of grade levels. *Psychology in the Schools, 43*, 527-535.

Silberglitt, B., & Hintze, J. (2005). Formative assessment using CBM-R cut scores to track progress toward success on state-mandated achievement tests: A comparison of methods. *Journal of Psychoeducational Assessment, 23*, 304-325. doi: 10.1177/073428290502300402

Smith, L. R. (2008). *Seventh grade students reading aloud: An examination of the relationship between rate and accuracy in oral reading and reading comprehension.* Doctorate, University of Oregon, Eugene. Retrieved from http://hdl.handle.net/1794/8417

Spencer, M. B. (1999). Social and cultural influences on school adjustment: The application of an identity-focused cultural ecological perspective. *Educational Psychologist, 34*(1), 43-57.

Strickland, D., & Alvermann, D. (Eds.). (2004). *Bridging the literacy achievement gap grades 4-12*. New York, NY: Teachers College Press.

Ticha, Espin, C. A., & Wayman. (2009). Reading progress monitoring for secondary-school students: Reliability, validity and sensitivity to growth of Reading-Aloud and Maze-selection measures. *Learning Disabilities Research, 24*, 132-142.

U.S. Department of Education. (2008). *A nation accountable: 25 years after a Nation at Risk*. Washington, DC: Author.

U.S. Department of Education, National Center for Education Statistics. (2009). *Digest of Education Statistics: 2008.* Retrieved from http://nces.ed.gov/programs/digest/d08/

Wanzek, J., Roberts, G., Linan-Thompson, S., Vaughan, S., Woodruff, T. L. & Murray, C. S. (2010). Differences in the relationship of oral reading fluency and high-stakes measures of reading comprehension. *Assessment for Effective Intervention, 35*, 67-77.

Wiley, H. I., & Deno, S. L. (2005). Oral Reading and Maze measures as predictors of success for English learners on state standards assessment. *Remedial and Special Education, 26*, 207-214.

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

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

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, 24*(3), 4-12.