PREDICTIVE VALIDITY OF READING AND MATHEMATICS CURRICULUM-BASED MEASURES ON MATHEMATICS PERFORMANCE AT THIRD GRADE

by

LINDA MARIE O'SHEA

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 2014

DISSERTATION APPROVAL PAGE

Student: Linda Marie O'Shea

Title: Predictive Validity of Reading and Mathematics Curriculum-Based Measures on Mathematics Performance at Third Grade

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:

Gerald Tindal, Ph.D.	Chairperson
Julie Alonzo, Ph.D.	Core Member
Ben Clarke, Ph.D.	Core Member
Roland Good, Ph.D.	Institutional Representative

and

Kimberly Andrews Espy	Vice President for Research and Innovation; Dean of the Graduate School

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

Degree awarded June 2014

DISSERTATION ABSTRACT

Linda Marie O’Shea

Doctor of Education

Department of Educational Methodology, Policy, and Leadership

June 2014

Title: Predictive Validity of Reading and Mathematics Curriculum-Based Measures on Mathematics Performance at Third Grade

In the current era of high stakes testing, educators use curriculum-based measures (CBMs) and large-scale benchmark assessments to inform instruction and monitor student performance. The Elementary and Secondary Education Act, The No Child Left Behind Act, and Race to the Top all require annual testing in grades 3 through 8 in mathematics and reading. Therefore, educators need appropriate assessments to make valid inferences about instruction and students’ current level of performance as well as risk. Consequently, construct validity is essential for both CBMs and large-scale tests to ensure they appropriately identify students’ current level of performance in reading and math, particularly in making inferences about proficiency (Adequate Yearly Progress). This study of third grade students explored the construct validity of a state math test by correlating it with both math and reading CBMs and determining the sensitivity and specificity of the CBM in predicting performance on the state test.

Results indicated a positive correlation and predictive relation between both CBM math and reading with the Oregon statewide benchmark assessment in mathematics at third grade. Regression analysis showed the strength of the predictive relation of CBM in the identification of students' current level of performance increased with the addition of CBM reading to the CBM math.

A Receiver Operating Characteristic (ROC) analysis indicated that CBM math and CBM reading (passage reading fluency and vocabulary) consistently predicted students who were on target to meet grade-level benchmarks on the statewide assessment. The study adds to the construct validity research on math and reading CBMs. The results may inform assessment development and accommodations needed to assess math content without the reading construct interfering with the interpretation of the results. In addition, it may be useful for educators seeking to identify students who are "at risk" for making grade level progress in mathematics.

CURRICULUM VITAE

NAME OF AUTHOR: Linda Marie O'Shea

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
Portland State University, Portland, Oregon
University of Puget Sound, Tacoma, Washington
University of Washington, Seattle

DEGREES AWARDED:

Doctor of Education, 2014, University of Oregon
Master of Education, 2005, Portland State University
Bachelor of Science, Social Science, 2003, Portland State University

AREAS OF SPECIAL INTEREST:

Science Technology, Engineering, and Mathematics (STEM) Education
Mathematic Intervention and Discourse
Educational Equity
Collaborative Learning for Educators and Students

PROFESSIONAL EXPERIENCE:

Assistant Principal, Arts and Technology Academy/Family School, 4J Eugene School District, 2013 - present

Math teacher, Sherwood Middle School, Sherwood School District, 2009-2013

Elementary Math Coach, Sherwood School District, 2008-2009

Fifth Grade Teacher, Hopkins Elementary School, Sherwood School District, 2005-2009

ACKNOWLEDGMENTS

I wish to express sincere appreciation to the Professors at the University of Oregon for broadening my perspective and deepening my understanding of education research, methodology, policy and leadership. A special thank you to Dr. Kathleen Scalise for providing me the knowledge and confidence for statistical analyses.

I would like to thank the members of my committee, Julie Alonzo, for her thoughtful edits and teaching of the writing process, Dr. Roland Good for his time and insight on the reading construct, and Dr. Ben Clarke for encouraging me to think about the practical implications in the education setting. A special thank you to Dr. Gerald Tindal, my advisor, for his patience, guidance, and knowledge as he guided me through the dissertation process. I will always remember his kindness and encouragement.

In addition, special thanks are due to my D.Ed. cohort for the interesting discussion and support through numerous hours of study.

A debt of gratitude to my students throughout the years, who pushed me to continue to look for new ways to support them and create opportunity for all students to learn.

A special thank you to my husband, Mike who provided the support, encouragement and edits needed. My sons, Kevin and Connor who inspire me to be a better person. My mother, Kathy, a role model who shares my passion for education and lifelong learning. Also my father, Glen, who instilled a strong work ethic in me.

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF FIGURES

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

In 1965, the Federal Government passed the Elementary and Secondary Education Act of 1965 to address inequalities in education that had become apparent during the Civil Rights Movement (Elementary and Secondary Education Act [ESEA] 1965). In 2002, the 107th Congress revised ESEA to focus on closing the achievement gap through increased accountability, flexibility, and school choice, and renamed it the *No Child Left Behind Act of 2001* (No Child Left Behind [NCLB], 2002). The achievement gap is the amount of difference in academic performance between subgroups of students and their peers (United States Department of Education [USDOE]; 2012).

In 2008, the Federal government funded President Obama's Race to the Top (RTTT), an initiative to allow states to be innovative and create programs to prepare students for college and career readiness (U.S. Dept. of Ed., 2011). Part of RTTT's agenda is to build data systems that measure student growth and success and inform teachers and principal how to inform instruction. These systems will address student achievement and provide information to guide the turn-around of low-performing schools (U.S. Dept. of Ed, 2010).

Critical Definitions for Measuring Student Achievement

To gauge student and school achievement, NCLB shifted the focus of education reform from teacher quality to student achievement data (O'Donnell & White, 2005). The increase in academic achievement made by students from different demographic subgroups measures the progress of teachers and schools in closing the achievement gap. (Davis, Darling-Hammond, LaPointe & Meyerson, 2005; Glanz, 2005: Louis, Dretzke & Wahlstrom, 2010; Stumbo & McWalters, 2011). These subgroups include students with disabilities, students from low socio-economic backgrounds, English Language Learners, and students from major racial/ethnic groups (e.g., White, Black, Asian/Pacific Islander, Latino, and American Indian/Alaskan Native). Student academic achievement is most commonly measured by students' performance on statewide achievement tests based on grade-level, academic content standards (Davis et al., 2005; Louis et al. 2010; USDOE, 2012).

State standardized tests are outcome assessments given at the end of the year and are used for school, district and state reporting purposes (Nese, Park, Alonzo & Tindal, 2011). To gauge students' success towards meeting yearlong content goals it is important to preview their progress toward end of the year proficiency. Interim curriculum-based measures and large-scale state standardized tests are two ways to collect data on student performance (Hosp & Hosp, 2003). Curriculum-based measures (CBMs) can provide evidence of students' progress towards academic benchmarks throughout the year to guide decision-making about the effectiveness of individual student performance in their instructional program (Deno, Marston & Tindal, 1986; Fuchs, L., 2004, Tindal, 2013). At the school level, CBMs can be used to evaluate each student's performance towards mastery of specific standards (annual targets), and at the district and state level they can provide feedback on the effectiveness of a school's instructional programs (Nese et al., 2011).

This study's focus is the relation(s) between student performance on curriculum-based measures (CBMs), and the statewide content standards-based achievement test in Oregon. Specifically, this study investigates whether curriculum-based easyCBM measures of third grade math and reading are valid predictors of student performance on the end-of-year, standards-based, Oregon Assessment of Knowledge and Skills (OAKS) in mathematics.

The literature synthesis discusses construct validity to identify if math measures are appropriate and valid methods for inferring student mathematical knowledge and skill level. Reading and mathematics measures used to identify student knowledge and skill level at the classroom (benchmark) and state (accountability) level follow. Measures are discussed in their capacity to inform both accountability and identification decisions. The purpose of the research is to identify whether adding a reading measure to mathematics benchmark screening improves the identification of students at risk for learning difficulties in mathematics.

Construct validity. Assessments need to provide trustworthy and appropriate inferences about a student's performance based on information specific to the construct being measured (Embretson, 1983; Kane, 2002; Kiplinger, 2008; Messick, 1989). *The Standards for Educational and Psychological Testing* state, "Validity is the most fundamental consideration in developing and evaluating tests" (American Education Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 1985). Validity is not a property of the test but the meaning of the test scores and function of the person's context of the assessment (AERA, APA, NCME, 1985; Messick, 1995). Kane (2013) suggests an effective assessment clearly defines the measured construct and provides guidance regarding the proposed interpretation and uses of the test scores. Ketterlin-Geller, Yovanoff, and Tindal (2007) note that the construct of the assessment needs to be clearly defined to accurately measure a student's knowledge and skills, and to reduce the impact of other factors on student performance. If assessment results guide decision-making and instructional consequences, it is important to make sure assessments accurately measure the identified construct (Kane, 2006; Ketterlin-Geller et al., 2007; Messick, 1989; Shaddish, Cook, & Campbell, 2001).

Threats to validity. Two threats to test validity, *construct-under representation* and *construct-irrelevant variance* are important to note when making interpretations of students' knowledge and skills (Kiplinger, 2008). *Construct underrepresentation* occurs when the tasks fail to measure important dimensions of the construct (Kiplinger, 2008).

The Oregon Assessment of Knowledge and Skills (OAKS) assess three mathematical domains each with multiple skills from the Common Core State Standards (CCSS) 11 mathematic domains. Three assessed in the third grade are: (a) numbers and operations, (b) geometry, and (c) number operations and algebra (http://www.core standards.org/Math/Content/3/introduction). CCSS also includes a mathematical practice standard that includes eight areas of processes and proficiencies important for the development of student competency in conceptual understanding and procedural fluency in mathematics (ODE, 2011). Each of these standard domains varies in complexity and proportion on the test, which may lead to under- or over- representations of a domain of the construct, resulting in threats to construct validity.

Construct-irrelevant variance occurs when multiple tasks or variables that are not part of the construct affect item difficulty (Ketterlin-Geller et al., 2007). Reading requirements in subject matter assessment, for example, can affect interpretations of scores and be a cause of *construct- irrelevant variance* (Messick, 1995). An example is reading within a mathematical word problem interfering with students' ability to demonstrate their mathematical knowledge and skill due to their inability to read proficiently and gain meaning from the text.

Accommodations. Providing individual accommodations may improve the interpretation of outcomes and thereby result in more construct valid inferences (Kane, 2001: Lai Berkeley; 2012). Accommodations are supports for students that provide equal opportunity to demonstrate knowledge and skills (Helwig & Tindal, 2003; Ketterlin-Geller et al., 2007).

Multiple studies have examined the use of accommodations to reduce the impact of individual disabilities and allow students to demonstrate their knowledge on the assessed construct (Fuchs, Fuchs, Eaton, Hamlett, & Karns, 2000; Helwig & Tindal, 2003; Hollenbeck, Rozek-Tedesco, & Tindal; 2000; Ketterlin-Geller et al., 2007). Examples of accommodation are alterations to the presentation or response formats, adjusting the testing environment, allowing extra time, or reading aloud items that are not measuring the reading construct (Helwig & Tindal, 2003; Ketterlin-Geller et al., 2007).

Appropriate accommodations can reduce interference from sources of *construct-irrelevant variance*. Administration guidelines and/or test environments are examples of efforts to control this source of variance. The OAKS math test has nine pages of guidelines for the read aloud accommodation to be validly administered (ODE, 2012).

The implementation of these guidelines is critical to maintain the construct validity. Computer administered easyCBM math measures have a read aloud accommodation built into them where students can click on a speaker icon and have both the question and responses read aloud to them (Alonzo & Tindal, 2012). In the recently adopted Common Core State Standards (CCSS), it implicitly states that appropriate accommodations are allowable to ensure maximum participation for students with a wide range of abilities.

Researchers suggest the current construct of math proficiency may include a combination of literacy and numeracy attributes (Crawford, Tindal & Steiber, 2001, Jiban & Deno, 2007; Thurber, Shinn, & Smolkowski, 2002). In assessing mathematics skill, however, it is important that reading difficulties are not a barrier for students or a potential source of *construct-irrelevant variance* (Alonzo, Lai, & Tindal, 2009). Considerable research, however, has found indications of a relationship between reading performance and math performance. Swanson and Jerman (2006) reported reading difficulties are an important correlate of math difficulties, and of outcomes on large-scale assessments in math. Additional researchers reported math disabilities are more likely to occur with reading disabilities than in isolation (Bryant, D.P. , Bryant, B.R., Gersten, Scammacca, & Chavez, 2008; Clarke, Smolkowski, Baker, Fien, Doabler & Chard, 2011; Fuchs, L.S., Fuchs, D. & Prentice, 2004; Vukovic, Lesaux, & Siegel, 2010). In their research, Jordan, Hanich, and Kaplan (2003) found reading abilities influence growth in mathematics achievement although mathematics abilities do not affect reading achievement. The association of reading and math suggests that a measure sensitive to reading difficulties may be useful in screening for math disabilities (Fletcher, 2005).

As a consequence, multiple studies have been conducted on read aloud accommodations for math assessments (Fuchs, L.S. et al, 2000; Helwig, Anderson & Tindal, 2002; Ketterlin-Geller et al, 2007; Lai & Berkeley; 2012; Weston, 2003). Studies found the complexity of the reading associated with a problem could interfere with a student's demonstration of their math knowledge or skill level (Helwig, Rozek-Tedesco, Tindal, Heath & Almond., 1999; Ketterlin-Geller et al., 2007). Weston (2003) examined the effects of read aloud accommodations on the *National Assessment of Educational Progress* (NAEP) math subtest and found all students showed improvement with those labeled learning disabled receiving the most significant benefits. Ketterlin-Geller et al. (2007) also found struggling readers to benefit from reading accommodations but proficient readers did not demonstrate the same benefit.

Curriculum-Based Measures

In some of this research on accommodations, researchers have used CBMs to determine the need for changes and adaptation in the state test; some of these changes have served as accommodations (reading the math test) and other served as modifications (reading the reading test). CBMs for benchmark screening and progress monitoring identify students' level of performance. These performance indicators guide further instruction and assessment. If CBM results indicate a reading disability, then accommodations such as a read aloud on a math assessment may increase the validity of the math results.

Initially developed for students in special education (Foegen, Jiban, & Deno, 2007; Fuchs, L. 2004), the increased emphasis on accountability and student measurement in federal education policy has expanded CBM use to include students from all education classifications (Alonzo et.al., 2009; Deno, 2003; Foegen et al., 2007; Park, Anderson, Alonzo, Lai & Tindal, 2012).

The purpose of CBM is to empirically measure changes in student performance over time towards mastery of a specific skill sequence (Deno, 1992; Fuchs, L., 2004; Nese, Park, Alonzo & Tindal, 2011). Students take both benchmark measures to determine students at risk and for those deemed at risk, repeated alternate forms of an assessment to measure changes in their progress on equivalent forms of the same task over a given period of time (Clarke, Baker, Smolkowski & Chard, 2008; Deno, 1992; Fuchs, L., 2004). A comparison of each student's data to his or her peers allows judgment of the degree of difference between his or her performance and that of his or her same-age peers in the same material at the same time (Deno, 1992; Fuchs, 2004). Using benchmark screening and progress monitoring, a data base for each individual student is formed to identify whether the student needs additional instructional support and/or is making adequate academic progress towards the expected benchmark (Deno, Fuchs, Marston & Shinn, 2001; Fuchs, L., 2004; Hosp & Hosp, 2003; Nese et al.; 2011). Teachers can analyze a student's data and make informed decisions allowing them to match instruction to student need (Deno, 1992; Foegen et. al., 2007; Fuchs, L., 2004; Nese et al, 2011).

Approaches to mathematics CBM. There are two different approaches used to develop CBMs in mathematics (Foegen et al., 2007; Fuchs, L., 2004; Tindal, 2013). Mastery monitoring (MM), referred to as sub skill mastery measure (SMM) in mathematics, is the systematic sampling of single skills within the year-long grade-level curriculum. General outcome measures (GOM) cover multiple skills that link growth on the CBM to greater understanding of the domain of mathematics (Christ & Vining, 2006; Clarke et al., 2008; Fuchs, L., 2004; Tindal, 2013).

SMM are highly sensitive and precise measurements used to assess the trend of achievement for specific skills such as operational fluency. (Fuchs, L.S. & Deno, 1991). When students are expected to make rapid gains and master the content in short periods of time, these are appropriate assessments (Fuchs, L.S. & Deno, 1991). All the measures are of equivalent difficulty and duration to evaluate performance on the specific skill. This form of CBM is a diagnostic tool to isolate skill deficits, evaluate mastery of a specific skill and determine instructional effects over a brief period of instruction (Christ & Vining, 2006; Nese et al., 2011)

A goal of GOM is to measure student achievement in the current grade level curriculum (Fuchs, L.S. & Deno, 1991). These CBMs consist of a few problems from each skill in the grade-level curriculum administered on a regular schedule to identify growth toward year-long competence of mathematics standards (Christ & Vining, 2006). Educators use multiple forms of the test with equivalent questions from each of the represented constructs. An example is annual curriculum at third grade, which would include the skills of multiplication and division, development of understanding of fractions, development of understanding of area and perimeter, and ability to describe and analyze two-dimensional shapes (National Council of Teachers of Mathematics, [NCTM], 2006). Students would take repeated assessments that are of equivalent difficulty throughout their third grade year comprised of questions assessing their understanding of these critical concepts to monitor their progress on these constructs within the yearly curriculum (Fuchs, L., Fuchs, D. & Courey, 2005; Tindal, 2013).

Screening, instructional placement, progress monitoring over extended periods of instruction, intervention evaluation, and identification of math disability use student GOM scores (Christ & Vining, 2006).

GOMs aligned with NCTM focal points. EasyCBM mathematic GOMs developers used the National Council of Teachers of Mathematics' Focal Point Standards (NCTM, 2006) in designing their assessments. These focal points ensure curriculum with adequate depth in the most important topics underlying success in school algebra (NMAP, 2008). The CBM assessments focus on students' conceptual understanding more than basic computational skills. Third grade level focal points include (a) numbers and operations, (b) geometry, and (c) numbers, operations and algebra. Numbers and operations focus on the basic use of operations (addition, subtraction, multiplication, and division) while numbers, operations and algebra, is designed to assess algebraic thinking using the four operations. EasyCBM reports a math benchmark score. Table 1 illustrates the NCTM third grade curriculum focal points measured on easyCBM.

Table 1

NCTM Third Grade Focal Points Assessed by easyCBM

Curriculum Focal Point	Standard
Numbers and Operations	Developing an understanding of fractions and fraction equivalence
Geometry	Describing and analyzing properties of two-dimensional shapes
Numbers, Operations, and Algebra	Developing understandings of multiplication and division and strategies for basic multiplication facts and related division facts

Reading curriculum-based measurement. EasyCBM reading measures can be used to monitor students' reading development over time, beginning with foundational skills (e.g., letter naming fluency, letter sound fluency) to more complex skills (e.g., vocabulary and reading comprehension). The basis for these skills and constructs is the National Reading Panel report (National Institute of Child Health & Human Development [NICHHD], 2000) and the five big ideas identified. Empirical findings of which assessments provided the most robust screening of content and skill for the grade level assessed inform measure selection. Yovanoff, Duesberry, Alonzo and Tindal (2005) report in their research findings that oral reading fluency, vocabulary, and comprehension CBMs implemented in the classroom are reliable and valid measures of student achievement. Reading fluency and vocabulary explained 40% to 50% of the variance in reading comprehension (Yovanoff et al., 2005) making them important across grade levels.

Fluency. The majority of reading research and classroom use of CBM has focused on reading fluency (Fuchs, L.S., et al., 2001; Good et al., 2001; Paris, 2005). In the research, oral reading fluency (ORF) was found to be a robust indicator of overall reading ability by its strong correlation with other criterion measures (Foegen et. al, 2007; Roehrig, Yaacov, Nettles, Hudson, & Torgeson, 2008; Shapiro, Keller, Lutz, Santoro & Hintze, 2006, Wood, 2006). Oral reading fluency (ORF) is a key component of reading CBM (Alonzo et al., 2006; Christ & Ardoin, 2009; Nese et al., 2011) and identified in the literature as one of the strongest predictors of reading comprehension in the early grades (Hasbrouk & Tindal, 2006). Fluency connects early reading skills as students move to reading words, phrases and sentences (Yovanoff et al., 2005). ORF is a strong indicator of overall reading ability by the completion of first grade or at the latest middle of second grade (Fuchs, L.S. et al., 2001).

Silberglitt, Burns, Madyun, and Lai (2006) report a deceleration in growth of reading fluency skills as grade level increases. Researchers suggest reading fluently becomes less of an instructional focus as students master reading and as instruction in the process of reading to gain information from content area materials gains importance (Nese et al., 2011; Wayman, Wallace, Wiley, Ticha & Espin, 2007; Yovanoff et al., 2005). Ehri (2005) describes a fluency threshold that readers achieve that allows them to focus on comprehension rather than decoding, at which point fluency is no longer sensitive to increases in reading comprehension.

Comprehension. As students reach a minimal fluency level, they attain a sufficient degree of automaticity, and vocabulary knowledge becomes a more informative indicator of reading comprehension (Yovanoff et al., 2005). Lack of vocabulary knowledge can lead to incorrect inferences of meaning and difficulties with reading comprehension. In addition, vocabulary is necessary for oral and written communication. For students to comprehend, they must be able to attach meaning to words.

Comprehension gives reading a purpose and allows the reader to interact and make sense of the text (Nese et al., 2011). Through metacognition, the process of thinking about one's thinking, readers use comprehension strategies to understand, remember and communicate what they read. The multiple choice reading comprehension measure on easyCBM assesses students' literal, inferential and evaluative comprehension of an original narrative fiction passage (Lai, Irvin, Park, Alonzo &Tindal, 2012).

Statewide Assessment

In Oregon, the state legislature passed the Oregon Education Act for the 21st Century in 1991 (Oregon School Boards Association [OSBA], 2005), which required schools and districts to inform the public about achievement in Oregon schools. To meet the requirements of the Oregon Education Act as well as the federal Improving America's Schools Act (IASA,) passed in 1994, the state established a statewide assessment program (Conley, 2007). The design of the Oregon Assessment of Knowledge and Skills (OAKS) tests provide comparison data of student performance and progress over time within their school, district, and state on grade-level content standards (Oregon Department of Education [ODE] 2008). To compare the performance of individual students and schools against state norms and growth models data is collected. A statewide report card shares school wide and district wide results with the public.

The OAKS is a criterion-referenced test tied to the Oregon content and achievement standards. The purpose of a criterion-referenced test is to identify the specific knowledge and skills each student can demonstrate (ODE, 2011b). An achievement scale ranging from 150 to 300 (ODE, 2010) reports scores for the assessment. Each point on the scale is an equal distance apart, allowing comparison from year to year. Score reports in specific skill areas provide educators information on areas of improvement.

Third grade OAKS mathematics. Third grade is the first year students take the reading and math OAKS. In math, student scores include three subscores along with an overall total sore. The three subscores are a) Numbers and Operations, b) Numbers and Operations, Algebra, and Data Analysis, and c) Geometry and Measurement (ODE, 2010b). Table 2 includes skills assessed from Oregon Content Standards and reported as subscores on OAKS mathematics.

Table 2

OAKS Math Sub Scores and Skills Assessed

OAKS Math Subscores	Skill Assessed from Standard
Numbers and Operations	• Develop an understanding of fractions • Represent and order common fractions • Add and subtract fractions with common denominators
Numbers and Operations, Algebra, and Data Analysis	• Demonstrate understanding of multiplication and concept of multiplication as repeated addition. • Demonstrate understanding of division and concept of division as repeated subtraction. • Use the inverse of operations to identify patterns and solve problems • Interpret graphs, tables, and charts
Geometry and Measurement	• Analyze two dimensional shapes using angle measurements and numbers of sides to classify them • Computing perimeter and area using both American standard and metric units of measurement

Summary of measures. CBMs and state summative assessments both provide information on student achievement, yet vary in purpose and design. The purpose of CBM is to document within-year achievement levels and monitor growth. Results for specific students guide instruction. OAKS is an end of the year assessment designed to hold schools and districts accountable for student growth and achievement.

How Do We Identify Students "At Risk" for Learning Difficulties?

Early identification of students with learning difficulties is critical for decision-making about student academic progress and changes in instruction. Educators use benchmark measures developed to identify students' academic progress and predict which students will need additional instruction (Fuchs, L. Fuchs, D., Compton, Bryant, Hamlett, & Seethaler, 2007). Screening tests are given to all groups to assess students' current level of proficiency and identify areas of difficulty (Steiner, 2003), such as fluency, comprehension, or computation. Screening assessments in reading and mathematics are used to measure if students are performing at the expected grade level given their current amount of instruction. Benchmark screenings are periodically (e.g., fall, winter, spring) given to all students to identify students who might be "at risk" at their current level of instruction (Anderson, Park, Irvin, Alonzo & Tindal, 2011; Wright, 2005). "At risk" students are students who are not making adequate progress towards proficiency of grade-level standards. Identified students should receive interventions, considered academic boosts, to bring the student to the level of their peers (Bryant, P. et al, 2008).

Benchmark screening tools. Multiple benchmark screening measures are available to screen reading and mathematics performance. Research on assessment and their ability to predict future performance in complex processes such as mathematics and reading is used to develop these measures (Hasbrouck & Tindal, 2006). Measures used for screening need to be valid, reliable, efficient, and classify accurately which students are "at risk" and "not at risk" for learning difficulties (Hasbrouck & Tindal, 2006; National Center on Response to Intervention [NCRTI], 2010 May). The screeners need to be developmentally and age appropriate for accurate data based decision-making to occur (NCRTI, 2010 May). Curriculum-based measures (CBM) are benchmark measures and screeners with good reliability and validity (Hosp M.K., & Hosp, J.L., 2003). At the third grade level, universal screening for reading ability should assess word and passage reading, oral reading fluency (Fuchs, L.S., Fuchs, D., Hosp & Jenkins, 2001; Good, Simmons, & Kame'enui, 2001), and reading comprehension (Hosp & Fuchs, 2005; Jenkins, Hudson & Johnson, 2007; Torgesen, 2002).

Benchmark screeners in response to intervention (RTI). One model of identification, intervention and instruction is Response to Intervention (RTI). RTI is an integrated evidence-based approach that includes general and special education students. The goal of RTI is to prevent learning problems through the early identification of students who are demonstrating learning difficulties, then provide evidence-based multi-tiered interventions (Bryant, P. et al., 2008). The RTI model for prevention of academic difficulties is composed of different tiers in which instruction varies by its intensity, explicitness, and individualization (Bryant, P. et al, 2008; Bryant, D. et al., 2011; Clarke et al., 2011; Fuchs, L. et al., 2007; National Center on Response to Intervention [NCRTI], 2010). The model has the potential to reduce the prevalence of reading and math disabilities and enhance student achievement (NCRTI, 2010). Furthermore, the model addresses the achievement gap for minority students by early identification of learning difficulties and focused intervention programs. Measurements of student performance to identify at risk students, monitor progress and predict positive learning outcomes are necessary for the model to work (Fuchs, L. et al., 2007).

Predictive Relationship of CBM on State Accountability Test Scores

Multiple studies researched the predictive qualities of reading fluency CBMs on state accountability tests. Crawford, Tindal and Stieber (2001) found a strong correlation between oral reading rate and test performance on a statewide reading test. Stage and Jacobson (2001) found oral reading fluency significantly predicted student scores on the Washington Assessment of Student Learning (WASL) at the fourth grade level, with an increased predictive power of 30%. McGlinchey and Hixson (2004) found similar results for fourth grade students on the Michigan Educational Assessment program. Hintze and Silberglitt (2005) found a strong predictive correlation between reading CBM and the Minnesota Comprehensive Assessments (a criterion- referenced standardized achievement test) at the third grade level. In their study of the correlation of reading fluency CBM with the Minnesota Comprehensive Assessments in reading at grades three, five, seven and eight, Silberglitt, Burns, Madyun, & Lail (2006) found the strongest correlation, 0.71, at third grade and weakest correlation, 0.51, for eighth grade students. The researchers suggest the correlation between fluency and reading ability becomes weaker as the reader advances grade level (Nese et al., 2011; Silberglitt et al., 2006).

In a study of the relation between easyCBM reading measures of oral reading passage fluency (PRF), vocabulary (VOC) and multiple-choice reading comprehension (MCRC), and the OAKs assessment for grade four and five, Nese et al. (2011) found all three CBMs were predictors of OAKS results. In further analysis, VOC was determined the strongest predictor, followed by MCRC, with PRF being the least significant (Nese et al, 2011).

Jiban and Deno (2007) found CBMs of cloze math facts and reading fluency to be the best predictor of math performance on the Minnesota Comprehensive Assessment in mathematics. They conclude that reading performance does correlate significantly with state math test results for third grade.

An extensive exploration of the relationship between easyCBM reading and statewide assessment of mathematics has not occurred. The OAKS assessment contains an academic vocabulary list for each grade-level mathematics test. There is documentation of vocabulary predicting reading proficiency (Nese et al., 2011). A better understanding of the relationship between reading development and its effect on mathematical learning may further inform benchmark screening and summative standards-based assessment in mathematics. Math disabilities are an “underestimated” research topic (Gregoire & Desoete, 2009), and researchers confirm the need to identify the “core deficits” of mathematic ability (Chiappe, 2005).

Research Questions

The purpose of this study is to examine the relationship of third grade students’ performance on fall mathematics and reading benchmark measures with their performance on summative mathematics assessment in the spring. Fall scores on easyCBM mathematics and reading are analyzed with spring scores of OAKS mathematics at the third grade level.

The specific research questions are as follows:

1. What is the correlation between performance on the fall easyCBM mathematics and reading measures and performance on the Oregon Assessment of Knowledge and Skills (OAKS) in mathematics?

2. Which fall easyCBM measures (math, reading fluency, vocabulary, or reading comprehension) best predict math performance on OAKS?

3. What is the accuracy (sensitivity and specificity) of easyCBM math and reading measures in correctly classifying students "at risk" on OAKS?

CHAPTER II

METHODOLOGY

This study included correlation, linear regression and Receiver Operator Characteristics analyses (ROC) to investigate the relationships of interest from existing grade 3 data sets collected by Behavioral Research and Teaching (BRT), a research institute at the University of Oregon. Data came from three district-wide standardized assessments: (a) easyCBM-mathematics (Alonzo et. al., 2006), (b) easyCBM reading assessments of PRF, VOC, MCRC (Alonzo et. al., 2006), and (c) OAKS- mathematics (ODE, 2008). The analyses included data from students who completed all five analyzed measures: the fall easyCBM reading assessments of PRF, VOC and MCRC, the fall easyCBM mathematic assessment, and the statewide (OAKS) mathematics assessment. Students had the opportunity to take the OAKS twice, and in this study, the analyses included the highest score. A description of specific (a) settings and participants, (b) measures, (c) procedures, and (d) analyses follows.

Participants and Setting

This study includes data collected during the 2011-2012 school year in a Pacific Northwest state with a sample of 16,207 third grade students from 29 school districts. There were 141 schools within those districts included in the study. Demographic information obtained from school records includes English language learners, special education status, race/ethnicity (“of color” or “not of color”) and gender.

Measures

Measures used in this study include CBMs publically available through the easyCBM system (www.easycbm.com) and the OAKS available from OAKS Online (http://www.oaks.k12.or.us/portal/). The CBMs used in this study assess the mastery of skills and knowledge deemed critical from the curricula at the third grade level. The OAKS mathematics used is an assessment of students' mastery of Oregon third grade mathematic content standards given to students annually.

EasyCBM development and alignment. The easyCBM mathematics and reading measures were developed following the guidelines for test development described in *The Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999) using the principles of Universal Design for Assessment (Alonzo et al., 2009). For each measure, developers sought to create tests that were sensitive to individual growth in short periods to allow for meaningful interpretation of growth over time (Alonzo & Tindal, 2007). The measures were written using experienced grade-level educators and reviewed by assessment researchers at the University of Oregon. Analyses of the items include bias/sensitivity. Grade-level students piloted items. Researchers used Item Response Theory (IRT) to analyze the specific items for each measure to increase the sensitivity of the measures in monitoring growth (Alonzo et al., 2009).

IRT is used to place students and items on the same scale, which can then be used to develop equivalent alternate forms based on difficulty, standard error, discrimination, and mean square outfit (Sage, Chapter 13, 2007). The Rasch model or the one-parameter logistic model (1PL) is the form of IRT used in designing easyCBM measures (Alonzo, Lai, & Tindal, 2009; Lai et al., 2012). This allows an individual’s response to an item be determined by the individual’s trait level and the item difficulty. (Sage, Chapter 13, 2007), providing equivalent forms of a test, so that students at the same trait level will have the same answer for a problem of the same difficulty. For each measure, multiple alternate forms are available at each grade level for benchmark screening and progress monitoring of students (Alonzo et al., 2009; Alonzo & Tindal, 2007, Nese et al., 2010)

Description of predictor variables. Student CBM math and reading scores in the fall of 2011-2012 school year are the predictor variables in this analysis. The curriculum-based measures used in this study are from the third-grade easyCBM benchmark assessments (Alonzo et al., 2006). The measures sample a year's worth of curriculum to determine the performance level of students towards mastery of the critical skills and knowledge for the grade level. Educators use easyCBM assessments as benchmark-screening tools to identify students at risk for math and reading difficulties and/or as progress-monitoring tools to understand how students perform over time (Alonzo, Park & Tindal, 2012). Analyses of data collected on individual students supports decision-making about student growth and instruction (Alonzo et al., 2006; Nese et al., 2011).

***EasyCBM measures of mathematics*.** The EasyCBM benchmark and progress monitoring mathematics measures are available for kindergarten through eighth grade. The NCTM (2006) curriculum focal points informed development of the math CBM.

EasyCBM measures of reading. At third grade, reading skills measured include fluency (passage reading fluency), vocabulary, and comprehension (multiple-choice reading comprehension). Developers of specific items for each measure utilize Item Response Theory (IRT) to increase the sensitivity of the measures to monitor growth (Lai et al., 2012). IRT is an advanced form of statistics that provides test developers with a tool to categorize an item by identifying the probability an individual with a specific trait will correctly answer a question at a specific level of difficulty (Item Theory, 2007). Developers create alternate forms of each measure with consistent equivalent difficulty (Lai et al, 2012; Nese et al., 2010).

EasyCBM reading skills measured. The easyCBM measures included as predictors in the analyses are the third grade level Passage Reading Fluency (PRF), Vocabulary (VOC), and Multiple Choice Reading Comprehension (MCRC) assessments.

PRF is an individually administered assessment of a student's ability to read connected narrative text accurately and fluently (Alonzo & Tindal, 2007; Nese et al., 2011). *VOC* provides an opportunity to evaluate a student's knowledge of words from the grade-level content standards (Alonzo, Andersen, Park, & Tindal, 2012). *MCRC* assesses a student's literal, inferential and evaluative comprehension of an original narrative fiction passage (Lai et al., 2012). Authors wrote passages specifically for use with the easyCBM progress monitoring and benchmark system. They were written for mid-year of a grade level (e.g., grade 3 is grade 3.5 level) and include approximately 250 words. Developers create each form to be of consistent length, and verified readability to fit grade level using the Flesch-Kincaid index feature available on Microsoft Word (Alonzo & Tindal, 2008).

Description of Criterion Variable

The Oregon Statewide Assessment (OAKS) is a criterion-referenced test tied to the Oregon content and achievement standards. The test is a summative assessment whose purpose is to identify the specific knowledge and skills each student can demonstrate at the end of an academic year (ODE, 2011b). Scores for the assessment are based on an achievement scale ranging from 150 to 300 (ODE, 2010). Each point on the scale is an equal distance apart, allowing comparison to be made from year to year. Scores are reported in specific skill areas to provide educators information on areas of improvement. The results are one piece of evidence of a student's level of performance and primarily used to compare achievement with Achievement Standards established by ODE (ODE, 2012).

OAKS is an online assessment without time constraints (ODE, 2008). The test is adaptive, and item difficulty varies with student performance on previous items. There is not a penalty for guessing answers. Two opportunities are available for each content area, and this study uses the highest score. Assessment scores are collected electronically and are scored against an answer key for a raw score. The raw score is converted to a scale score labeled in Rasch units or RIT scores (ODE, 2008). The Rasch unit scale accounts for students' response to items relative to the item difficulty. Educators use scores to measure student achievement and provide data to follow a student's educational growth (ODE, 2008).

OAKS mathematics. The OAKS math assessment contains multiple choice and constructed-response questions. There are three primary core standards per grade level. In third grade, they are (a) Numbers and Operations, (b) Numbers and Operations: Algebra, and Data Analysis, and (c) Geometry and Measurement. There are four to nine content standards associated with each content strand. Table 3 includes the content strands, common core curriculum goal and percentage of questions for the third grade test.

Table 3

Proportion of Content Strand Assessed on Third Grade OAKS Mathematics

Content Strand	Percent of Questions	Common Core Curriculum Goal
Numbers and Operations	35%	Develop an understanding of fractions and fraction equivalence.
Number and Operations, Algebra, and Data Analysis	35%	Number and Operations, Algebra, and Data Analysis: Develop understandings of multiplication and division, and strategies for basic multiplication facts and related division facts. 999999999999999999
Geometry and Measurement	30%	Describe and analyze properties of two-dimensional shapes, including perimeters.

The OAKS mathematics blueprint includes specific vocabulary used in the assessment at each grade level. Table 4 includes a list of the academic vocabulary expectations for third grade students on the third grade OAKS mathematics test.

Reliability and validity of study variables. Mathematics and reading easyCBM, as well as OAKS used in this study, are reliable and valid measures of student achievement, when used for their intended purpose (Lai et al, 2012; ODE, 2007; Park, Anderson, Alonzo, Lai, & Tindal, 2012). Researchers at Behavioral Research and Teaching (http://www.brtprojects.org) continue to investigate the technical adequacy of easyCBM measures.

Table 4
Assessable Academic Vocabulary Summary for Mathematics Third Grade OAKS

acute angle	denominator	frequency table	Mile	product	Tools
Add	diagram	geometric pattern	Millimeter	quadrilateral	Total
Addition	difference	greater than	mixed number	quotient	Transformation (transform)
Altogether	dimensions	growing pattern	Model	rectangle	Translation (translate)
Angle	distance	Hexagon	Multiple	reflection (reflect)	Trapezoid
area models	distributive	hundreds grid	Multiples	repeated addition	triangle (triangular)
Array	divide	Identity	Multiplication	repeated subtraction	two-dimensional
Associative	dividend	improper fraction	Multiply	result	Units
Attributes	division	Inch	number line	rhombus (rhombi)	vertex (vertices)
Axis	divisor	increasing sequence	number pattern	right angle	Whole
bar graph	dozen	Inside	Numerator	rotation (rotate)	whole number
Centimeter	equal	inverse (opposite)	obtuse angle	rule	Yard
Closed	equal groups	Isosceles	Octagon	scalene	zero property
Combine	equation	Key	Order	set	Degrees
Commutative	equiangular	kilometer	Parallel	side	Fraction
Compare	equilateral	Legs	Parallelogram	similar	Meter
Compose	equivalent	less than	Part	skip counting	Polygon
Congruent	expression	line of symmetry	Pentagon	slide	Table
Data	factor	line plot	Perimeter	square	
Decompose	flip	line segment	Perpendicular	subtraction	
decreasing sequence	foot (feet)	List	picture graph	sum	

Third grade easyCBM mathematics. Nese, Lai, Anderson, Park, Tindal, & Alonzo (2010) researched the relationship of third grade easyCBM mathematics measures and OAKS mathematics. For the overall student population, Crohnbach's alpha ranged from .70 to .80 for all three easyCBM third grade math benchmark assessments. The split-half reliability estimates were in the moderate range of .50 – .80 (Nese et al., 2010). Nese et al. (2010) reported all third grade easyCBM mathematics benchmark assessments were predictive of spring OAKS mathematics (Fall, R^2 =. 48, n = 3302, β = 0.69; Spring, R^2 = .54, n = 3119, β= .74), to support the construct validity associated with easyCBM.

Third grade easyCBM reading. Researchers examined easyCBM reading fluency measures and found moderate test-retest reliability and moderate to high alternate form reliability (Lai et al, 2012; Park et al., 2012). Table 5 reports specific correlation values.

Table 5

Third-grade Fluency Measure Reliability for easyCBM

EasyCBM Measure	Test- Retest Reliability	Alternate Form Reliability
Word Reading Fluency	.67 to .92	.72 to .92
Passage Reading Fluency	.84 to .94	.92 to .90

In a convenience sample of third graders (*n* = 288) from a 10 school sample, easyCBM was compared to Gates-MacGinitie Reading tests and DIBELS to evaluate criterion validity. The EasyCBM VOC demonstrated a low to moderate correlation (r_s =

.39) with Gates- MacGinitie Word knowledge (Lai, Alonzo, & Tindal, 2013). The easyCBM MCRC and CCRC also had a moderate correlation (r_s = .41) to the Gates-Mac Ginitie Reading Comprehension (Lai et al., 2013). In addition, the easyCBM passage reading fluency had a strong correlation (r = .91) to DIBELS (Lai et al., 2013).

The validity of an assessment is determined by the use of the results or valid inferences made from the assessment results (Kane, 2002; Messick, 1995). Park, Anderson, Irvin, Alonzo, and Tindal (2011) researched the diagnostic efficiency of 3rd grade easyCBM reading measure in relation to the statewide OAKS summative assessment. Using Receiver Operating Curve Analysis (ROC) they analyzed the probability of students being classified correctly "at risk" or "not at risk" for meeting the OAKS reading benchmark. In their analyses they attempted to maximize sensitivity (approximately .85) while maintaining a high level of specificity (approximately .71) (Park et. al., 2011). The Area Under the Curve (AUC) was used to determine how well the measures were properly classifying students with a value of 1.0 being perfect and .50 representing chance. The overall accuracy was high (PFR, AUC = .90-.91, MCRC, AUC = .69-.74, VOC, AUC = .83-.84), with PRF showing the greatest classification accuracy.

Third grade OAKs mathematics. In multiple studies, third grade OAKS mathematics has been determined to be a reliable and valid measurement of student level of achievement on third grade benchmarks (ODE, 2007). To determine reliability, an analysis of standard error of measurement provided evidence of reliable test scores for students of all abilities except the extreme ends of the distribution (ODE, 2007).

The relationship of third grade OAKS with state and nationally normed tests was examined to provide evidence of concurrent validity. ODE (2007) reports third grade

OAKS mathematics has a strong correlation to the California Achievement Test (.74) and the Iowa Test of Basic Skills (.76). A moderate correlation with the Northwest Evaluation Association math content test (0.66) was reported (ODE, 2007).

Procedures

I used extant data from Grade 3 Fall 2011 benchmark scores on easyCBM math and reading assessments of PRF, VOC, and MCRC, and the OAKS – math assessment administered in the spring of 2012. Students who took the state's extended assessment were not included. Standard accommodations for each measure were allowed for all students; data on specific students were not recorded. The easyCBM PRF is administered individually with the student reading to a trained assessor. The easyCBM VOC and MCRC measures and OAKS assessments were computer-based and administered by trained educators.

Training and administration for easyCBM math and reading assessments. Training for the easyCBM assessments is online, and a teacher's manual is available to all educators. For all of the assessments, students on individualized education plans (IEP) are given the opportunity to use allowable accommodations. This includes printing a test and answering the questions with paper and pencil before inputting them. For the math assessment, students are allowed to have problems read aloud to them.

Passage reading fluency (PRF) is a measurement of the number of words read correctly in connected text by a student in a given amount of time. PRF is measured by having a student read grade level passages and counting the number of words read correctly in one minute. Passage administration is standardized in that the same administration protocol is used with all students: the assessor places a copy of the passage in front of the student, reads the directions verbatim from the assessor's copy of the passage, and begins his/her stopwatch as soon as the student reads the first word in the passage. The assessor scores his or her copy of the passage. Words read incorrectly or omitted are marked with a slash. Self-corrected words are marked with "SC" and counted as correct. (Alonzo & Tindal, 2012). A student's score is reported in words correctly read per minute (wcpm).

The VOC assessment is administered on the computer and can take place in a computer lab or classroom with a 1:1 student ratio of computers to students. Students are given a multiple-choice assessment with 20 questions with three possible answers for each one. The correct answer is a synonym of the tested word, and then there is a nearly correct and far from correct answer. Each correct answer is worth one point for a possible total of 20 points.

MCRC is also administered on a computer in the same setting as the VOC assessment. In third grade, students read a 1500 word passage and then answer 20 multiple-choice questions based on the story. The questions include seven questions each targeting literal and inferential comprehension and six questions targeting evaluative comprehension (Lai et al, 2012). Each form of the test includes easy, moderate, and difficult questions for each level of comprehension. It takes students approximately 20 to 30 minutes to read the story and 10 to 20 minutes to answer the questions. Students are encouraged to first read the entire story, and then answer the questions. Students are allowed to move back and forth between the passage text and question, and students can change their responses during a testing session. Tests are computer scored. Students receive one point for each correctly answered question.

Training and administration for OAKS math assessments. OAKS is administered to Oregon students annually (ODE, 2008) online using a 1:1 student to computer ratio. Educators must receive training and sign an agreement to facilitate the assessment during a predetermined test window established by the Oregon Department of Education (ODE). The test is untimed, and students determine their pace during testing sessions. Students can review the questions and their answers during each session, but are not allowed to revisit problems from previous testing sessions. On the mathematics assessment, students are allowed to have the questions read using predetermined specifications on precise wording for all symbols and number names (ODE, 2012).

Variable codes for analyses. Performance variables included in the data sets are fall 2011-2012 benchmark scores from easyCBM reading and math and 2011-2012 school year OAKS mathematics scores. The benchmark scores are used for the correlation analysis and to analyze the variance of the CBMs in their effect on the statewide assessment. For the ROC and AUC analyses, easyCBM benchmarks are also recoded into a new variable with two values (0 = meets; 1 = does not meet) to designate the relationship to the 50th percentile. The average grade-level scores for progress monitoring within the easyCBM assessment system corresponds with the 50th percentile. (http://www.easycbm.com/static/files/pdfs/info/ProgMonScoreInterpretation.pdf). In addition, OAKS mathematics was recoded with two variables (0 = meets; 1= does not meet) to designate results that met the third grade benchmark score of 212 for the analyses (ODE, 2012). Table 6 lists the variables and their variable codes for this study.

Table 6

Performance Variables' Names, Description, and Coding Definitions

Variable Name	Variable Description	Coding
PRF_fall	Results of Fall easyCBM -Passage Reading Fluency	Words read correctly per minute
VOC_fall	Results of Fall easyCBM –Vocabulary	0 - 25 continuous
MCRC_fall	Results of Fall easyCBM -Multiple choice Reading Comprehension	0 – 20 continuous
MATH_fall	Results of Fall easyCBM – Mathematics	0 - 45 continuous
OAKS Math	Results of 3rd grade OAKS –math	0 = Meets 1= Does not meet
PRF_Ben	Achievement of PRF of 85	0 = Meets 1 = Does not meet
VOC_Ben	Achievement of VOC of 17	0 = Meets 1=Does not meet
MCRC_Ben	Achievement of MCRC of 12	0 = Meets 1 = Does not meet
MATH_Ben	Achievement of MATH of 31	0 = Meets 1 = Does not meet
OAKS Math_Ben	Achievement of OAKS- math of 212	0 = Meets 1 = Does not meet

Statistical analysis of the restructured data provided mean, median, and standard deviation of each individual measure. In addition, the study includes analysis to determine information about student performance in relationship to established benchmarks for each assessment. The study provides a correlation analysis to determine the relationships between the CBM reading and mathematics measures and their relationship to the OAKS mathematics test. Next, a simple linear regression was conducted to look at a possible predictive relationship between the CBM measures and OAKS mathematics. ROC and AUC analyses were performed to analyze the specificity and sensitivity of each CBM in predicting the outcomes on the OAKS mathematic assessment.

Analyses

The data initially gathered by Behavioral Research and Teaching (BRT) included 16, 207 students in Oregon from 29 districts and 141 schools. This data file was restructured to only include students who had completed all five assessments: Fall easyCBM PRF, VOC, MCRC, and MATH and OAKS mathematics. This inclusion rule eliminated 61.5% (*n* = 9,961) of the students from the data. Then a frequency analysis with boxplots was conducted to eliminate any outliers. Two student records for PRF fluency with scores of 246 and 273 were eliminated for being unrealistically high as the 90% score is 178. In addition, one student's record was deleted with a VOC score of 55, which is 30 points higher than the total possible.

An *a priori* decision was made to eliminate students with scores of '0' on easyCBM assessments. Four students had a score of '0' on the easyCBM math assessment. There were 69 students who had a single '0' score on easyCBM reading assessments (2 PRF, 14 VOC, 53 MCRC). In addition, six students had two or more assessments with '0' scores. These cases also were eliminated from the data set. Thirty-eight percent (*n* = 6,164) of the students in the study had valid scores for all five assessments. The students included in the clean, restructured data represented 27 school districts and 134 schools. The data were recoded to analyze demographic non-performance variables. Table 7 lists the variables and their code.

Table 7

Non-performance Variables for Students in the Participating Schools

Variable Name	Description	Coding
Gender	Male, Female, Did not report	0=Female 1 = Male
SpED	Special Education Eligibility	0 = Identified 1 = Not identified
ELD	English Language Development	0 = Not eligible 1 = Eligible
Race/Ethnicity		0 = "of color" 1 = "not of color"

Table 8 provides demographic data of students in the study with scores reported. Percentages are rounded to the nearest whole number.

Statistical analysis of the restructured data provided mean, median, and standard deviation of each individual measure. In addition, analysis was completed to determine information about student performance in relationship to established benchmarks for each assessment. A correlation analysis was conducted to determine the relationships between the CBM mathematics and reading measures, and the measures' relationship to the OAKS mathematics test.

Table 8

Demographics for Students Included in the Study

Gender	*Number of Students*	*Percent of Students*
Female	2987	48.5
Male	2946	47.8
Did not report	231	3.7
Special Education		
Yes	569	9.2
No	3129	50.8
Did not report	2466	60.0
English Language Learner		
Yes	461	7.5
No	2682	43.5
Did not report	3021	49.0
Ethnicity/Race		
"of color"	1347	21.9
"not of color"	2491	40.4
Did not report	2326	37.7

Multiple regression analyses were conducted to investigate the best predictors of mathematics performance on OAKS from fall easyCBM reading and mathematics performance. This provided information on the explanation in scores attributed to the different easyCBM measures. Sequential analyses examined the validity of each set of the predictors with math measures added in first, then PRF, followed by VOC and MCRC. Specifically, students' scores on CBMs of math, reading fluency (PRF), vocabulary (VOC), and reading comprehension (MCRC), and overall reading (PRF,VOC, MCRC) were entered as predictors in separate regression models with state-wide mathematics scores as outcome measures.

To determine the classification accuracy of easyCBM reading and mathematics measures' prediction of performance level on the OAKS mathematics, ROC analyses was performed. AUC analysis is used to identify classification accuracy. Classification efficiency statistics reported include sensitivity and specificity. Sensitivity refers to those identified "at risk" who are "not at risk" for math difficulties on OAKS, and specificity is those labeled "not at risk" who actually are "at risk."

CHAPTER III

RESULTS

The study provides descriptive statistics for all of the performance variables and frequency reports for the number of students meeting and not meeting the benchmark scores for all of the CBM assessments and statewide summative assessment. A correlational analysis yielded information about the association between the easyCBM measures and OAKS mathematics assessment, and between the individual easyCBM measures themselves. A multiple regression analysis provided information on the possible predictive nature of the easyCBM math, reading (PRF, VOC, MCRC) fluency (PRF) and comprehension (VOC, MCRC) measures on OAKS mathematics scores. Finally, a Receiver-Operator Characteristic (ROC) including Area under the Curve (AUC) analysis revealed the specificity and sensitivity of the easyCBM assessments in their predicted achievement on the OAKS mathematics benchmark.

Cases Included and General Description

This study included results reported to BRT of students who completed all five measures: easyCBM PRF, VOC, MCRC, Math, and OAKS Mathematics during the 2011-2012 school year. The easyCBM measures and OAKS scores are normally distributed (see appendix A) with a skew between negative one and positive one, allowing for parametric statistics. Descriptive statistics for easyCBM and OAKS mathematics follow in Table 9.

Table 9

Descriptive Statistics for Each Grade 3 Assessment

Measure	N	Mean	Std. Dev	Minimum	Maximum
PRF_fall	6164	87.37	38.684	1	232
VOC_fall	6164	14.75	4.329	1	20
MCRC_fall	6164	10.57	3.674	1	19
Math_fall	6164	29.65	6.196	1	45
OAKS_math	6164	214.72	10.478	178	272

For students included in the study, the mean score for PRF and OAKS math are above benchmark scores. Students' mean score was below the benchmark for all other assessments.

The study provides frequency analysis to establish information about student performance on specific benchmark tests. Student scores on the outcome variable (OAKS) in this study were higher than the overall scores reported by ODE. In 2011-12, 63.9% of Oregon third grade students received an achievement level of *meets* or *exceeds* on the OAKS benchmark (ODE, 2012). The scores ranged from 140 to 275. In this study, student scores ranged from 178 to 272, and 67.2% scored at the *meets* or *exceeds* achievement level. The distributions of students who *met* or *did not meet* benchmark scores follow in Table 10.

Table 10

Number of Students Making Grade 3 Performance Variable Benchmarks

Assessment	N		Percent	
	Meet	Did not meet	Met	Did not meet
PRF_Ben	3049	3115	49.5	50.5
VOC_Ben	2789	3375	45.2	54.8
MCRC_Ben	2626	3538	42.6	57.4
Math_Ben	2744	3420	44.5	55.5
OAKSmath_Ben	4134	2021	67.2	32.8

The highest number of students not meeting benchmarks occurred during the fall easyCBM MCRC assessment followed by easyCBM math and VOC. The highest number of students achieving the benchmark was the OAKS mathematics assessment.

Research Question 1: Relationship Between easyCBM Fall Assessments and OAKS Math

The first research question addressed the relationship between student results on the easyCBM assessments and OAKS mathematics assessment. Scatterplots (see Appendix B) prepared with each pair of data determine whether there is a possible linear relationship. The assumption of linearity was not markedly violated, and the five performance variables were normally distributed. Therefore, the study includes Pearson correlations to examine the intercorrelation of the variables.

The strongest positive correlation, which would be considered a very large effect size according to Cohen (1988), was between passage reading fluency and vocabulary, r (6162) =.73, p < .001. This means students who read a higher number of correct words per minute were more likely to have higher vocabulary scores. Passage reading fluency was also positively correlated with multiple choice reading comprehension (r =.63), multiple choice reading comprehension with vocabulary (r = .63), and easyCBM math with OAKS mathematics (r = .66), all larger than typical effect sizes (Cohen 1988). In addition, easyCBM math had a larger than typical effect according to Cohen (1988) on passage reading fluency (r=.51), vocabulary (r=.56), and multiple choice reading comprehension (r = .51). Positive correlations with larger than typical effect sizes between OAKS mathematics and passage reading fluency (r= .54) and vocabulary (r=.55) were also reported. The weakest correlation in the data, between multiple-choice reading comprehension and OAKS mathematics (r = .47), is considered a medium effect size according to Cohen (1988). Table 11 shows the correlation between performance variables.

Table 11

Correlation of easyCBM Fall Reading and Math Measures with OAKS Mathematics

	OAKS math	PRF_fall	VOC-Fall	MCRC_fall	Math_fall
OAKSmath	1.0	.541	.55	.47	.66
PRF_fall	--	1.0	.73	.63	.51
VOC_fall	--	--	1.0	.63	.56
MCRC_fall	--	--	--	1.0	.51
Math_fall	--	--	--	--	1.0

Research Question 2: Predictive Ability of Fall easyCBM Assessments for OAKS Mathematics Performance

The study includes regression analysis to investigate the predictive relation of fall easyCBM mathematics and reading performance and spring performance on OAKS mathematics. This analysis provides further explanation, about the proportion of variance in state test scores explained by the different easyCBM measures. The Pearson-product correlation coefficient (R) provides information on the predictor variables ability to predict correctly the criterion variable. The range for R is zero to one, zero means there is no linear relationship and one means the predictor variable accurately predicts the criterion variable in all cases.

To predict the OAKS-mathematics scores, two regression analyses follow. One analysis included the three fall reading CBMs (PRF, VOC, MCRC) as predictors, while the second analysis included fall math CBM as the predictor variable. The regression of OAKS on reading CBMs was significant R^2 =.353, F (4, 6157) = 1119.32, $p < .001$.

The four easyCBM assessments of math, passage reading fluency, vocabulary and multiple choice reading comprehension were included in a multiple regression analysis with OAKS- Mathematics. The linear combination of the four easyCBM measures was statistically significant and one or more of the variables significantly predicted OAKS - math. ANOVA results F (4, 6159) = 1547.79, $p < .001$ indicated that the combined reading and math measures contributed to variance in OAKS math results. The R value was .708 and R^2 =.501 indicating that 50.1 % of the variance in OAKS math scores could be explained by the combination of all four CBMs. Table 12 shows results from the multiple regression analysis with OAKS- math as the constant and the four easyCBM assessments as the predictor variables. The standardized coefficients indicated that Math (β = .483) is more predictive than the reading CBMs.

Table 12

Regression of OAKS – Mathematics on Combined Fall easyCBMs

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	Std. Error	Beta	*t*	*p*
(Constant)	180.80	.479		377.48	.000
PRF	.05	.004	.18	12.92	.000
VOC	.30	.04	.12	8.56	.000
MCRC	.09	.04	.03	2.56	.010
Math	.82	.02	.48	42.825	.000

Results of OAKS math alone regressed on fall easyCBM math were also statistically significant. ANOVA results, F (1, 6162) = 4819.06, $p < .001$, indicated fall math CBM contributed to variance in OAKS math results. The R value was .662 and R^2 =.439, indicating 43.9% of the variance in OAKS math scores could be explained by easyCBM fall math scores. Table 13 shows the results of the regression on fall easyCBM math.

Table 13

Regression of OAKS Mathematics on Fall easyCBM Math

Model	Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	*t*	*p*	95% CI *LL*	*UL*
(Constant)	181.5	.49		371.27	<.000	180.5	182.46
Math_fall	1.12	.02	.541	69.42	<.000	1.09	1.15

The regression of OAKS on reading CBMs was significant R^2 =.353, F (4, 6157) = 1119.32, *p* < .001. Results indicated that one or more of the reading CBMs contributed to variance in OAKS math results with 35.3% of the variance in OAKS-mathematics scores explained by the combined reading measures. Table 14 shows results of the regression of the easyCBM reading measures on the constant OAKS- mathematics. The standardized coefficients indicate that vocabulary (β = .28) is slightly more predictive than passage reading fluency (β = .25).

Table 15 shows partial correlations associated with four predictor variables. The semi-partials indicated that the fall easyCBM math (.48) accounted for more of the variance than the other variables. Squaring the semi-partial correlation coefficient reveals that easyCBM math uniquely accounted for 23% of the variance in OAKS Mathematics performance.

Table 14

Regression of OAKS- Mathematics on Fall easyCBM Combined Reading

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	*t*	*p*
(Constant)	194.71	.40		485.38	.000
PRF	.07	.004	.25	15.61	.000
VOC	.67	.04	.28	17.42	.000
MCRC	.40	.04	.14	10.07	.000

Table 15

Part Correlations: OAKS-Mathematics on Fall CBMs

Model	Correlations	
	Zero- Order	Part
PRF	.54	.12
VOC	.55	.08
MCRC	.47	.02
Math	.66	.39

The differences between the three models (all four CBMs, math CBM and three reading CBMs) can be determined by analyzing R^2 and R^2 change values. The fall math assessment (R^2 change = .44) is greater than the combined reading CBM (R^2 change = .35). The combined measures (R^2 change = .50) indicates 50% of the variance in OAKS – Mathematics can be accounted for with the four easyCBM measures. Therefore, there is a 9% increase in predictive ability when math is added to the combined reading CBMs and a 6% increase when the reading CBMs are added to the math CBM.

Research Question 3: Consistency of Fall easyCBM Prediction of OAKS Mathematics

Many districts use CBM results to establish categories of risk to identify students who need additional intervention to meet state benchmarks. The scores can also provide information on current practice and changes needed to instruction. The current study examines the consistency of fall easyCBM's ability to predict student achievement on the OAKS benchmarks.

The study provides a ROC analysis (see appendix C) to determine if the easyCBM measures would consistently predict which students would meet the state mathematics benchmarks on OAKS. Students who score 212 or higher meet the third-grade state benchmark in math on OAKS. Therefore, OAKS data was recoded so values 212 and above were coded as zero, "not at risk" of meeting the state standards. Student scores 211 or below were coded as one, "at risk" to not meet the standard. The ROC analysis was designed so larger test values indicated stronger evidence for a positive actual state of zero. This meant the higher the easyCBM assessment score, the higher the likelihood of a high score on the OAKS. Results for all the easyCBM measures were consistent. Results indicate that 4,143 students had an accurate prediction of "not at risk" or "at risk," while 2,021 students had an inaccurate classification of "at risk" or "not at risk."

To calculate the consistency with which easyCBM predicted accurately that a student was "not at risk" to meet the benchmark on OAKS, an Area Under the Curve

(AUC) was calculated with a nonparametric assumption. A diagonal line on the graph with a slope of .5 indicates that the predictive value of the easyCBM scores on OAKS would be no greater than chance. A statistic of 1.0 for the AUC indicates a perfect prediction of "not at risk" and "at risk." The fall easyCBM assessments all had AUC >.5, results that contradicted the null hypothesis AUC =.5. Therefore, they all had some predictive ability.

The fall easyCBM math assessment (AUC =.83 with standard error of .005, 95% CI [.823-.843]) had a notably strong predictive ability to identify students who were "not at risk." The results in Table 16 illustrate a fair to good predictive ability of the Fall easyCBM assessments for identifying students who will make benchmarks on OAKS-Mathematics in the spring.

Table 16

Area Under the Curve to Predict OAKS Mathematics Outcomes

				Asymptotic 95% CI	
Test Result Variable	Area	Std. Error	Asymptotic Sig.	*LL*	*UL*
Fall easyCBM PRF	.770	.006	.000	.757	.782
Fall easyCBM VOC	.800	.006	.000	.789	.812
Fall easyCBM MCRC	.748	.006	.000	.736	.761
Fall easyCBM math	.833	.005	.000	.823	.843

The ROC analysis provides information about the sensitivity and specificity related to the easyCBM assessments in relation to the OAKS assessment results. Sensitivity indicates the true positive rate (i.e., students meeting the OAKS benchmark who were identified as "not at risk" by their easyCBM scores) among all of the positives indicated within the analysis. Specificity indicates the true negative rate. In this case, the students classified as “at risk” who did not meet the state benchmark.

The 3rd grade easyCBM Math cut score used in the study as a possible predictor for students’ ability or lack of ability to meet standards for the OAKS math assessment was 31. This score is the 50th percentile score on easyCBM math. However, in the Coordinates of the Curve shown in this analysis (see Appendix D) the value of 30.5, which would translate to a cut score of 31 for practical purposes, showed sensitivity of .61 and specificity of .88 (1-specificity = .116). Therefore, if using a cut score of 31 (representing those below the 50th percentile) only 61 out of every 100 students who met the OAKS benchmark would have been identified accurately as not being at risk.

CHAPTER IV

DISCUSSION

This study investigated the relationship between CBM - math and reading and OAKS - mathematics to (a) determine the association between easyCBM fall math and reading performance and student performance on OAKS math, (b) investigate the predictive relations between fall CBM assessments individually and combined on student OAKS mathematics performance, and (c) examine the consistency with which Fall easyCBM predicts which students may or may not attain the OAKS mathematics benchmark. In this section, I address findings and limitations, interpret study results, and outline implications and areas for further research.

Main Findings

There was a statistically significant positive correlation between OAKS math performance and student fall CBM performance. The correlation between students' OAKS math performance and fall CBM math, PRF, and VOC measures was larger than typical according to Cohen's guidelines (1988), with math showing the largest effect size followed by VOC. The positive correlation means that, in general, students who scored higher on the reading and math CBMs, tended to score higher on the OAKS mathematics. As expected, the strongest positive correlation occurred between CBM reading measures VOC and PRF.

In analyses of the predictive relationship of CBM on student performance on OAKS math, I found CBM math a stronger predictor than the reading CBM measures individually and in combination. The CBM math scores predicted greater variance in OAKS math scores. The addition of the combined fall reading CBM scores increased the predictive power of outcomes on the OAKS mathematics. In examining the individual fall reading measure, I found VOC to be the best predictor of OAKS mathematics outcomes. Therefore, the addition of reading CBMs influences the variance in OAKS math scores, suggesting that a reading construct is a component of student OAKS math performance. In my interpretations, I note that these findings both confirm the construct validity of the math measures as well as introduce the possibility of construct irrelevant variance that may be important to counter, particularly for students with reading (or word meaning which I interpret as vocabulary skills) deficits.

In the investigation of the Area Under the Curve (AUC) easyCBM math was sensitive in identifying students who were “at risk” as well as “not at risk” in meeting OAKS benchmarks. Reading measures of VOC and PRF were also credible indicators of students’ risk levels in predicting OAKS math performance. This result supports the previous findings using regression analysis with the reading construct and the relation with the OAKS mathematics. CBMs showed both sensitivity (positive predictive power) and specificity (negative predictive power).

In all three analyses, vocabulary was the strongest reading skill associated with student OAKS math performance which is consistent with vocabulary being identified as a significant variable in the OAKS mathematics as part of the test blue print. The addition of vocabulary demonstrates construct irrelevant variance as a reading skill within the math assessment.

Limitations

As the study used an extant data set, a number of limitations were present, many of which could not be controlled, including (a) mortality, (b) grade level and location, (c) standardization and accommodations, (d) curriculum and instruction, and (e) differences in the assessments.

Mortality. Originally, data from 16,207 students were obtained. Analyses of the initial data set revealed a significant portion of the initial students (61.5%, n= 9,961) in the data set, had not completed one or more of the assessments. To be included in the current study's data set, students had to complete all five assessments within the school year. This included the (a) one-to-one Passage Reading Fluency measure, (b) computer-based Vocabulary CBM, (c) computer-based Multiple Choice Reading Comprehension CBM, and (d) computer-based Math CBM all administered in the fall. In addition, students had to complete the Oregon Assessment of Knowledge and Skills - Mathematics in the spring. The reasons students did not participate are unknown. Also, note that an *a priori* decision resulted in the deletion of students who scored a "0" or appeared to be outliers on any assessment.

Analyses of both the initial data set and restructured study data set resulted in slightly higher results for students who completed all five assessments. Students who completed all five assessments read approximately three words more per minute in the fall than the original participants did. Mean scores on the other four assessments range from .24 to .39 points higher than the initial data. The distribution of scores for all the measurement variables reflected a normal curve (See Appendices A and B) except the fall vocabulary CBM for the initial data set. The fall reading measure was significantly skewed to the right (1.17) in the initial data set, yet skewed to the left (-.77) for the restructured data. The restructured data set more closely resembled a normal curve for fall easyCBM VOC.

Grade level and location. This study included only Oregon students enrolled in third grade. The sampling plan was used to decrease potential confounds that may arise when using data from multiple grade levels. While providing control, using a single grade level of students reduced the generalizability to other grade levels. In addition, a single state summative assessment was used to decrease potential confounds that could arise with multiple statewide summative assessments. Therefore, the results are only generalizable to third grade students within the state of Oregon. Although this limitation restricts the external validity, it creates opportunity for future research across grade levels and in other states with summative assessments.

Standardization and accommodations. Training is required to administer all the assessments in this study, but there is no proof of adequate training of assessors. EasyCBM provides training materials and the state requires teachers to read the administration manual and attend a presentation on administering OAKS. Teachers administering OAKS are required to sign a fidelity to the test agreement. Educators are generally, however, in isolation with the students during test administration so documentation of consistency in administration is not possible. Therefore, the assumption is made that all personnel administering the assessments abide by the rules. The PRF CBM requires assessors to listen to students read passages in a one-to-one administration and to record correctly their performance. In contrast, the computer-based CBMs (VOC, MCRC, and Math) and OAKS administration are in a group setting, making environmental factors likely to be different across classrooms, schools, and districts. The difference in settings, nevertheless, may be cause for some differences in the test administration, which would in turn influence students' performance.

Furthermore, adjustments of the environment are allowable accommodations for students taking the OAKS. These changes can be in the presentation provided to students, their setting, time, and the responses they make. Though the state has a clear list of allowable accommodations, no documentation was available to determine which students received which types of accommodations. Although the implementation of accommodations is to help students access the skill assessed, there may be some variance accounted for in the math assessment due to these accommodations. In particular, the read aloud and computer-based reading accommodations may differ depending on the assessor's fidelity to the assessment guidelines.

Curriculum and instruction. The lack of education program information within the extant data used for the analyses added an additional limitation to the study. The study does not account for differences in instructional approach, curriculum materials, teacher experience and credentials, or time/intensity of instruction. Although all participants were third grade students, different school districts adopt their own curriculum materials. Teachers highlight the areas they feel best meet the state standards. This allows variance in the interpretation of the standards and emphasis on each skill.

Teachers using CBM benchmark assessments identify students' current level of mathematics or reading performance. Students identified below benchmark may be placed in intervention courses or receive a different curriculum than their peers. The instructional changes may lead to "better than expected" growth, which would reduce the power of the fall CBM to predict spring statewide test.

Information on teacher background and practice also are not part of the analysis. To be highly qualified, teachers meet specific criteria. Teachers at the third grade level teach multiple subjects and may not devote the same attention to mathematical or reading pedagogy as other content areas. Teachers also attend professional development in areas of interest and need. Therefore, the study does not account for differences in training and expertise, which can affect student achievement.

Finally, although the state of Oregon has guidelines requiring students to attend school, varying attendance, changes in teacher (substitutes, mid-year resignations, etc.), and class schedules affect time of instruction. In a study of student attendance on academic achievement in Pennsylvania, Gottfried (2009) found a strong correlation between time in class and statewide summative assessments.

Differences in assessments. Significant differences are present in the two types of measures used in this study. OAKS is an annual assessment with two opportunities to obtain the highest score, which is the one used for Adequate Yearly Progress (now referred to as Annual Measurable Objective). Students take **easy**CBM benchmark measures three times per year (fall, winter, spring) and easyCBM progress-monitoring assessments throughout the year. Because OAKS is high stakes for districts, it may receive more attention than easyCBM.

OAKS is a summative assessment designed to assess student learning annually at the end of an academic year and use the results for educational accountability purposes. It is a computer-based adaptive assessment so individual students receive different questions based on their level of performance.

In contrast, the purpose of easyCBM measures is the assessment of individual students' mastery of knowledge and skills in specific content areas. Students take the same grade-level benchmark assessment as their peers. There is also an option of progress monitoring with alternate forms designed to be of equivalent difficulty to monitor learning gains throughout the year. Students are accustomed to the assessments, and some students practice them regularly.

Interpretations

My study extended information provided by previous research regarding (a) validity associated with math CBM, (b) the predictive relationship between math and reading CBMs and standardized summative math assessments, and (c) the use of CBM in the identification of students at risk for making adequate progress in mathematics.

Validity evidence for math CBMs. Results of this study add to the existing body of research on the technical adequacy of CBMs by showing a moderately strong association between easyCBM math and a statewide summative math assessment. My findings fit well with prior research on math CBMs summarized by Foegen et al. (2007) who found relations between CBM math measures and other criterion variables in the same range as I found. In both my study and her analysis of previous studies, math CBM correlates to summative assessments supporting the construct validity of math CBM. The correlation strength may improve with consideration to construct irrelevant variance (removing written language).

Research by Shapiro et al. (2006) also reports a statistically significant correlation between CBM and the statewide summative assessment (Pennsylvania System for School Assessment, [PSSA]). In their research, they report CBM as two separate scores: math computation and concepts/applications. Their study of third grade students (n=380) from two districts in Pennsylvania found the strength of the correlation varied with time of year. Analyses of spring and winter CBM results indicted a moderate range of correlation (computation r=.52-.53; concepts/application (r = .61-.64) while fall CBM had somewhat lower correlations (r= .40 to .46).

Fall easyCBM math had a stronger correlation to OAKS Mathematics than the CBM measures used in Shapiro et al.'s study (2006). Winter and Spring benchmark screeners were not analyzed in my study. The representation of constructs within the CBM may explain the difference in strength of the fall correlation in relationship to winter and spring. In the fall, students may not have exposure to the skills assessed. The instructional sequence may vary between classrooms. In addition, students identified as below benchmark in the fall may receive additional instruction to boost their math and reading performance. Therefore, boost in performance by this group of students would increase correlation.

Predictive relationship of CBM on summative assessment. In this study CBM math and reading both had positive predictive relationships with the math statewide assessment. Together math and reading CBM explained 50% of the variance on the OAKS mathematics. This finding supports previous research. In their discussion of basic computation CBM, Jiban and Deno (2007) reflected on previous studies of basic facts measures with moderate correlations to standardized math achievement tests, pondering their lack of use by educators. This led to research using multiple CBMs to predict math performance. Jiban and Deno (2007) studied third grade students (n=38) in Minnesota using three CBM measures (basic math facts, cloze math facts and maze reading) and their predictive relationship to the Minnesota Comprehensive Assessment in Mathematics. Basic math facts and cloze math facts were computational only (no written language) with students identifying the missing numbers after (basic) and before (cloze) the equal sign. The Maze reading assessment required students to read silently for one minute and circle correct word choices each time they came to a missing word. This measure is similar to the VOC CBM, as students are choosing proper words to develop coherent and meaningful sentences. The measure also assesses fluency as the more a student reads the higher the opportunities to respond.

In the end, Jiban and Deno (2007) found math and reading CBM combined accounted for 52% of the variance on a statewide math test at fifth grade, a level that is very similar to what I report in the current study. The researchers reported data on reading performance, particularly silent reading, may improve prediction of performance on standardized math assessments. They suggested examining the concepts of numeracy and literacy separately allows examination of math without text and gives insight into both constructs as they jointly contribute to prediction of math performance.

In my study, reading measures VOC and PRF both are strongly correlated to OAKS math performance. They also increase predictive ability of fall CBM on OAKS Mathematics. It is important to note that correlation and predictive power do not mean causation. Therefore, other constructs, not measured, may share some variance captured by the reading measures. Possibilities could include processing speed, memory or ability to maintain attention. Information on student cognitive abilities or learning styles was not available within the extant data set used in this study.

The positive correlation of reading and math CBM to OAKS mathematics may also inform the use of accommodations on future assessments and reduce construct irrelevant variance. Helwig &Tindal (2003) found teachers did not correctly identify students for accommodations. Identifying skill deficits may help determine appropriate accommodations. Ketterlin-Geller et al. (2007) suggest read aloud accommodations for math assessments to mediate poor silent reading comprehension, or simplified language to reduce the linguistic demands of word problems. Students who score below benchmark in reading skills assessed may benefit from one of these accommodations. If a read aloud accommodation is used, listening comprehension can become another skill adding to the variance of the assessment outcome.

Identification of students at risk. Results of this study found CBM math to be a strong predictor of student identification of "not at risk" and "at risk" of meeting the statewide achievement levels. This supports Shapiro et al.'s (2006) study in Pennsylvania, which found classification rates between 66% to 85%, and specificity and sensitivity of .6 or .7 in math. Both studies' results demonstrate a strong positive correlation between CBM and statewide math summative assessments. This finding allows educators to infer that students who are progressing at established growth levels on CBM are on target for meeting end of the year achievement levels. Therefore, student scores on CBM can inform student risk levels for making adequate yearly progress.

While results of my study indicate that fall CBM math can consistently predict students who were on target to meet the OAKS benchmark, it also had a high number (n = 2021) of students with an inaccurate risk classification. This may lead to students receiving intervention who are not at risk. With limited resources, teachers, schools and districts need to target students with the greatest level of need. Therefore, a cut score of 31 based on the 50th percentile may be too high. Using decision-making rules set forth by Silberglitt and Hintze (2005) for establishing cut scores, I determined the optimal sensitivity (.79) and specificity (.70) provided within the coordinates of the curve (see appendix D) in the ROC analysis was associated with a cut score of 27.5 (or 28 for practical purposes). A cut score of 28, rather than 31, would be a more consistent predictor of students achieving the OAKS benchmark.

Implications for Practice

States and their school districts are accountable for student performance based on statewide assessment scores (NCLB, 2002). Therefore, schools and districts need meaningful data to predict student performance and adjust instruction to support student growth and performance. It is essential for these assessments to lead to valid inferences of the addressed construct.

Although the statewide test is a summative measure, the addition of CBMs to measure student progress is a means to track student progress throughout the year. Results from studies of CBMs and their prediction of student achievement in reading demonstrate their effectiveness as screening measures for identifying students who need additional instruction and intervention to meet state benchmarks levels (McGlinchey & Hixson, 2004; Nese et al., 2011; Silberglitt et al., 2006; Stage & Jacobson, 2001). Findings from this study indicate a positive prediction of math benchmark levels with easyCBM math on a state summative math assessment. The addition of reading CBM scores increased the ability to predict student math performance. This study supports further research of the use of math and reading CBM to identify students for intervention to make grade-level state benchmarks.

The ability to read is necessary to access math content and perform tasks in math (Crawford et al., 2001). Helwig and Tindal (2003) suggest math and reading screening when combined with teacher knowledge of their students, may help identify students who need reading accommodations on math assessments. The use of easyCBM PRF, VOC, and MCRC to identify students' reading difficulties could help identify students for possible read aloud accommodations within the math classroom and on summative assessments. Increasing students' reading proficiency or access to the written language within math tests containing word problems and multiple-choice questions is essential to decrease construct irrelevant variance. OAKS mathematics requires substantial reading skill, including assessed vocabulary, while easyCBM math does not. The difference in vocabulary assessed within OAKS and CBM math may affect the strength of their association. Whether reading skill is part of math competence or viewed as construct irrelevant variance in the context of math assessment, it appears the difference in linguistic load of easyCBM math and OAKS mathematics may shape student performance.

The positive correlation of reading and math CBM to OAKS mathematics may inform the use of accommodations on further assessments and reduce construct irrelevant variance. Helwig &Tindal (2003) found teacher often did not correctly identify students for accommodations on assessments. Identifying skill deficits may help determine appropriate accommodations. Ketterlin-Geller et al. (2007) suggest read aloud accommodations for math assessments to mediate poor silent reading comprehension, or simplified language to reduce the linguistic demands of word problems. Students who score below benchmark in reading skills assessed may benefit from one of these accommodations. If a read aloud accommodation is used, listening comprehension can become another skill adding to the variance of the assessment outcome.

Areas for Future Research

With the increased expectations that all students will master grade-level math standards, it is important to create effective tools for measuring student progress. Throughout the year, short simple assessments are needed that can inform instructional practice. Math concepts build upon each other making it important to pinpoint student strength and weakness for intervention. Educators need accurate assessments that are specific to skills within the math construct and guide them toward an understanding of student thinking errors.

In monitoring student progress, it is also important to understand the overlapping areas of math and reading constructs to identify students who struggle in one construct due to their disability in another construct. The debate continues whether computational fluency or concepts/applications discover the most robust indicator of student proficiency in math. Perhaps both areas need separate assessments, in the same manner, as fluency, vocabulary, and comprehension are separate within the reading construct for CBM. Fuchs (2004) refers to this approach to math CBM as "robust indicators." This approach identifies effective measures with strong correlation to math proficiency criteria and does not account for specific curricula or standards. One suggestion by Foegan et al., (2007) in their summary of math CBM is the identification of "numeracy" as a focus and development of robust indicators or early numeracy assessments comparable to PRF in reading. This addresses issues of construct underrepresentation by refining the focus to just those measures determined to demonstrate strong predictive ability of student math performance. Construct irrelevant variance is also addressed by removing written language.

In the math CBM debate, educators often address computation and application/word problems separately. This leads to a question of the literacy component within math. Vocabulary's predictive ability on math performance needs further research. The role of language in "real-world" math is relevant, yet computational performance is an important element too. In math, the difference between an accurate or inaccurate answer is often a computational error within a multi-step problem. In "real world" applications such as cutting siding for a house or calculating interest on a loan, slight errors can lead to enormous costs.

In this study, fall CBM VOC was the strongest reading skill associated with student OAKS math performance. Nese et al. (2011) found VOC to be the strongest reading skill associated with student performance on OAKS- reading at grades four and five. This suggests the importance of vocabulary development (word meaning) in both reading and mathematics for student growth. Vocabulary CBM may be a strong indicator of students at risk. Therefore, additional CBM research on the role of vocabulary development and the identification of students at risk in both math and reading could lead to changes in instructional practice. Implicit instruction of vocabulary may be needed for both constructs.

Teachers deserve information on effective ways to intervene and build competence with students struggling in math. Students and educators need tools to monitor and increase student growth. As researchers begin to better identify students "not at risk" and "at risk" a better understanding of intervention design and instruction needs investigation. Researchers should continue to identify measures sensitive to student growth. An understanding of student growth and learning progression can further enhance CBM research. Furthermore, teachers' use of the data and changes to their instructional approach requires more investigation. We need continued research to refine CBM for screening (computation and/or application) and progress-monitoring students' math development, and evaluate changes to instructional programs to support improved student outcomes.

APPENDIX A

DISTRIBUTION OF EASYCBM AND OAKS WITH NORMAL CURVE FOR STUDENTS WHO COMPLETED ALL FIVE ASSESSMENTS

Figure 1

Distribution of Grade 3 Fall easyCBM Math Results

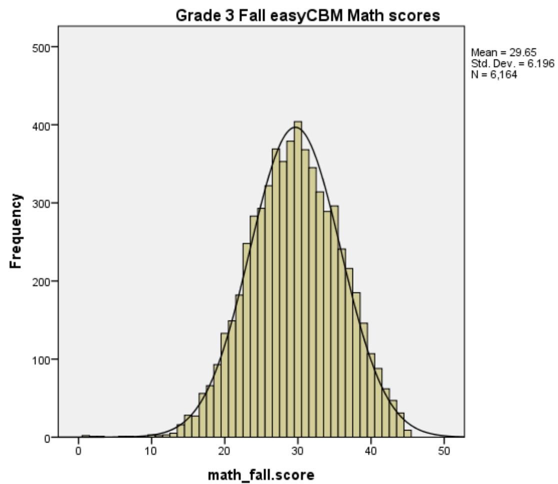


Figure 2

Distribution of Grade 3 Fall easyCBM Passage Reading Fluency Results

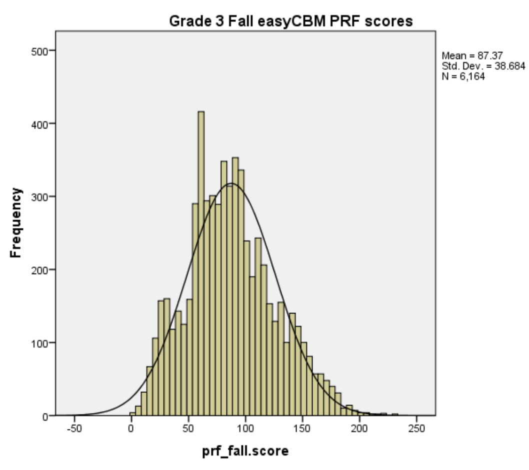


Figure 3

Distribution of Grade 3 Fall easyCBM Vocabulary Results

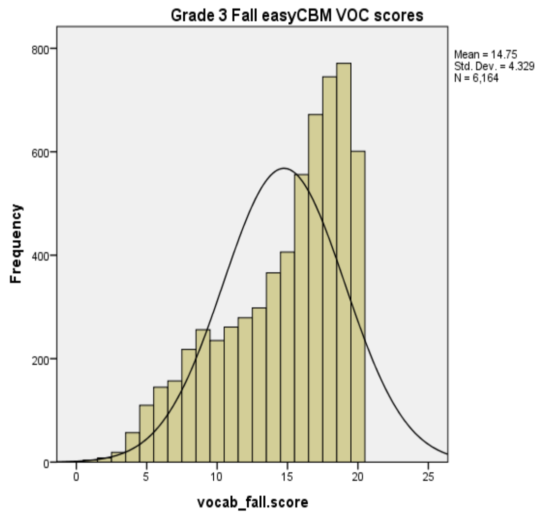


Figure 4

Distribution of Grade 3 Fall easyCBM MCRC Results

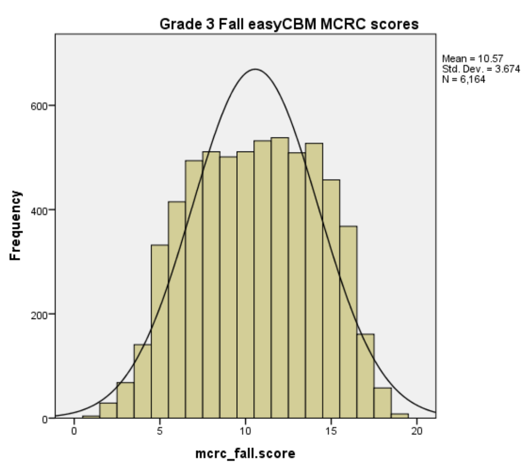


Figure 5

Distribution of Grade 3 OAKS Mathematics Results

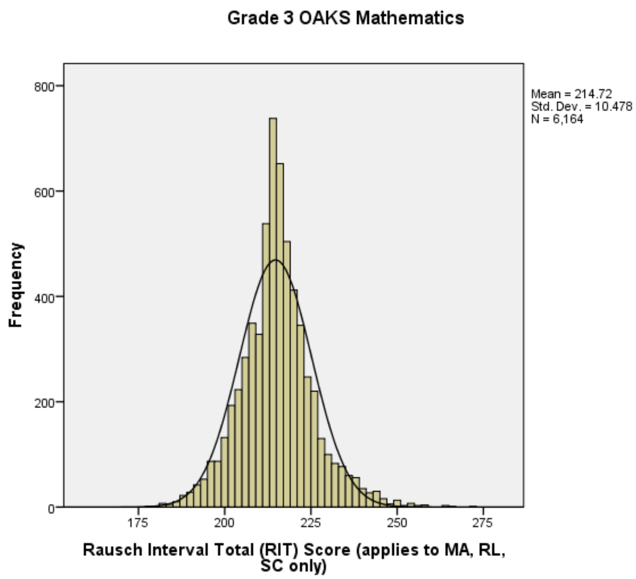


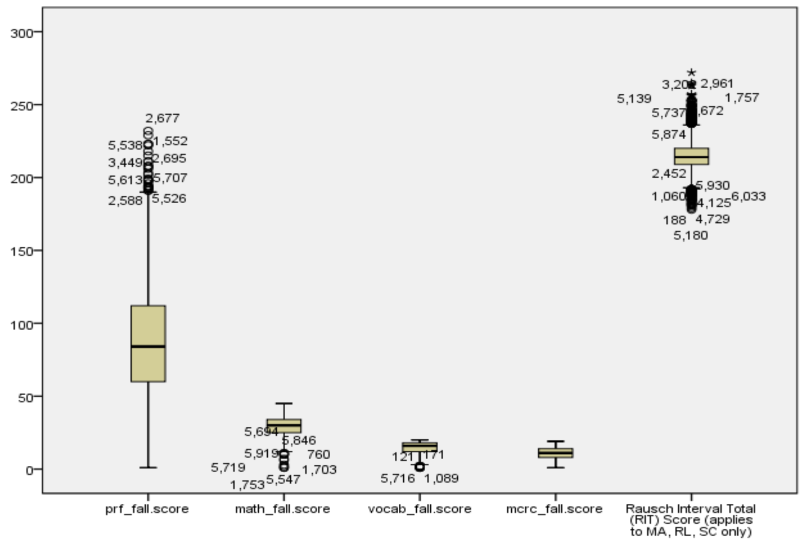
Table 17

Distribution of Restructured Grade 3 Data Set for Fall easyCBM and OAKS

	N	Mean	Skewness	
	Statistic	Statistic	Statistic	Std. Error
prf_fall.score	6164	87.37	.373	.031
vocab_fall.score	6164	14.75	-.766	.031
mcrc_fall.score	6164	10.57	-.065	.031
math_fall.score	6164	29.65	-.103	.031
Rausch Interval Total (RIT) Score (applies to MA, RL, SC only)	6164	214.72	.369	.031
Valid N (listwise)	6164			

Figure 6

Boxplot of Distribution in Fall easyCBM Assessments and OAKS Assessment Results



APPENDIX B

DESCRIPTIVE STATISTICS AND DISTRIBUTION OF INITIAL GRADE 3 DATA SET

Table 18

Descriptive Statistics of Initial Grade 3 Data Set for easyCBM and OAKS

	N	Minimum	Maximum	Mean	Std. Deviation
prf_fall.score	8611	0	273	84.54	39.715
vocab_fall.score	8472	0	100	14.51	4.767
mcrc_fall.score	9542	0	40	10.23	4.003
math_fall.score	13609	0	45	29.26	6.405
Rausch Interval Total (RIT) Score (applies to MA, RL, SC only)	14479	171	272	214.41	10.534
Valid N (listwise)	6234				

Table 19

Distribution of Grade 3 Fall easyCBM measures and OAKS Mathematics

	N	Skewness	
	Statistic	Statistic	Std. Error
prf_fall.score	8611	.338	.026
vocab_fall.score	8472	1.169	.027
mcrc_fall.score	9542	-.213	.025
math_fall.score	13609	-.286	.021
Rausch Interval Total (RIT) Score (applies to MA, RL, SC only)	14479	.249	.020
Valid N (listwise)	6234		

APPENDIX C

SCATTERPLOT OF EASYCBM AND OAKS, AND CORRELATION OF

ASSESSMENT RESULTS

Figure 7

Scatterplot of Grade 3 Fall easyCBM math with OAKS math

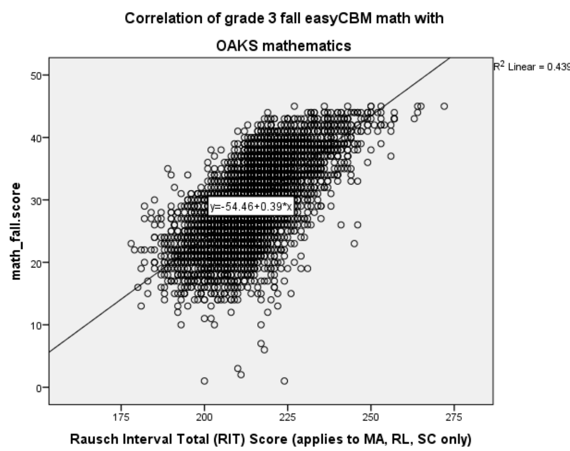


Figure 8

Scatterplot of Grade3 Fall easyCBM PRF with OAKS math

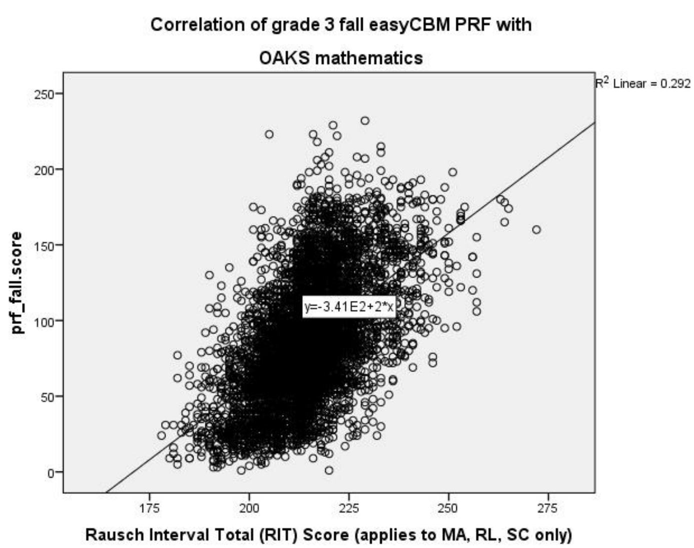


Figure 9

Scatterplot of Grade 3 Fall easyCBM Vocabulary with OAKS math

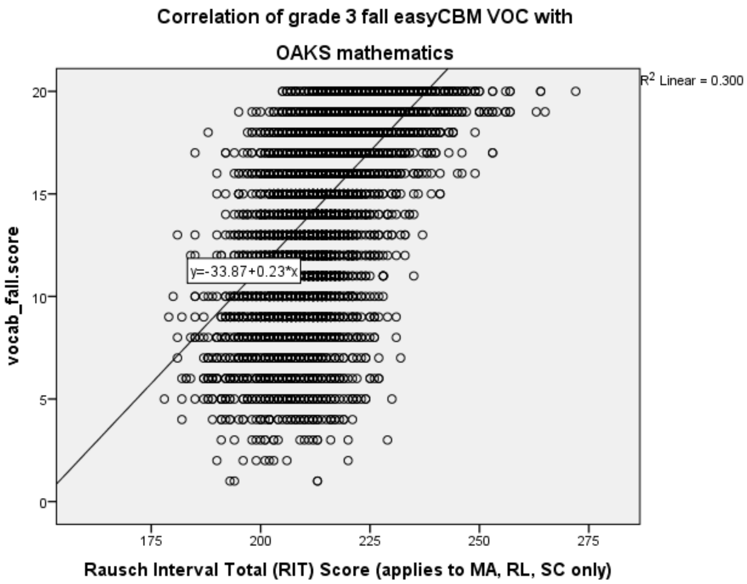
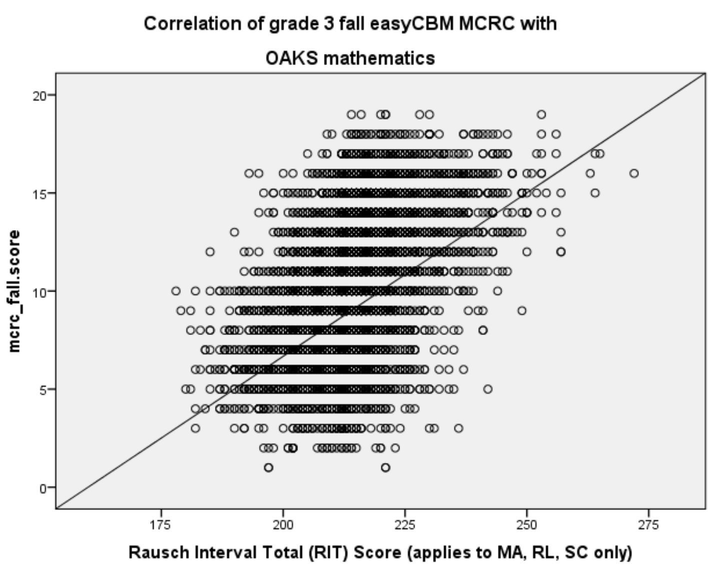


Figure 10

Scatterplot of Grade 3 Fall easyCBM Multiple Choice Reading Comprehension with OAKS math



APPENDIX D

ROC CURVES INCLUDING EASYCBM ASSESSMENT RESULTS AND OAKS MATH

Figure 11

ROC Curve Including Fall easyCBM Math and OAKS Math

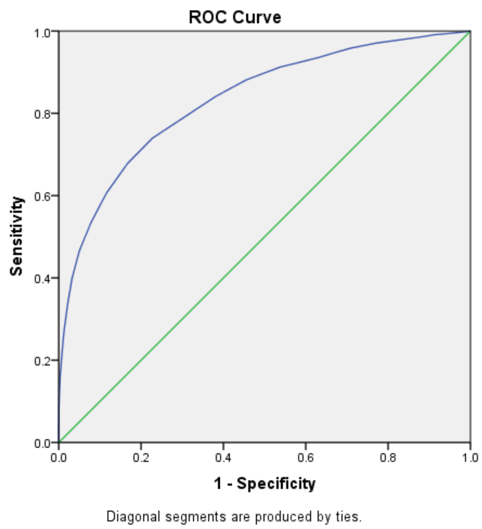


Figure 12

ROC Curve Including Fall easyCBM PRF and OAKS Math

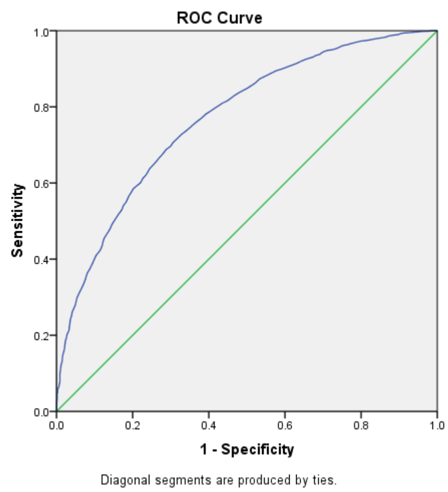


Figure 13

ROC Curve Including Fall easyCBM VOC
and OAKS Math

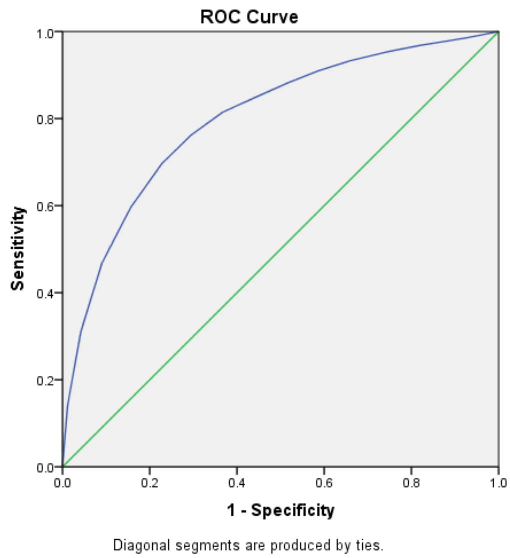


Figure 14

ROC Curve Including Fall easyCBM MCRC and OAKS Math

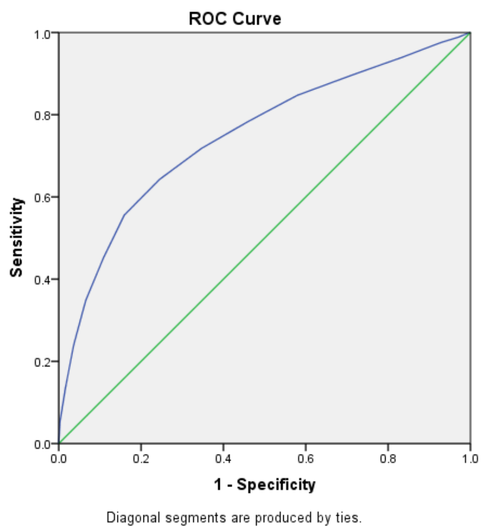


Table 20

Coordinates of the Curve with Test Result Variable Fall CBM Math

Coordinates of the Curve

Test Result Variable(s): math_fall.score

Positive if Greater Than or Equal To[a]	Sensitivity	1 - Specificity
.00	1.000	1.000
1.50	1.000	1.000
2.50	1.000	.999
4.50	1.000	.999
6.50	1.000	.999
8.50	.999	.999
10.50	.999	.998
11.50	.999	.997
12.50	.999	.996
13.50	.999	.994
14.50	.998	.987
15.50	.997	.975
16.50	.996	.964
17.50	.993	.941
18.50	.991	.913
19.50	.986	.879
20.50	.979	.827
21.50	.971	.770
22.50	.958	.706
23.50	.935	.629
24.50	.912	.537
25.50	.881	.455
26.50	.840	.380
27.50	.789	.302
28.50	.740	.228
29.50	.678	.167
30.50	.606	.116
31.50	.535	.079
32.50	.466	.050

33.50	.399	.032
34.50	.334	.021
35.50	.267	.012
36.50	.211	.007
37.50	.161	.003
38.50	.118	.001
39.50	.083	.000
40.50	.057	.000
41.50	.036	.000
42.50	.021	.000
43.50	.010	.000
44.50	.002	.000
46.00	.000	.000

The test result variable(s): math_fall.score has at least one tie between the positive actual state group and the negative actual state group.

a. The smallest cutoff value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values.

REFERENCES CITED

Alonzo, J., Anderson, D., Park, B. J., & Tindal, G. (2012). *The development of CBM vocabulary measures: Grade 3* (Technical Report No. 1210). Retrieved from Behavioral Research and Teaching website: http://www.brtprojects.org/publications/technical-reports

Alonzo, J., Ketterlin-Geller, L.R., & Tindal, G. (2006). *Curriculum-based measurement in reading and math: providing rigorous outcomes to support learning*. In L. Florian (Ed.), The Sage Handbook of Special Education (pp. 307-318). Thousand Oaks, CA: Sage

Alonzo, J., Lai, C. F., & Tindal, G. (2009). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 3* (Technical Report No. 0902). Retrieved from Behavioral Research and Teaching website: http://www.brtprojects.org/publications/technical-reports

Alonzo, J., Park, B.J., & Tindal, G. (2012). *The development of the easyCBM CCSS reading assessments: Grade 3* (Technical Report No. 1221).). Retrieved from Behavioral Research and Teaching website: http://www.brtprojects.org/publications/technical-reports

Alonzo, J., Park, B.J., & Tindal, G. (2013). *Examining the construct validity and internal structures of the easyCBM CCSS reading measures* (Technical Report No. 1227). Retrieved from Behavioral Research and Teaching website: http://www.brtprojects.org/publications/technical-reports

Alonzo, J., Park, B. J., Tindal, G. (2013). *An examination of the internal structures of the easyCBM CCSS reading measures.* (Technical report #1304). Retrieved from Behavioral Research and Teaching website: http://www.brtprojects.org/publications/technical-reports

Alonzo, J., & Tindal, G. (2007). *The development of word and passage reading fluency measures in a progress monitoring assessment system* (Technical Report No. 40). Retrieved from Behavioral Research and Teaching website: http://www.brtprojects.org/publications/technical-reports

Alonzo, J., & Tindal, G. (2008). *Examining the technical adequacy of fifth-grade reading comprehension measures in a progress monitoring assessment system* (Technical Report No. 0807). Retrieved from Behavioral Research and Teaching website: http://www.brtprojects.org/publications/technical-reports

Alonzo, J., & Tindal, G. (2012). *Teachers' manual for regular easyCBM: Getting the most out of the system.* Retrieved from http://www.easycbm.com/static/files/pdfs/info/easyCBM_Teachers_Manual.pdf

Alonzo, J., Tindal, G., Ulmer, K., & Glasgow, A. (2006). easyCBM online progress monitoring assessment system. Eugene, OR: Behavioral Research and Teaching. Retrieved from http://easycbm.com.

Anderson, D., Park, B. J., Irvin, P. S., Alonzo, J., & Tindal, G. (2011). Diagnostic efficiency of easyCBM reading: Washington state (Technical Report No. 1107). Retrieved from Behavioral Research and Teaching website: http://www.brtprojects.org/publications/technical-reports

Anderson, D., Lai, D.-F., Alonzo, J., & Tindal, G. (2011). Examining a grade-level math CBM designed for persistently low-performing students. *Educational Assessment*, 16, 15-34.

Anderson, D., Tindal, G. & Alonzo, J. (2009). Internal consistency of general outcome measures in grades 1-8 (Technical Report No. 0915). Retrieved from Behavioral Research and Teaching website: http://www.brtprojects.org/publications/technical-reports

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

Baker, S. K., Smolkowski, K., Katz, R., Fien, H., Seeley, 1. R., Kame'enui, E. J., & Thomas Beck, C. (2008). Reading fluency as a predictor of reading proficiency in low-performing, high-poverty schools. *School Psychology Review, 37,* 18-37

Bryant, D. P., Bryant, B.R., Gersten, R., Scammacca, N., & Chavez, M. (2008). Mathematics intervention for first-and second-grade students with mathematics difficulties: the effects of tier 2 intervention delivered as booster lessons. *Remedial and Special Education, 29*, 20-32.

Burns, M. Codding, R.S., Boice, C., Lukito, G. (2010). Meta-analysis of acquisition of fluency math interventions with instructional and frustration level skills: Evidence for a skill-by-treatment interaction. *School Psychology Review, 39,* 69-83

Chiappe, P. (2005). How reading research can inform mathematics difficulties: The search for the core deficit. *Journal of Learning Disabilities, 38 (4)*, 313.

Christ, T.J., & Ardoin, S.P. (2009) Curriculum-based measurement of oral reading: Passage equivalence and probe-set development. *Journal of School Psychology, 47,* 55-75

Christ, T.J. & Vining, O. (2006) Curriculum-based measurement procedures to develop multiple-skill mathematics computation probes: Evaluation of random and stratified stimulus set arrangements. *School Psychology Review, 35(3)*, 387-400

Clarke, B., Baker, S., Smolkowski, K., & Chard, D. J. (2008). An analysis of early numeracy curriculum-based measurement: Examining the role of growth in student outcomes. *Remedial & Special Education*, *29*, 46-57.

Clarke, B., Smolkowski, K., Baker, S., Fien, H., Doabler, C. & Chard, D. (2011) The impact of a comprehensive tier I core kindergarten program on the achievement of students at risk in mathematics. *The Elementary School Journal, 111(4),* 561-584.

Conley, D. T. (2007). *Policy analysis: Oregon student assessment systems.* Eugene, Oregon: Education Policy Improvement Center (EPIC).

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

Davis, S., Darling-Hammond, L., LaPointe, M., & Meyerson, D. (2005). *School leadership study: Developing successful principals* [Review of Research]. CA: Stanford University, Standard Educational Leadership Institute.

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

Deno, S.L. (1992) The nature and development of curriculum-based measurement. *Preventing School Failure, 36 (2),* 5

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

Deno, S., Fuchs, L.S., Marston, D. & Shinn, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30 (4),* 507-524

Deno, S., Marston, D., & Tindal, G. (1986). Direct and frequent curriculum-based measurement: An alternative for education for educational decision-making. *Special Services in the Schools, 2,* 5-27

easyCBM (2012). Curriculum based measurement for every tier [Assessment software]. Retrieved from http://www.easycbm.com

Ehri, L.C. (2005). Phases of development in learning to read words by sight. *Journal of Research in Reading, 18 (2),* 116-125

Elementary and Secondary Education (ESEA) Act of 1965, Pub. L. No. 89-10, § 79

Embretson, S.E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin*, *93*,179-197

Fletcher, J.M. (2005) Predicting math outcomes: Reading predictors and comorbidity. *Journal of Learning Disabilities, 38 (4),* 308-312

Foegen, A. Jiban, C. & Deno, S. (2007) Progress monitoring measures in mathematics. *The Journal of Special Education, 41,* 121-139

Fuchs, L. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review 33,* 188-192

Fuchs, L.S., & Deno, S.L. (1991) Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children, 58,* 232-243

Fuchs, L.S., & Fuchs, D. (2002). Mathematical problem solving profiles of students with mathematical disabilities with and without comorbid reading difficulties. *Journal of Learning Disabilities, 35,* 563-573

Fuchs, L., Fuchs, D., & Courey, S. (2005). Curriculum-based measurement of mathematics competence: From competence, to concepts and applications to real life problem solving. *Assessment for Effective Intervention, 30 (2),* 33-46. doi: 10.1177/073724770503000204

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

Fuchs, L., Fuchs, D., Compton, D., Bryant, J., Hamlett, C., & Seethaler, P. (2007). Mathematics screening and progress monitoring at first grade: Implications for responsiveness to intervention. *Exceptional Children, 73*(3), 311-330.

Fuchs, L.S., Fuchs, D. Eaton, S.B., Hamlett, C.L. & Karns, K.M. (2000). Supplementing teacher judgments of mathematics test accommodations with objective data sources. *School Psychology Review, 29,* 65-85

Fuchs, L.S., Fuchs, D., & Prentice, K. (2004) Responsiveness to mathematical problem-solving instruction: Comparing students at risk of mathematics disability with and without risk of reading disability. *Journal of Learning Disabilities, 37(4),* 293-306

Glanz, J. (2005). Action research as Instructional Supervision: Suggestions for principals. *National Association of Secondary School Principals Bulletin, 89*(643), 17-27.

Good, R.H., Kaminski, R.A., Smith, S. Laimon, D.& Dill, S. (2001). Dynamic Indicators of Basic Early literacy Skills (5^{th} Ed.) Eugene: University of Oregon.

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.

Grégoire, J., & Desoete, A. (2009). Mathematical disabilities – an underestimated topic?" *Journal of Psychoeducational Assessment, 27,* 171-174

Grimm, K. J. (2008). Longitudinal associations between reading and mathematics achievement. *Developmental Neuropsychology, 33(3),* 410-426.

Harlaar, N., Kovas, Y., Dale, P., Petrill, S., Plomin, R. (2012). Mathematics is differentially related to reading comprehension and word decoding: Evidence from a genetically sensitive design. *Journal of Educational Psychology, 104,* 622-635. doi:10.1037/a0027646

Hasbrouck, J.A. & Tindal G. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. Reading Teacher 59(7), 636-44

Haworth, C. A., Kovas, Y., Harlaar, N., Hayiou-Thomas, M. E., Petrill, S. A., Dale, P. S., & Plomin, R. (2009). Generalist genes and learning disabilities: A multivariate genetic analysis of low performance in reading, mathematics, language and general cognitive ability in a sample of 8000 12-year-old twins. *Journal of Child Psychology & Psychiatry*, *50*, 1318-1325. doi:10.1111/j.1469-7610.2009.02114.x

Helwig, R., Anderson, L. & Tindal, G. (2002). Using a concept-grounded curriculum-based measure in mathematics to predict statewide test scores for middle school students with LD. *Journal of Special Education*, *36*, 102.

Helwig, R., Rozek-Tedesco, M. A., & Tindal, G. (2002) An oral versus standard administration of a large-scale mathematics test. *Journal of Special Education, 36,* 39-47

Helwig, R., Rozek-Tedesco, M.A., Tindal, G. Heath, B. & Almond, P. (1999). Reading as an access to math problem solving on multiple-choice tests. *Journal of Educational Research, 93*, 113-125

Helwig, R., & Tindal, G. (2003) An experimental analysis of accommodation decisions on large-scale mathematics test. *Exceptional Children, 69,* 211-225

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

Hollenbeck, K., Rozek-Tedesco, M., & Tindal, G. (2000). The influence of reading and listening comprehension of a math read-aloud accommodation. Eugene, OR: Behavior Research and Teaching

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

Hosp, M. K. & Hosp, J. L. (2003). Curriculum-based measurement for reading, spelling, and math: How to do it and why. *Preventing School Failure 48*, 10-17

Interpreting the easyCBM Progess Monitoring Test Results (2013). Riverside Houghton –Mifflin. Retrieved from http://www.easycbm.com/static/files/pdfs/info/ProgMonScoreInterpretation.pdf

Item Response Theory and Rasch Models (2007). In Advanced Psychometric Approaches (Chapter 13), Retrieved from http://www.sagepub.com/upm-data/18480_Chapter_13.pdf

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

Jiban, C., & Deno, S. (2007). Using math and reading curriculum-based measures to predict state mathematics test performance: Are simple one-minute measures technically adequate? *Assessment for Effective Intervention, 32* 78-89

Jordan, N.C., Hanich, L.B., and Kaplan, D. (2003) A longitudinal study of mathematical competencies in children with specific mathematic difficulties versus children with comorbid mathematics and reading difficulties. *Child Development*, *74*, 834-850

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

Kane, M. T. (2001), Current Concerns in Validity Theory. *Journal of Educational Measurement*, 38 (4), 319–342.

Kane, M. (2002). Validating high-stakes testing programs. *Educational Measurement: Issues and Practices, 21,* 31-41. doi:10.1111/j.1745-3992.2002.tb00083.x/pdf

Kane, M. T. (2006) Current concerns in validity theory. *Journal of Education Measurement, 38 (4),* 319-342. doi: 10.1111/j.1745-3984.2001.tb01130x.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Education Measurement*, 50, 1-73

Ketterlin-Geller, L. R., Yovanoff, P., & Tindal, G. (2007). Developing a new paradigm for conducting research on accommodations in mathematics testing. *Exceptional Children, 73, (*3), 331-347.

Kiplinger, V.L. (2008). Reliability of large-scale assessment and accountability systems. In K. E. Ryan & L. A. Shepard, (Eds) *The future of test-based educational accountability*. (3-24). New York: Routledge.

Kovas, Y. Y., Haworth, C. A., Harlaar, N. N., Petrill, S. A., Dale, P. S., & Plomin, R. R. (2007). Overlap and specificity of genetic and environmental influences on mathematics and reading disability in 10-year-old twins. *Journal of Child Psychology & Psychiatry*, *48*, 914-922. doi:10.1111/j.1469-7610.2007.01748.x

Lai, C.F., Alonzo, J. & Tindal, G. (2013). EasyCBM reading criterion related validity evidence. (Technical Report No. 1310). Retrieved from Behavioral Research and Teaching website: http://www.brtprojects.org/publications/technical-reports

Lai, S.A. & Berkeley, S. (2012). High-stakes test accommodations research and practice. *Learning Disability Quarterly, 35(3),* 158-169

Lai, C.F., Irvin, S.P., Park, B.J., Alonzo, J. & Tindal, G. (2012). *Analyzing the reliability of the easyCBM reading comprehension measures: Grade 3*. (Technical Report no. 1202). Eugene, Oregon: Behavioral Research and Teaching.

Louis, K., Dretzke, B., & Wahlstrom, K. (2010). How does leadership affect student achievement? Results from a national US survey. *School Effectiveness & School Improvement*, *21*, 315-336.

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

Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice, 14(4),* 5-8. doi: 10.1111/j.1745-3992.1995.tb00881.x.

Messick, S. (1989). Validity. In R.L. Linn (Ed.), Educational measurement (3rd ed., pp. 13-103). New York: Macmillan.

National Center for Education Statistcs . Retrieved from http://nces.ed.gov/nationsreportcard/states/

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

National Center on Response to Intervention. (2010, March). *Essential components of RTI – a closer look at Response to Intervention*. Washington, DC: U.S. Department of Education, Office of Special Education Programs, National Center on Response to Intervention.

National Center on Response to Intervention. (2010, May). *Users guide to universal screening tools chart.* Washington, DC: U.S. Department of Education, Office of Special Education Programs, National Center on Response to Intervention.

National Council of Teachers of Mathematics. (2006). Curriculum focal points for pre-kindergarten through grade 8 mathematics: A quest for coherence. Reston, VA: Author. Retrieved from https://www2.bc.edu/solomon-friedberg/mt190/nctm-focal-points.pdf

National Mathematics Advisory Panel. (2008) *Foundations for Success: The Final report of the Mathematics Advisory Panel*. Washington DC: U.S. Department of Education.

Nese, J.F., Park, B. J., Alonzo, J. & Tindal, G. (2011).   Applied curriculum-based measurement as a predictor of high-stakes assessment: Implementations for researchers and teachers. *Elementary School Journal*, *3*, 608-624 doi: 10.1086/659034

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

O'Donnell, R. & White, G. (2005). Within the accountability era: Principals' instructional leadership behaviors and student achievement. *National Association of Secondary School Principals Bulletin*, *89 (643)*, 56-68.

Oregon Department of Education (2007). Reliability and Validity. (Technical Report No. 4) Retrieved from http://www.ode.state.or.us/teachlearn/testing/manuals/2007/asmttechmanualvol4_validity.pdf

Oregon Department of Education (2008). Background on the Oregon Statewide Assessment Program. Retrieved from http://www.ode.state.or.us/search/page/?id=451

Oregon Department of Education. (2010). Assessment scoring: Frequently asked questions about scoring statewide assessments Retrieved April 21, 2010, from http://www.ode.state.or.us/apps/faqs/index.aspx?=88

Oregon Department of Education. (2011a). How Oregon Set Achievement Standards in 2006-2007 Retrieved April, 29, 2011, from http://www.ode.state.or.us/search/page/?id=849

Oregon Department of Education. (2011b). Test Specifications and Blueprints Retrieved April, 29, 2011, from http://www.ode.state.or.us/search/page/?id=496

Oregon Department of Education (2011c). Oregon common core state standards for mathematics (CCSSM) Retrieved from http://www.ode.state.or.us/teachlearn/real/newspaper/newspaper_section.aspx?subjectcd=MA

Oregon Department of Education, Office of Learning and Special Partnerships. (2012). *Guidelines for the read aloud accommodation.* Retrieved from http://www.ode.state.or.us/teachlearn/testing/admin/alt/ea/2-guidelines-for-the-math-read-aloud-accommodation-for-2012-2013-(3).pdf

Oregon School Boards Association. (2005). Education Improvement in Oregon. Retrieved May 15, 2008, from http://www.osba.org/covered/curricul/edimprov.htm.

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40 (2),* 184-202

Park, B.J., Anderson, D., J., Lai, C.F., & Tindal (2012). An examination of test-retest, alternate form reliability, and generalizability theory study of the easyCBM reading assessments: Grade 3 (Technical Report No. 1218). Retrieved from Behavioral Research and Teaching website: http://www.brtprojects.org/publications/technical-reports

Park, B.J., Anderson, D., Irwin, P.S., Alonzo, J. & Tindal, G. (2011) *Diagnostic efficiency of easyCBM reading: Oregon* (Technical Report No. 1106) Retrieved from eric.ed.gov website: http://www.eric.ed.gov/PDFS/ED531920.pdf

Roehrig, A.D., Yaacov, P., Nettles, S., Hudson, R.F., & Torgeson, J.K. (2008).Accuracy of the DIBELS oral reading fluency measure for predicting third grade reading comprehension outcomes. *Journal of School Psychology, 14,* 343-366.

Shadish, W.R., Cook, T.D., Campbell, D.T. (2001). Statistical Conclusion Validity and Internal Validity In *Experimental and Quasi-experimental Designs for Generalized Causal Inference* (pp. 33-63). Boston: Houghton-Mifflin

Shapiro, E. S. Keller, M. A., Lutz, J.G., Santoro, L.E. & Hintze, J.M. (2006). Curriculum-based measures and performance on state assessment and standardized tests: Reading and math performance in Pennsylvania, *Journal of Psychoeducational Assessment, 24,* 19-35.

Silberglitt, B., Burns, M. K, Madyun, N. H., & Lail, K E. (2006). Relationship of reading fluency assessment data with state accountability test scores: A longitudinal comparison of grade levels. *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: 10.1177/07342829050230040

Stage, S.A. & Jacobson, M.D. (2001). Predicting student success on a state-mandated performance-based assessment using oral reading fluency. *School Psychology Review, 30 (3),* 407-419.

Streiner, D. (2003). Diagnosing tests: Using and misusing diagnostic and screening tests. *Journal of Personality Assessment 81*, *209*-219.

Stumbo, C., & McWalters, P. (2011). Measuring effectiveness: what will it take? *Educational Leadership, 68,* 10-15.

Swanson, H.L., & Jerman, O. (2006). Math disabilities: A selective meta-analysis of the literature. *Review of Educational Research, 76(2),* 249-274.

Thurber, R.S., Shinn, M.R., & Smolkowski, K. (2002). What is measured in mathematics tests? Construct validity of curriculum-based mathematics measures. *School Psychology Review, 31(4),* 498-513.

Tindal, G. (2013). Curriculum-based measurement: A brief history of nearly everything from the 1970s to the present. *ISRN Education,* 1-29. doi:10.1155/2013/958530

Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology, 40,* 7–26.

United States Department of Education (2102) *Glossary of Terms: No Child Left Behind.* Retrieved from http://www2.ed.gov/nclb/index/az/glossary.html?src=az#2

U.S. Department of Education, (2010) . *A blueprint for reform: The reauthorization of the elementary and secondary education act.* Retrieved from *http://www2.ed.gov/policy/elsec/leg/blueprint*.

U.S. Department of Education, (2011). *Race to the top fund*. Retrieved from *http://www2.ed.gov/programs/racetothetop/index.html*

U.S. Department of Education. (2011, September 23). Obama administration sets high bar for flexibility from no child left behind in order to advance equity and support reform. Retrieved from http://www.ed.gov/news/press-releases/obama-administration-sets-high-bar-flexibility-no-child-left-behind-order-advanc

Vukovic, R., Lesaux, N. & Siegal, L. (2010) The mathematic skills of children with reading difficulties. *Learning and Individual Differences, 20,* 639-643.

Vukovic, R. & Siegel, L. (2010) Academic and cognitive characteristics of persistent mathematic difficulty from first through fourth grade. *Learning Disabilities & Practice*, *25*, 25-38.

Wayman, M.M., Wallace, T., Wiley, H.I., Ticha, R., & Espin, C.A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education, 41(2)*, 85-120.

Weston, T.J. (2003). The validity of oral accommodations in testing: NAEP validity studies. (NCES-WP-2003-06), Washington, DC: National Center for Education Statistics

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

Wright, P.W.D. (2005). *U.S. Department of Education's commentary and explanation about proposed regulations for IDEA 2004*. Washington, DC: US Department of Education

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,* 4 - 12.