



2013

Examining Predictive Validity and Rates of Growth in Curriculum-Based Measurement with English Language Learners in the Intermediate Grades

Ella Farmer
Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_diss



Part of the [Educational Psychology Commons](#)

Recommended Citation

Farmer, Ella, "Examining Predictive Validity and Rates of Growth in Curriculum-Based Measurement with English Language Learners in the Intermediate Grades" (2013). *Dissertations*. 663.

https://ecommons.luc.edu/luc_diss/663

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Dissertations by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](#).
Copyright © 2013 Ella Farmer

LOYOLA UNIVERSITY CHICAGO

EXAMINING PREDICTIVE VALIDITY AND RATES OF GROWTH IN
CURRICULUM-BASED MEASUREMENT WITH ENGLISH LANGUAGE
LEARNERS IN THE INTERMEDIATE GRADES

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE GRADUATE SCHOOL
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

PROGRAM IN SCHOOL PSYCHOLOGY

BY

ELLA FARMER

CHICAGO, ILLINOIS

AUGUST, 2013

Copyright by Ella Farmer, 2013
All rights reserved.

ACKNOWLEDGEMENTS

First, I would like to thank Dr. Pamela Fenning, my dissertation chair, for her ongoing support and encouragement. There were times when I did not think that I would be able to finish this project, and she helped me to believe that I could do this. Dr. Fenning was extremely flexible and understanding my struggles to balance work, family, and my dissertation research.

I would also like to thank my other committee members, Dr. Terri Pigott and Dr. Diane Morrison. I am grateful to Dr. Pigott for her help and guidance with statistical and technical aspects of the current study, and for her support and encouragement throughout my time at Loyola. I would also like to thank Dr. Morrison for her support and encouragement during the dissertation process.

I would also like to thank several people for their help and support along the way at Loyola. Dr. Prasse encouraged me to apply to the doctoral program and I will always appreciate his encouragement and support. Thank you to Dr. Wynne and Dr. Golomb for providing opportunities to be involved in their research projects. I am also grateful to Dr. Kathy Pluymert and Dr. Laura Swanlund for their collaboration and support with an earlier research project that provided some of the framework for my dissertation research.

Thank you to Dr. Charlene Cobb the administrators of East Maine School District 63 for being willing to provide me with the data for the current study and supporting me through this project.

Finally, I would like to thank my family and friends for all of the support and encouragement they provided throughout this process. Thank you to my parents, for always believing in me and supporting me. Thank you to my husband, Dan, for giving me the time and space to work and for believing that I would finish even when I thought I might not. Thank you also to my children, Calum and Soren, for being understanding on all of those Saturdays when I was at the library.

For my children, Calum and Soren

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iii
LIST OF TABLES	viii
CHAPTER ONE: INTRODUCTION	1
English Language Learners	1
Federal Legislation and Education Reform	6
Curriculum-Based Measurement	9
Statement of the Problem	12
Research Questions	13
CHAPTER TWO: REVIEW OF THE LITERATURE	15
Literacy Development	16
Curriculum-Based Measurement	20
Curriculum-Based Measurement with English Language Learners	26
Summary	33
CHAPTER THREE: METHODOLOGY	35
Introduction	35
Subjects and Setting	36
Procedures	38
Measures/ Instrumentation	39
Data Analysis	44
Summary	46
CHAPTER FOUR: RESULTS	48
Descriptive Statistics	49
Research Question 1	56
Research Question 2	60
Research Question 3	66
Research Question 4	72
CHAPTER FIVE: DISCUSSION	75
Research Question 1	76
Research Question 2	78
Research Question 3	82
Research Question 4	86
Implications for Practice	87
Limitations	89
Implications for Future Research	91
Summary	92

REFERENCES

94

VITA

102

LIST OF TABLES

Table 1: Number of Students Assessed During the 2010-2011 School Year	36
Table 2: Overall Demographic Information for the School District	37
Table 3: Overall Demographic Data for the Current Sample	49
Table 4: Demographic Data for ELL Students	50
Table 5: Home Languages for ELL Students	52
Table 6: ACCESS for ELLs Scores	53
Table 7: Standardized Test Scores for Fourth Grade	54
Table 8: Standardized Test Scores for Fifth Grade	54
Table 9: Standardized Test Scores for Sixth Grade	55
Table 10: Fixed Effects for All Students	57
Table 11: Estimation of Variance Components for All Students	57
Table 12: Level Two All Students	59
Table 13: Fixed Effects for ELL Students	61
Table 14: Estimation of Variance Components for ELL Students	61
Table 15: Level Two for ELL Students	64
Table 16: Pearson Correlations	67
Table 17: Cross Tabulations	69
Table 18: Cross Tabulation Interpretation	70
Table 19: Diagnostic Accuracy	71

CHAPTER ONE

INTRODUCTION

As diversity in the United States increases, so does diversity in the public schools. Many students from diverse ethnic backgrounds, especially those who are Hispanic or Asian, speak a language other than English in the home (KewelRamani, Gilbertson, Fox, & Provasnik, 2007). Some of these students enter school proficient in English, but others enter school with limited English proficiency. Teachers and administrators are responsible for ensuring that all students meet standards for proficiency in reading and mathematics as measured by their performance on high stakes tests. This is a challenge for school staff. In order to formatively assess students and determine whether students are predicted to meet state standards, it is important that school personnel use assessment tools that are valid for English Language Learners (ELLs) as well as the overall population of students. This chapter will describe the current educational climate and provide a rationale for the current study.

English Language Learners

The population of the United States has changed significantly over the past thirty years, and is predicted to change even more over the next few decades (Federal Interagency Forum on Child and Family Statistics, 2011). Between the years of 2005 and 2020, the white population is expected to grow by four percent, while the minority population is predicted grow by 32 percent (KewelRamani, Gilbertson, Fox, &

Provasnik, 2007). The percentage of children identified as Hispanic has increased from nine percent in 1980 to 23 percent in 2010, and is predicted to increase to 39 percent by the year 2050 (Federal Interagency Forum on Child and Family Statistics, 2011). Also, the foreign born population of the United States has increased in the years from 1990 to 2005 (KewelRamani, Gilbertson, Fox, & Provasnik, 2007).

The demographic composition of American public schools has changed along with the general population. In 2009, 21% of school age children in the United States spoke a language other than English in the home (Federal Interagency Forum on Child and Family Statistics, 2011). This was a significant increase from 1980 when only 10% of school age children lived in homes where English was not the primary language (U.S. Department of Education, 2011). Some of these children enter school with adequate proficiency in English, while others develop English proficiency at school.

Approximately 5 percent of school age students speak another language at home and do not have proficiency in English (Federal Interagency Forum on Child and Family Statistics, 2011). This is an increase from 4 percent in 1980, although down from a high of 7 percent in 2000 (U.S. Department of Education, 2011). The number of ELL students in the public schools grew by 57.17% in the years from 1995-2006, while the total school population grew by only 3.66% (National Clearinghouse for English Language Acquisition, n.d.).

Students who come from homes where a primary language other than English is spoken are known as Language Minority (LM) students (Kieffer, 2008). Students who have not yet developed English proficiency are referred to as Limited English Proficient

(LEP) or English Language Learners (ELLs). In the present study, the phrase English Language Learners, or ELLs, will be the primary term used to refer to those students who speak another language at home and have yet to develop adequate proficiency in English. English Language Learner is the preferred by many in education because it does not define students based on a deficit (Keiffer, 2008). Limited English Proficient is the term used in federal reporting documents, and can be used interchangeably with English Language Learner. In the No Child Left Behind Act of 2001, the federal government defines a student who is Limited English Proficient as an individual:

- (A) who is aged 3 through 21;
- (B) who is enrolled or preparing to enroll in an elementary school or secondary school;
- (C)
 - (i) who was not born in the United States or whose native language is a language other than English;
 - (ii)
 - (I) who is a Native American or Alaska Native, or a native resident of the outlying areas; and
 - (II) who comes from an environment where a language other than English has had a significant impact on the individual's level of English language proficiency; or
 - (iii) who is migratory, whose native language is a language other than English, and who comes from an environment where a language other than English is dominant; and
- (D) whose difficulties in speaking, reading, writing, or understanding the English language may be sufficient to deny the individual —
 - (i) the ability to meet the State's proficient level of achievement on State assessments described in section 1111(b)(3);
 - (ii) the ability to successfully achieve in classrooms where the language of instruction is English; or

(iii) the opportunity to participate fully in society (§ 9101, 25).

English Language Learners have been identified as a group of students who are significantly at-risk for difficulties with developing early reading skills (Snow, Burns, & Griffin, 1998). Gaps are also present when comparing the performance of English Language Learners (ELLs) and non-ELLs through the elementary years. In 2007, 30 percent of fourth grade ELL students performed at the Basic level or above on a national test of reading, compared to 69 percent of non-ELL students. At the eighth grade level, 29 percent of ELL students performed at or above the Basic level, compared to 75 percent of non-ELLs (U.S Department of Education, 2007). While language minority children who enter school with adequate English proficiency show similar growth trajectories to their English-speaking peers in kindergarten through fifth grade, those who enter school with limited proficiency demonstrate lower rates of growth and fall further behind by the end of elementary school (Keiffer, 2008).

Based on data from the 2009 National Assessment of Educational Progress (Hemphill & Vanneman, 2011), reading scores for all fourth and eighth grade students have increased significantly since 1992. While scores for most student groups have increased since 1992, significant gaps persist between minority and majority groups. Gaps between Hispanic and white students have not changed significantly since 1992, although the Hispanic population in the schools has increased significantly in this time period (Hemphill & Vanneman, 2011). Hispanic students who are proficient in English achieve significantly higher scores in reading than their ELL peers in both fourth and eighth grade, and this gap has been consistent since 1998 (Hemphill & Vanneman, 2011).

As described below, many ELL students fall into this and other categories that are associated with increased risk of academic difficulties.

Students who speak another language at home also have higher rates of other risk factors than are seen in the general population. They are more likely to be poor, to be from racial or ethnic minority backgrounds, and to attend segregated schools. As of 2000, 71 percent of ELL students in elementary school were Hispanic, and 14 percent were Asian (Capps, Fix, Murray, Ost, Passel, & Herwantoro, 2005). Many students who speak another language at home are the children of immigrant parents. While one in five school children in the United States has immigrant parents, one in four children who participate in free or reduced lunch programs is the child of immigrants. Two thirds of ELL students are considered low income. Fifty-three percent of ELLs attend schools where over 30 percent of students have limited English proficiency. This is likely related to the fact that many of these students live in areas that are residentially segregated based on income, race, and ethnicity. Finally, six out of seven ELL students in elementary school live in linguistically isolated households, which means that there everyone in the household over the age of 14 has limited English proficiency (Capps, et al., 2005). English Language Learners also have a higher prevalence of negative adult outcomes in adolescence. Graduation rates for ELL students are significantly lower than non-ELL peers in some states that provide data to the federal department of education, but this data is not currently reported by all districts and states (Zehr, 2009).

Federal Legislation and Education Reform

No Child Left Behind

In 2002, the Elementary and Secondary Education Act was reauthorized and signed into law as the No Child Left Behind Act of 2001. This legislation was intended to address persistent problems in education in the United States. A significant segment of school children do not develop basic reading and mathematics skills, and high numbers of adolescents drop out of high school before graduating. In addition, a significant achievement gap has consistently been demonstrated when students from certain subgroups to the overall population. NCLB stresses the importance of all students developing adequate academic skills in the areas of reading and mathematics by 2014 (No Child Left Behind, 2002).

In order to assess student progress, schools and districts are mandated to assess all students using a standardized test that is linked to state educational standards. All students are tested in reading and mathematics each year in grades three through eight. States are responsible for setting yearly targets toward the goal of 100 percent of students achieving success by 2014. Schools and districts are evaluated by the state based on the performance of all students in each grade, and also the performance of specific subgroups. Subgroups include racial and ethnic groups, ELLs, low income students, and students receiving special education. Schools and districts that do not meet targets for Adequate Yearly Progress (AYP) may be responsible for a range of school improvement measures, such as offering school choice, providing supplementary tutoring for students,

or ultimately being subject to restructuring or school closure. Since the passage of NCLB, schools and districts have been accountable for ensuring that all of their students meet state standards, without regard for income, race, ethnicity, special education, or ELL status (No Child Left Behind, 2002). Public schools are now in a position where they are responsible for making sure that all students meet basic standards, and cannot allow at-risk populations to fail.

Title III of the No Child Left Behind Act also includes specific goals for identification and instruction of students who are English Language Learners, and is known as the English Language Acquisition, Language Enhancement, and Academic Achievement Act. The goals of Title III are that all English Language Learners, including immigrant children, meet standards for English language proficiency and meet grade level academic standards in English (§3102(1)). States are responsible for selecting assessments of English language proficiency, setting annual goals for growth in English proficiency, English proficiency attainment, and academic achievement in English (U.S. Department of Education, 2004).

Individuals with Disabilities Education Act

The federal Individuals with Disabilities Education Act of 2004 (IDEA, 2004) is an education law that provides funding and details the requirements for public schools that provide special education services for students with disabilities. This most recent reauthorization contains some changes that make the law consistent with No Child Left Behind and add provisions for early intervening services for students who are at-risk of

academic failure. Specifically, portions of special education funding received under IDEA are permitted to be used for school wide programs and activities under NCLB. In addition, in order to meet eligibility criteria for special education services, school staff must demonstrate that a lack of appropriate reading instruction, based on the definition under NCLB is not a determinant factor for that student's educational deficits. Lack of appropriate mathematics instruction and limited English proficiency must also be ruled out as determinant factors (IDEA, 2004).

Another significant change in the 2004 law involves changes in the definition of a Specific Learning Disability (SLD). Previously, states required that students demonstrated a discrepancy between their academic performance and their ability, generally measured by a standardized intelligence test. The present law mandates that states establish eligibility criteria that do not require an ability/achievement discrepancy. Instead, states must allow for a process that determines if a student's performance is discrepant from his peers or state educational standards, and how well a student responds to scientific, research-based intervention (IDEA, 2004). In other words, school teams have the option of providing students who are performing below academic standards with research-based intervention, and documenting their progress with these interventions as part of an evaluation to determine eligibility for special education.

Related to changes in both special education law and No Child Left Behind, many states and school districts have begun using systems that emphasize frequent screening of all students to identify those who are at risk of failure, and the use of targeted

interventions for those identified as at-risk. One specific problem-solving model is called Response to Intervention (Fuchs & Fuchs, 2006). This approach has been advocated as an alternative to the traditional approach to identification of students with disabilities (Deno, 2003; Fuchs, 1999). Traditionally, students were not able to receive intervention until their academic performance was well below grade level expectations. This is often described as a “wait to fail” model. Response to Intervention allows for the provision of supplementary services for students on an as-needed basis, without requiring extensive individual testing or special education eligibility. Response to Intervention also promises to alleviate problems with overrepresentation of specific at-risk groups in special education by providing educational opportunities for these students (National Center for Culturally Responsive Educational Systems (NCCRESt), 2005).

Curriculum-Based Measurement

One of the origins of the model of RTI as currently practiced is Deno’s data-based program modification model (Batsche, Elliott, Graden, Grimes, Kovaleski, Prasse, Reschly, Schrag, & Tilly 2005). As part of this model, formative assessments were developed that were sensitive to growth and allowed for direct measurement of academic skills (Batsche et al., 2005). These assessments, developed in the 1970’s by Stanley Deno and colleagues, were intended to be more authentic forms of assessment than the norm-referenced tests that were widely used at the time (Deno, 1985; Reschly, Busch, Betts, Deno, & Long, 2009). They are generally known as Curriculum-Based Measurement (CBM), and continue to be frequently used in universal screening and progress

monitoring of students who are receiving academic interventions. Curriculum Based Assessments, of which CBMs are one type, are taken directly from standard grade level curriculum, or closely matched to curricular expectations and are sensitive to small changes in performance due to instruction Curriculum-Based Measurement consists of timed standardized assessments that measure a specific skill and can be used repeatedly for ongoing measurement (Howe & Shinn, 2003).

There are several types of CBM that can be used in the area of reading. At the kindergarten and primary level, early literacy measures are frequently used. The most widely used reading CBM for first grade and above is a brief assessment that uses a one-minute fluency probe as an indicator of overall reading performance, which is often known as R-CBM or Oral Reading Fluency (ORF) (Reschly et al., 2009). Another commonly used reading CBM is called Maze. This is a multiple-choice cloze task, in which a student is asked to read a passage silently and circle the correct missing words from three choices presented in parentheses (Shinn & Shinn, 2002b). For the purposes of this study, all standardized oral reading CBM measures will be commonly known as R-CBM, regardless of the publisher. R-CBM has been advocated as an indicator of overall reading performance, called a General Outcome Measure (GOM) (Shinn & Shinn, 2002a). As such, R-CBM is intended to assess all areas of reading, including reading comprehension, although it does not directly measure comprehension skills (Fuchs, Fuchs, Hosp & Jenkins, 2001; Good, Simmons & Kame'enui, 2001).

R-CBM has been widely studied as a screening tool and as a of individual student growth over time. Scores have been shown to adequately reliable and have been evaluated for validity as well (Deno, 1992). R-CBM has been advocated as a useful tool for identifying students who are at-risk of academic failure, in part because it avoids the biases that may be part of traditional teacher-referral processes (Marston, Mirkin, & Deno, 2001). Previous studies have examined the reading performance of students in the primary grades, and have demonstrated that R-CBM is sensitive to growth in reading for ELLs, but that rates of improvement are different for ELLs than for the general education population (Dominguez de Ramirez & Shapiro, 2006; Farmer, Swanlund, & Pluymert, 2010).

In the current climate of accountability and high-stakes testing, it is significant to note that R-CBM has been demonstrated to have good predictive validity when used to predict students' performance on state standards tests. A meta-analysis by Reschly et al. (2009) investigated findings from over 100 studies conducted over a period of thirty years, and found an overall strong correlation between R-CBM and a range of achievement measures. Several studies included in this meta-analysis have specifically investigated predictive validity with regard to state standards tests and have found significant correlations (Baker, et al., 2008; Good, Simmons, & Kame'enui, 2001; Hintze & Silberglitt, 2005; Stage & Jacobsen, 2001; Wiley & Deno, 2005). At least three additional studies have found positive predictive relationships between R-CBM and the

Illinois Standards Achievement Test (ISAT) (Sibley, Biwer & Hesch, 2001; Caywood-Rukas, 2010; Ditekowsky, n.d.).

The research regarding predictive validity of R-CBM for ELL students with relation to state standards tests is much more limited. Small-scale studies have been conducted with groups from the upper Midwest that are primarily Hmong- and Somali-speaking (Wiley & Deno, 2005), and students in California (Cabrera, 2008) and Arizona (Stokes, 2010) who are primarily Spanish-speaking. While these studies reported moderate to strong correlations between R-CBM and results on state-standards tests, the demographic make-up of samples in the studies, definitions of English-Language Learners and methodologies varied across studies. It is important to note that the ELL population is heterogeneous, and it is reasonable that differences in predictive validity may be seen depending on home language, socioeconomic status, race, and ethnicity of students in any given group.

Statement of the Problem

A significant problem facing many educators today is finding a way to reliably identify English Language Learner students who are at-risk of not meet academic standards. Schools and districts are held accountable for the academic performance of all students, including those who have not yet attained adequate proficiency in English. Schools that do not meet performance targets for English Language Learners and other groups may be subject to state or federal penalties (NCLB, 2001). Because of these changes in federal legislation, many schools are attempting to provide additional

instruction and intervention for students who are at-risk of academic failure. In order to ensure that the correct students are identified and provided with intervention, it is crucial to have assessment tools that are validated for use with the ELL population at all grade levels.

R-CBM is commonly used in universal screening and identification of students who are at-risk and assessing individual student progress in reading, and there is an increasing body of literature that supports their use for these purposes (Reschly et al., 2009). While the use of R-CBM in a Response to Intervention model is a promising approach for English Language Learners (NCCRESt, 2005 Vanderwood & Nam, 2007), questions remain. Much of the existing research focuses on students in the primary grades, who are developing reading proficiency (Baker, Plasencia-Peinado, & Lezcano-Lytle, 1998). Additional research should be conducted involving the utility for monitoring progress and predictive validity of R-CBM with English Language Learners at all grade levels, across states, and with a broader range of ELL students.

Research Questions

The current study will add to the research regarding the predictive validity of R-CBM for students from diverse backgrounds. Building on previous research that supports the use of reading fluency as a general outcome measure at all elementary grade levels, the current study investigates the utility of R-CBM as a measure of growth and relationship between performance on R-CBM and the Illinois state standards test. The sample used in this study consists of students in intermediate grades who come from a

wide range of ethnic and linguistic backgrounds and addresses the following research questions:

1. What are average initial scores and rates of growth in R-CBM for students in grades 4, 5, and 6 in a particular school district, and are there significant differences when comparing ELL and non-ELL students?
2. When examining performance on R-CBM for ELL students, does initial status or rate of improvement vary based on demographic factors such as level of language proficiency, ethnicity, and home language?
3. To what extent are R-CBM scores predictive of performance on a state standards test for 4th through 6th grade students, considering ELL status?
4. When comparing R-CBM and MAP, which is the most significant predictor of performance on a state standards test for sixth grade ELL and non-ELL students?

CHAPTER TWO

REVIEW OF THE LITERATURE

Over the past two decades, the number of school children who speak another language in the home has increased (U.S. Department of Education, 2011). Many of these children enter school with limited proficiency in English, and many have other risk factors as well, such as poverty and racial or ethnic minority status (Capps, Fix, Murray, Ost, Passel, & Herwanto, 2005). These students are at risk for academic difficulties, and have historically underperformed when they are compared to peers who enter school speaking English proficiently (U.S. Department of Education, 2007). Recent federal education legislation, such as the reauthorization of the Individuals with Disabilities Education Act and No Child Left Behind, has increased the emphasis on ensuring that all students, regardless of English Language Learner (ELL) status or other risk factors, achieve grade-level academic standards. ELL students underperform when compared to their peers, particularly in the area of reading (Keiffer, 2008). Therefore, the development of literacy in English Language Learners must be closely investigated, as well as the most appropriate methods to use for assessing and identifying ELL students who are at risk for problems with reading.

The following chapter will address research that has been conducted regarding literacy development for the general population of elementary school students, as well as

literacy development for English Language Learners. Research regarding the use of oral reading fluency curriculum-based measurement as a general outcome measure for assessing overall reading will be discussed. The predictive validity of R-CBM will be examined for all students, and specific attention will be given to research involving ELLs. Finally, research conducted in Illinois predicting performance on the Illinois Standards Achievement Test will be summarized.

Literacy Development

Learning to read is a complex process involving the development of multiple skills and the use of a wide range of knowledge (Adams, 1990; Snow, Burns, & Griffin, 1998). Through the 1980s and 1990s, the “reading wars” were waged in the field of education. Researchers and practitioners debated between the relative merits of phonic-based or whole language instruction (Anderson, 2000). In order to settle these debates and inform national education policy, the U.S. Congress established a panel of educators and researchers to review the existing literature and report on their findings (National Institute of Child Health and Human Development, 2000). The National Reading Panel (NRP) was convened in 1997 in response to a congressional charge to the National Institute of Child Health and Human Development (NICHD) in consultation with the Secretary of Education. Due to the extent of research that has been published in the area of reading in recent years, the panel chose to focus on several specific topics. These were alphabets, fluency, vocabulary, comprehension, teacher education, and computer instruction and technology. Meta-analyses were conducted by subgroups in each of these

areas, and the results were summarized in a summary report (NICHD, 2000) as well as separate reports from each of the subgroups.

In the area of alphabets, the NRP found significant support for phonemic awareness instruction. They found that students who received specific instruction in phonemic awareness showed improvements in reading and spelling that lasted beyond the end of training. Phonics instruction was also examined, and it was determined that systematic phonics instruction has a significant benefit for students in kindergarten through sixth grade, as well as for older students who are struggling readers. In the area of fluency, the NRP investigated the efficacy of guided oral reading and independent silent reading. The researchers found that guided oral reading and repeated reading had a significant positive impact on word recognition, fluency, and comprehension. However, they did not find enough support to indicate that independent silent reading had a significant impact on reading achievement. The panel also found that vocabulary instruction can lead to significant gains in comprehension. For optimal learning, vocabulary should be taught both directly and indirectly, and instruction should include repetition and active engagement of the student. Directly teaching text comprehension techniques using a multiple-strategy method was found to be effective as well (NICHD, 2000).

The results of the National Reading Panel report have been summarized elsewhere as the “5 Big Ideas in Beginning Reading” (University of Oregon Center on Teaching and Learning, n.d.) or the five components of reading (No Child Left Behind,

2001). These principles have been used in the definition of reading instruction in No Child Left Behind and the Individuals with Disabilities Act. Because NCLB has a focus on improving beginning reading so that all children are reading by third grade, there are many federally and state programs such as Reading First that focus on students in kindergarten through grade 3. Although the National Reading Panel focused on reading instruction for students in kindergarten through grade 12, there has been less emphasis on applying the findings to students in grades 4 and above. The National Institute for Literacy (2007) published a report that summarizes the research findings relating to reading instruction for adolescents, based on the report from the National Reading Panel. They list the five key components of reading for adolescents as decoding, morphology, fluency, vocabulary, and comprehension. They discuss how the concepts of phonemic awareness and phonics can be incorporated in instruction on reading multisyllabic words and understanding the component parts of words.

Literacy Development for English Language Learners

Around the same time that the National Reading Panel's recommendations were released, another panel was formed to examine research regarding literacy development in English Language Learners. This panel, called the National Literacy Panel on Language-Minority Children and Youth, was funded by the U.S. Department of Education, Office of English Language Acquisition and the National Institute of Child Health and Human Development and was comprised of experts from the fields of reading, language, research methodology, and bilingualism (McCardle, 2006). The

resulting report was published in a volume edited by August and Shanahan (2006). The panel investigated the development of literacy in language minority children, the role of first-language proficiency in second-language development, the role of environmental variables, instruction and professional development, and student assessment.

August and Shanahan report that the review of the literature indicated that language minority students follow a similar path of development as do their English-speaking peers. Word-level skills develop before reading comprehension. Research indicates that language-minority students may develop word-level components of literacy at similar levels to their peers, but that text-level skills such as vocabulary and comprehension lag behind their peers in the higher grades (Lesaux, 2006; August & Shanahan, 2006). The authors of this meta-analysis indicate that this suggests that oral language proficiency is likely associated with reading comprehension skills in English Language Learners (August & Shanahan, 2006; Geva, 2006). Additional studies have confirmed the linear nature of the development of word reading skills in ELL students.

In a longitudinal study of children from Spanish-speaking homes who received English-only instruction beginning at the preschool level, Mancilla-Martinez and Lesaux (2011) found that growth in English word reading was linear and was equivalent to word-reading skills of English-only students from kindergarten through the age of eleven. However, oral language proficiency in both Spanish and English, as measured through expressive vocabulary and sentence repetition, remained significantly below that of English-only peers. The rates of growth on both measures for ELLs was significantly

lower than that of English-only peers, such that the gap between ELLs and non-ELLs widened over time rather than closing. As vocabulary and overall oral language proficiency have been found to be related to reading comprehension in English Language Learners, these findings have significant implications for the validity of early literacy measures that focus on word-reading skills in determining risk status and predicting long term outcomes for ELLs.

Curriculum-Based Measurement

Curriculum Based Assessment is a broad category of measures that involve direct observation of a student's academic skills, using the classroom curriculum (Deno, 1985; Shinn, 2009). Curriculum Based Measurement (CBM) is a particular type of Curriculum Based Assessment, but is differentiated in that the assessments are standardized, using specific materials, administration procedures, and scoring (Deno, 2003, Shinn, 2008). CBM started as a method of progress monitoring to make educational decisions for student in special education (Deno, 2003). In subsequent years, CBM has been widely researched and has been used for screening, identification and placement in remedial and special education programs, formative evaluation, and mainstreaming decision-making (Deno, 2003). In addition, more recent research is developing to support the use of CBM in predicting outcomes on high-stakes testing as well as in early childhood and secondary settings (Deno, 2003; Fuchs, 2004). Many studies have been conducted to evaluate the predictive validity of R-CBM measures when compared to other measures of reading performance. Earlier studies were primarily conducted with relatively homogeneous

groups of students, and primarily used individually administered achievement tests as outcome measures. Later studies used larger, more diverse samples of students and investigated questions of variability between grade levels and subgroups.

Fluency is an essential component of reading, and has an impact on overall reading ability and comprehension. Students who are not fluent readers demonstrate difficulties with comprehension because their reading is not yet automatic, and their attention cannot be entirely focused on the meaning of the words that they are reading (NICHD, 2000). R-CBM is a one-minute reading fluency probe. A student is asked to read a grade-level passage aloud while the examiner records the number of words read correctly and number of errors (Fuchs, Fuchs, Hosp & Jenkins, 2001). In order to be able to read fluently and accurately, students engage in a complex process involving translating letters into sounds, putting sounds together into whole words, making connections between words and sentences, using prior information and making inferences (Fuchs, Fuchs, Hosp, & Jenkins, 2001; Fuchs, 2004). R-CBM is described as a dynamic indicator of basic skills (DIBS) (Shinn, 2008) or a general outcome measure (Fuchs & Deno, 1994). Some critics claim that this is only a measure of decoding, in that words read correctly are counted (Hamilton & Shinn, 2003). It is important to note that R-CBM is not intended to be a comprehensive reading assessment, and is not intended to be diagnostic (Shinn, 2008). Instead, it is a general outcome measure, and an indicator of overall reading performance. The predictive validity of R-CBM assessments is discussed below.

One criticism of R-CBM as a general outcome measure is that it does not identify students who are “word callers”. This is a term that is widely used by teachers and others in education, but not clearly defined (Meisinger, Bradley, Schwanflugel, Kuhn, & Morris, 2009). It is generally understood to mean students who are able to read accurately or fluently, but who do not have adequate comprehension skills. Hamilton and Shinn (2003) and Meisinger et al. (2009) investigated the phenomenon of word callers and found that they are not as widespread as teachers report. At the second and third grade level, both studies found that the number of students who fit criteria to be considered a word caller (adequate fluency but below average comprehension) was not significant. Meisinger et al. also found that teacher nominations of students as word callers did relate to those who met the criteria developed by researchers. In addition, teachers did not share a common definition of what characterized a word caller. However, Meisinger et al. did find that a significant number of students met the word caller criteria at the fifth grade level. This may suggest that the number of students who are fluent readers but who struggle with comprehension increases through the intermediate grades.

Several studies have investigated the relationship between R-CBM and state standards tests and found significant positive relationships between the two assessments, indicating that R-CBM has adequate predictive validity with relation to these exams (McGlinchey & Hixson, 2004; Sibley, Biwer, & Hesch, 2001; Stage & Jacobson, 2001). A meta-analysis by Reschly, Busch, Betts, Deno, and Long (2009) examined 108 studies that were conducted between the late 1970’s and 2008. Each of these studies investigated

the relationship between R-CBM measures and a variety of norm-referenced achievement tests and state standards tests. These included group administered state-standards tests, group administered national standards tests, and individually administered standardized achievement tests. Their findings supported a strong overall correlation between R-CBM results and standardized tests of reading achievement, and did not find significant differences depending on grade levels. They did find differences in correlations depending on the source of the test, administration format (individual or group) and reading subtest type. While R-CBM was a significant predictor of performance on state-specific tests of reading standards, the magnitude of the correlation was not as high as the overall correlation when including all types of tests, and not as high the correlation between R-CBM and tests based on national standards.

Some of the questions that have been raised about the predictive validity of CBM measures with specific groups involve grade levels, ethnic groups, socioeconomic status, and ELLs. Jenkins and Jewell (1993), in an early small study of students in grades 2-6 in elementary schools in the Pacific Northwest, found that the predictive validity of R-CBM measures decreased at higher grade levels. In this study, a standard set of passages was used for students at all grade levels, rather than using passages with increasing difficulty in the higher grades, as is more commonly seen in more recent research in reading CBM. At the time that this study was conducted, standardized CBM passages such as those published by DIBELS and AIMSweb were not as widespread. Similar findings were not seen in the meta-analysis conducted by Reschly, et al. (2009), where no differences were

seen in predictive validity when comparing grade levels, either individually or in groups (e.g. grades 1-3 and grades 4-6). However, it should be noted that the number of studies reviewed decreased at the higher grade levels, with only four studies included at grade 6.

Published studies have specifically investigated the validity of R-CBM measures with racial or ethnic minority students. Two examined the predictive bias when comparing African American and Caucasian students, and came to different conclusions. Kranzler, Miller, and Jordan (1999) conducted a series of simultaneous multiple regression analyses, using a sample of students in grades 2-6 who were administered grade-level specific reading passages. They found significant predictive bias in grades 5 and 6, indicating that R-CBM overestimated the reading comprehension skills of African American students and underestimated the skills of Caucasian students. Hintze, Callahan, Matthews, Williams and Tobin (2002) used a similar sample, but administered the same passages to students in all grade levels and conducted a series of hierarchical regression analyses using age, sex, socioeconomic status (SES), and ethnicity as predictors. Hintze et al. did not find significant predictive bias when comparing ethnic groups, and did not find that R-CBM over- or underpredicted reading comprehension skills when controlling for age, sex, and SES for African American or Caucasian students. They did find that a model based on age, SES, and R-CBM scores accounted for 25% more of the variability in reading comprehension scores in African American students than in Caucasian students

Pearce and Gayle (2009) also examined the predictive validity of R-CBM for a sample of White and American Indian third-grade students in the Upper Plains. They found that R-CBM had adequate predictive validity for both groups, but that the American Indian group performed at a significantly lower level than the White group. Negative predictive power was found to be robust for both groups, but positive predictive power was stronger for the American Indian group. There were also significantly more false negative results in the American Indian group, indicating that a greater number of American Indian students may perform adequately on the R-CBM assessment, but not meet standards on a state test. The authors argue that this is a concern, because this suggests that some American Indian students may not be correctly identified as at-risk for reading failure if R-CBM is used as a screening measure.

Baker et al. (2008) investigated the use of R-CBM as a predictor of overall reading proficiency for students in schools that were low-performing and high poverty. Thirty-four schools that qualified for a statewide “Reading First” initiative were involved in the study. They found moderate to moderately strong correlations between R-CBM scores in grades 1-3 and performance on standardized achievement tests in second and third grade. They also found that including the students’ rate of improvement, or slope, in the analysis improved the predictive validity over using a single R-CBM data point.

Curriculum Based Measurement with English Language Learners

Early Literacy Measures

There is a developing body of research that supports the use of early literacy CBM measures to identify ELL students in kindergarten and first grade that would benefit from additional intervention. Vanderwood, Linklater, and Healy (2008) examined the predictive accuracy of an early literacy CBM assessment for ELL students. Students were administered a brief phonics assessment (Nonsense Word Fluency) in first grade, and the results were analyzed with relation to three outcome measures in third grade. The first grade scores were demonstrated to have significant correlations to all three third grade measures. Fien, Baker, Smolkowski, Mercier Smith, Kame'enu, & Smith (2008) also examined the validity of Nonsense Word Fluency (NWF) with a large sample of students in grades K-2 in Oregon and found that correlations between NWF and criterion measures were generally equivalent, when comparing ELLs and their non-ELL peers. They found that NWF had a significantly stronger correlation for non-ELL students at the initial kindergarten administration, indicating that this measure should be used with caution for ELLs at that point. Linklater (2007) also found that CBM measures of phonemic awareness (initial sound fluency and phoneme segmentation fluency) have adequate predictive validity for ELL students in kindergarten.

Another study by Betts, Reschly, Pickart, Heistad, Sheran, and Marston (2008) investigated the predictive validity of early literacy measures for ELLs and ethnic subgroups. They found that kindergarten early literacy assessments were significant

predictors of performance on a reading assessment at the end of second grade. The authors also found no predictive bias when comparing ELL and non-ELL groups. Home languages spoken by ELLs included Spanish, Hmong, Somali, and Oromo. Predictive bias was found when comparing the European American and Hispanic American subgroups. The authors suggest that this finding may have implications for schools with populations of ELLs who are primarily Hispanic American and speak Spanish as their home language.

Reading CBM

Dominguez De Ramirez and Shapiro (2006) investigated the use of R-CBM with students in Spanish bilingual education programs in grades 1-5. They found that compared to students in general education, the bilingual students demonstrated lower overall scores in both Spanish and English, and also lower rates of growth. While the rates of growth in English for the bilingual students was close to that of general education peers by fifth grade, a significant gap in overall scores remained. They did find that students in bilingual programs made significant growth in English reading fluency, even when that was not the primary language of instruction. Therefore, the authors concluded that English R-CBM could be a viable measure to use for Spanish-speaking students in bilingual programs.

Betts, Bolt, Decker, Muyskens, and Marston (2009) investigated the relationship between the length of time that a student's family had been in the country, home language, and reading development. Based on data from a group of third grade students

whose home languages were Somali and Spanish, they found that R-CBM scores were highly correlated with performance on a test of broad reading achievement. They did not find that language group (Somali or Spanish) had a moderating effect between time in the United States and reading performance. However, they did find that for Somali students, their increased time in the United States has a significant direct impact on the development of oral reading fluency, while the same effect was not seen with Spanish-speaking students. The authors note that the opposite finding was expected, given that the Spanish language is more similar to English and shares the same alphabetic code. However, they argue that the findings may be due to the fact that Hispanic communities have longer and deeper ties in the United States, and that Spanish speakers therefore have more opportunities to engage with Spanish language and home culture. Somali speakers likely have fewer native language opportunities and may be exposed to more English language through the media and mainstream culture than Spanish speakers.

Farmer, Swanlund, and Pluymert (2010) examined the performance of ELL and non-ELL students in grades 1-3 on R-CBM and used a growth model to examine initial scores and rates of improvement. They found that initial scores for ELL students were significantly lower than non-ELL peers. The relationship between ELL status and rate of growth was inconsistent across grade levels, with ELL students in first grade improving at a rate that was lower than non-ELL peers, and the reverse relationship in second grade. Other factors such as reading intervention, socioeconomic status, and ELL programming

(bilingual or English-only instruction) were also found to be inconsistently related to initial R-CBM scores or rate of growth, depending on grade level.

Limited research has investigated the predictive validity of R-CBM measures for students in the intermediate elementary grades. Wiley and Deno (2005) examined a sample of students in third and fifth grades who were administered R-CBM and Maze measures in fall and the Minnesota Comprehensive Assessment in the spring of the same year. Approximately half of the students in each group were ELLs, and the home language for the majority (80 percent) of the students was Hmong. For ELLs, significant moderate correlations were found between both R-CBM and Maze and the state test, but they were not as strong as those seen with non-ELLs at the third grade level. At the fifth grade level, a moderately strong correlation was found between ELL's performance on R-CBM, while the correlation for non-ELLs was moderate. For Non-ELL students, stronger correlations were found between Maze and the state standards test when compared to R-CBM at the fifth grade level. Mean scores on R-CBM were higher for the non-ELL group at both levels, and the mean score for fifth graders was 142 words per minute, compared to 123 words per minute for ELLs.

One reason for this finding may be that R-CBM continues to be significantly correlated with overall reading performance for ELLs is that many have not yet obtained adequate fluency, while non-ELL peers at the same grade level have attained adequate fluency. As discussed above, previous research has consistently found that ELL students lag behind their non-ELL peers in overall reading at the intermediate grade levels

(Kieffer, 2008; Lesaux, 2006; August & Shanahan, 2006). It is possible that when most students within a given group have attained adequate fluency, R-CBM is no longer a useful measure. In effect, students are near the “ceiling” of the R-CBM task. Because ELLs tend to perform at a lower level than their non-ELL peers, it is reasonable that they would not, as a group, reach this ceiling as soon as their peers. Findings are inconsistent regarding the predictive validity of R-CBM at the intermediate and middle school levels (grades 4-6 and beyond).

Cabrera (2008) investigated the relationship between R-CBM scores and outcomes on a California state language arts test for third through sixth graders at a primarily low income elementary school. Cabrera looked at language minority (LM) and English-only students, and he classified any student whose parents indicated on a home-language survey that they spoke a language other than English at home as an English-Language Learner, regardless of language proficiency. The findings of this study indicate that R-CBM scores predicted outcomes on the state test for language minority students, but did not predict outcomes for English-only students who adequate scores (“low risk”) on R-CBM. However, the sample of English-only students used in this study was small (N=42). He also did not disaggregate his data based on grade level, and therefore did not address whether predictive validity remains consistent for language minority students at higher grade levels.

Stokes (2010) also investigated the relationship between performance on R-CBM measures and performance on the Arizona state standards test in reading at the middle

school level. She found that initial performance on R-CBM was significantly lower for Spanish-speaking ELL students than for non-ELL peers. No significant differences were found in rates of growth when comparing the two groups. R-CBM was found to be a significant predictor of performance on the AIMS assessment for both ELL and non-ELL students at the sixth grade level.

A recent study by Hosp, Hosp, and Dole (2011) examined predictive bias when using CBM assessments to predict performance on a state standards test in Utah. The authors found predictive bias when comparing ELL and non-ELL students in grades 1-3 when using Nonsense Word Fluency and DIBELS ORF measures. While the ORF measures resulted in adequate sensitivity in identifying those at risk of not meeting standards on the state test, they did not have adequate specificity for grades two and three. This suggests that ELL students might be over-identified as at-risk based on CBM measures in these grade levels.

R-CBM and the Illinois Standards Achievement Test.

In Illinois, three unpublished studies have been conducted that investigate the use of R-CBM to predict outcomes on the Illinois Standards Achievement Test (ISAT). Sibley, Biwer, and Hesch (2001) examined correlations between performance on R-CBM, the ISAT, and a standards-based assessment used by the district. The authors found a strong correlation between fall third grade R-CBM scores and and spring third grade ISAT scores (0.750, $p < 0.001$). They also found moderate to strong correlations when comparing the previous year's spring R-CBM scores to the RIT scores from Level

Achievement tests in the fall for grades 3-5. Caywood-Rukas (2010) also examined the predictive validity of kindergarten early literacy CBM measures and first grade R-CBM with regard to ISAT and another frequently used achievement test in grades three through five. A moderate correlation was found when comparing first grade R-CBM performance and third grade ISAT. She argued that the reason for the relatively weak correlation could be that first grade screening data was used to place students in reading interventions, which, if successful, may have reduced the risk status of many of the students initially identified as below standards in first grade.

Ditkowsky (n.d.) conducted an unpublished study of the relationship between R-CBM, Maze scores and ISAT in order to develop cut scores that would accurately predict performance on the state standards test. Data was collected in a medium-sized suburban district, and logistic regression and linear discriminant function analysis were used as the methods of analysis. Ditkowsky reports that when determining which students are at-risk of performing below standards on the ISAT, there are false negative rates of 20-30% and false positive rates of 10-15%. The author states that these scores were confirmed with data from additional districts in Illinois and ISAT data from later years. However, additional data to support those statements are not provided. Ditkowsky (2009) also reported updated cut scores based on a larger sample of students in another unpublished paper on his website. He reported a sample size of 1300 to 2600 students from eight districts in northern Illinois, and indicated that the predictive validity of R-CBM to ISAT ranged from .71 to .74. Specific details regarding descriptive statistics, methods, and data

analysis were not reported. The resulting cut scores are widely used across Illinois for the purpose of program evaluation.

Ditkowsky and Koonce (2010) used a sample of third-grade students in Illinois, and the relationship between growth and performance on the ISAT. They divided students into three groups based on their growth in R-CBM from fall to spring, designated as “adequate progress”, “questionable progress” and “not adequate progress”. Initial fall R-CBM scores were significant predictors of ISAT scores for all three groups. When comparing groups, students who made adequate progress in R-CBM also had significantly higher scores on the ISAT than students in the other two groups. For students who received special education services, making adequate growth on R-CBM was also a significant predictor of passing the ISAT.

Summary

Research in the development of literacy has identified five key components that should be in place in any instructional reading program (NICHD, 2000). In general, students develop from reading at the individual word level to high level skills involving vocabulary, fluency, and comprehension. Although these five components have been used primarily with reading programs in the primary grades, similar components should be in place in literacy programs for young children through adolescents (National Institute for Literacy, 2007). Some research indicates that development of literacy in English Language Learners or Language Minority students generally follows the same trajectory (August & Shanahan, 2007). However, additional research indicates that while

ELLs may develop word-level reading skills at similar levels to their peers, vocabulary and comprehension may lag behind, especially in the intermediate grades and above (August & Shanahan, 2007), and that growth in reading comprehension in ELLs may follow a different growth curve than non-ELL peers (Mancilla-Martinez & Lesaux, 2011).

Research supports the use of R-CBM for a variety of different purposes. Most relevant to the current study, R-CBM has been shown to have good predictive validity when used to predict performance on a variety of different types of reading achievement tests (Reschly et al., 2009). A number of state-standards tests have been studied in this regard, including the Illinois Standards Achievement Test, and adequate predictive validity has been demonstrated (Sibley, Biwer & Hesch, 2001; Caywood-Rukas, 2009; Ditkowsky, n.d.). Limited research indicates that R-CBM may be a reliable and valid measure for English-Language Learners, but questions remain regarding predictive validity, especially at the intermediate grades and above. Previous research indicates that the prevalence of “word-callers” may increase in the intermediate grades (Meisinger, 2009). Given that ELLs may have adequate word-level reading skills, but may lack comprehension, it is possible that some ELLs may fit the description of “word-callers”. Therefore, the predictive validity of this type of assessment for ELLs in upper elementary and beyond remains in question.

CHAPTER THREE

METHODOLOGY

Introduction

The purpose of the current study was to examine the utility of curriculum-based measurement in reading as a universal screener and a measure of individual growth in reading performance over time for English Language Learners (ELLs) and their non-ELL peers in the intermediate grades. In the current study, English Language Learners are defined as those students with limited English proficiency as identified through their performance on state-mandated assessments of English-language proficiency. The predictive validity of R-CBM measures were examined as well, with relation to both a state standards test and a widely-used, norm-referenced assessment of overall reading. All students in the current study were assessed using reading curriculum-based measurement (R-CBM) assessments, three times per year. This is a short, standardized assessment that uses the number of words read per minute as an indicator of overall reading performance. The current study examines differences in initial R-CBM scores when comparing ELL and non-ELL students. Differences between ELL and non-ELL students when comparing rates of growth in R-CBM performance were examined as well. For ELL students, further analyses examined whether factors such as language proficiency, home language, or race/ethnicity can account for any of the variance in initial scores or rates of improvement. Predictive validity was examined

with respect to the Illinois Standards Achievement Test, which is a state standards test. Predictive validity of R-CBM was also compared to that of the NWEA Measures of Academic Progress (MAP) assessment, a computer-based assessment administered three times per year.

Subjects and Setting

The participants in the current study were 4th, 5th, and 6th grade students in an elementary district in a suburb of Chicago, Illinois. The approximate number of students assessed at each grade level is shown in the table below.

Table 1. Number of Students Assessed during the 2010-2011 School Year

Grade level	Number of students assessed
4 th grade total	378
5 th grade total	356
6 th grade total	370
4 th grade ELL	82
5 th grade ELL	59
6 th grade ELL	59

There are six elementary schools and one junior high in the district. Demographic data for the entire district grades K-8 was obtained from the Illinois District Report Card

from 2011. As of 2011, the district has been identified as not making Adequate Yearly Progress in the areas of reading and mathematics and has been identified for district improvement under state and federal guidelines. Four schools (three elementary and one junior high) are identified as Title I schools, and two of these schools are in Federal School Improvement Status. See Table 2 for additional demographic data for all students in the school district in the proposed study.

Table 2. Overall Demographic Information for the School District

Total Enrollment	3,530
Percent White	35.4%
Percent Black	4.9%
Percent Hispanic	22.4%
Percent Asian/ Pacific Islander	35.5%
Percent Native American	0.2%
Percent Multiracial/Ethnic	1.6%
Percent Low Income	43.3%
Percent Limited English Proficient	33.7%
Percent IEP	13.4%
Chronic Truancy Rate	0.1%
Mobility Rate	9.6%
Attendance Rate	95.1%
Percent Meeting and Exceeding Standards on Illinois Standards Achievement Test	85%

All students in grades four, five, and six during the 2010-2011 school year were eligible to be participants in the study. Cases with significant missing data were eliminated from the study. Any students who were missing demographic information, missing ISAT scores, or missing all R-CBM data were removed from the sample. For the purposes of this study, any students who had ACCESS for ELLs scores were counted as English Language Learners. This may have included some students had been dismissed from receiving ELL services and were being monitored. Students whose parents refused ELL services were also included in the sample.

Procedures

Archival data from the 2010-2011 school year was be used in this study. Assessment procedures occurred as part of the standard assessment schedule for all students in the district, unrelated to the current study. Confidentiality was maintained by coding all students and removing any identifying information. Student test data was transmitted to the author via email with identifying student information, such as student ID numbers and student names. Downloaded files were password-protected. Student identifying data was used to match cases in order to merge data files. Subsequently, each student was randomly assigned a code number and identifying student information was deleted. All electronic files were password protected, and all paper copies with student identifying information were stored in a locked filed cabinet for the duration of the study, and will be shredded. Only the author will have access to electronic or paper files.

Each participant completed R-CBM assessments using standardized measures published by AIMSweb, an internet-based assessment and data management system. R-CBM assessments were administered three times throughout the given school year, in September, January, and May. All sixth grade students also were administered the Measures of Academic Progress (MAP) assessment during similar time periods. This is a norm-referenced, computer-based, adaptive assessment of reading, mathematics, and language usage. All ELL students were administered the Assessing Comprehension and Communication in English State-to-State for English Language Learners (ACCESS for ELLs) assessment in January of 2011 in order to assess English language proficiency.

Measures/Instrumentation

AIMSweb Reading Curriculum-Based Measurement (R-CBM)

AIMSweb Oral Reading CBM (R-CBM) is a general outcome measure of overall reading development (Shinn & Shinn, 2002a). Students are individually asked to read aloud selected passages that are graded and meant to represent reading curriculum. After one minute, the total number of words read correctly and total errors are recorded (Shinn & Shinn, 2002a). Independent studies have examined test-retest and alternate form reliability and results range from .82 and .99 (Shinn & Shinn, 2002a). Construct validity and predictive validity have been widely researched and found to meet acceptable standards (Shinn & Shinn, 2002a; Marston, 1989; Reschly, et al., 2009). Marston (1989) reported correlations between CBM reading measures and published measures of reading comprehension as ranging from .63 to .90, with most above .80.

AIMSweb scores are reported as number of words read correctly per minute. The AIMSweb system also allows for a variety of reporting options, including comparisons to building or district-level peers, or comparisons to national norms. AIMSweb provides default cut scores that are set at the 45th percentile based on national norms (AIMSweb, 2011). The cut scores are set at the 45th percentile based on research indicating that students who scored at or near this level generally had a success rate of 80% on state standards tests (AIMSweb, 2011). This research was conducted using data from twenty states, and is not specific to any particular state test (AIMSweb, 2011).

Assessing Comprehension and Communication in English State to State for English Language Learners (ACCESS for ELLs)

ACCESS for ELLs is a language proficiency test for students in kindergarten through twelfth grades. It is designed to monitor English language proficiency on a yearly basis and help determine when full proficiency has been achieved. It is designed to be representative of social and academic language as defined by the World-class Instructional Design and Assessment (WIDA) Consortium's standards (WIDA, 2011). All students identified as ELLs in Illinois are required to take this test on a yearly basis in order to comply with the federal requirements under the No Child Left Behind (NCLB) legislation (Illinois State Board of Education (ISBE), n.d.). Test forms are provided for different grade level clusters (Kindergarten, 1-2, 3-5, 6-8, and 9-12). Within each grade level cluster from first grade on, there are three forms of the test available (A= Beginning; B= Intermediate; C= Advanced). Scores are reported as raw scores, scale

scores, and proficiency scores. Scale scores are based on a range of 100-600, allowing comparisons across grade levels and tiers. Proficiency scores are interpretive, noninterval scores, and are grade specific. Scores are reported in the domains of listening, speaking, reading, and writing. Composite scores in the areas of oral language, literacy, comprehension, and overall English Language proficiency are reported as well (WIDA, 2011). In Illinois, proficiency scores are used to determine whether students continue to qualify as ELL status (ISBE, n.d.). For the purposes of this study, overall composite scale scores will be used. The overall score is a combination of listening (15%), speaking (15%), reading (35%), and writing (35%).

A technical report based on testing from the 2009-2010 school year indicates that reliability was high, ranging from .928 to .972 depending on grade level cluster (Yanosky, Yen, Longuit, MacGregor, Zhang, & Kenyon, 2011). Moderate to strong correlations were found when comparing student performance on ACCESS for ELLs to performance on older language proficiency tests (Yanosky, et al., 2011). The reading and writing sections of ACCESS for ELLs have been shown to be accurate predictors of performance on a state standards-based test in the areas of reading, writing, and mathematics in fifth and eighth grade (Yanosky, et al., 2011).

Measures of Academic Progress (MAP)

Measures of Academic Progress (MAP) tests are norm-referenced, computer-based, adaptive assessments in reading, mathematics, and language usage developed by the Northwest Evaluation Association (Northwest Evaluation Association (NWEA),

2011a). Scores are provided using an equal-interval Rasch-unit (RIT) scale, which allows for comparisons throughout the school year or across grades (NWEA, 2011a). The areas assessed by MAP reading are Literary Works, Literature, Reading Strategies and Comprehension, and Word Analysis and Vocabulary. In the current study, MAP was administered three times per year in sixth grade only.

A study of the alignments between MAP reading scores and performance on the Illinois Standards Achievement Test (ISAT) was conducted using a sample of 83,000 students in over 290 schools in Illinois (NWEA, 2011b). MAP scores accurately predicted whether students in sixth grade would pass the ISAT in 90.1% of cases (NWEA, 2011b). The correlation between MAP and ISAT reading scores for sixth grade students was .735 (NWEA, 2011b).

Illinois Standards Achievement Test

The Illinois Standards Achievement Test (ISAT) is a state standards test designed to measure individual student performance with relation to state standards in reading, mathematics, writing, and science. For the purposes of this study, only performance on the reading portion of the assessment will be analyzed. The ISAT reading test includes multiple-choice items from the Stanford Achievement Test, Tenth Edition (SAT-10) as well as items written by Illinois educators. The four standard areas of reading assessed include Vocabulary Development, Reading Strategies, Reading Comprehension, and Literature (Illinois State Board of Education, 2010a). There are also two open-ended items requiring written responses (one pilot and one scored). The extended response item

accounts for 10% of the overall score in reading (Illinois State Board of Education, 2010a). Scores are reported on a vertical scale across grades. Cut scores, which vary from year to year, place students in the following performance categories (Illinois State Board of Education, 2010a):

Exceeds Standards: Student work demonstrates advanced knowledge and skills in the subject. Students creatively apply knowledge and skills to solve problems and evaluate the results.

Meets Standards: Student work demonstrates proficient knowledge and skills in the subject. Students effectively apply knowledge and skills to solve problems.

Below Standards: Student work demonstrates basic knowledge and skills in the subject. However, because of gaps in learning, students apply knowledge and skills in limited ways.

Academic Warning: Student work demonstrates limited knowledge and skills in the subject. Because of major gaps in learning, students apply knowledge and skills ineffectively. (p.1)

Analysis of the 2010 ISAT test administration indicates adequate reliability and validity (Illinois Board of Education, 2010b). Internal consistency across items on the reading test ranged from .89 to .92 depending on grade level. Interrater reliability for the extended response item on the reading test ranged from 98-99%. The reliability of the classification categories yielded p values between .843 and .997. Correlation with the SAT-10 was examined for the entire ISAT (which contains portions of the SAT-10) and the portions of the ISAT that are not derived from the SAT-10. Correlations were strong for both the entire ISAT and the non-SAT-10 derived portion. This indicates adequate concurrent validity with the SAT-10 (Illinois State Board of Education, 2010b).

Data Analysis

Growth Model: Hierarchical Linear Modeling

Hierarchical linear modeling is a statistical method that can be used to examine rates of growth (Raudenbush & Bryk, 2002). This allows for analysis of the slopes of individual student growth, rather than looking at average growth within a sample (Raudenbush & Bryk, 2002). HLM 7 software published by Scientific Software International was used to conduct the analyses. The analyses examined initial R-CBM performance and rate of growth based on the three data points during the school year (fall, winter, and spring). The analyses determined whether there is significant variance in initial status and rate of growth. A separate analysis was conducted for each grade level. At each grade level, the Level 1 analysis determined the average initial score and rate of growth. If there is significant variance, additional factors can be added to the model in order to determine whether they contribute to the variability in either initial status or growth. When significant variance was determined to be present, ELL status (yes/no) was added as a predictor at both Level 1 (initial status) and Level 2 (growth).

When it was determined that there were significant differences when comparing ELL and non-ELL students at all grade levels, an additional set of analyses examined initial status and rate of growth for ELL students only. When it was determined that there was significant variance in initial status and rate of growth for ELL students, additional factors were included in the model to determine whether they contribute to the

variance. The factors that were included in this model were level of language proficiency, home language, and race/ethnicity. Level of language proficiency was measured using ACCESS scale scores. Given the current sample, home language was represented as Spanish/ non-Spanish. However, after initial analyses, this predictor was removed due to problems with multicollinearity. Race/ethnicity was represented with dummy variables representing the categories of Asian and white. The comparison category was Hispanic..

Predictive Validity of R-CBM Scores

Predictive validity of R-CBM scores was examined in two ways. First, the correlation between R-CBM scores (words read correctly per minute) from a single point in time and ISAT scale scores was examined. The analysis was conducted separately for all students at each grade level, for ELL students at each grade level, and for non-ELL students at each grade level. At grade 6, correlations was also calculated for MAP scale scores and ISAT scale scores. These calculations were conducted for R-CBM scores and MAP scores from both the fall and winter benchmark periods.

In addition, AIMSweb default cut scores (AIMSweb, 2011) from both fall and winter were examined in order to determine diagnostic accuracy. Diagnostic accuracy was investigated for each grade level overall, non-ELL students, and for ELL students only at each grade level. Diagnostic accuracy refers to the following five measures (Stage & Jacobsen, 2001):

1. Sensitivity: Percentage of students below the cut score who do not meet standards on ISAT

2. Specificity: Percentage of students above the cut score who meet or exceed standards on ISAT
3. Positive predictive power: Probability that a student who scores below the cut score will not meet standards on the ISAT
4. Negative predictive power: Probability that a student who scores above the cut score will meet or exceed standards on the ISAT
5. Overall accuracy: The overall accuracy of the classification categories when corrected for chance.

Summary

The following research questions were addressed using the sample, measures, and methods described above.

1. What are average initial scores and rates of growth in R-CBM for students in grades 4, 5, and 6 in a particular school district, and are there significant differences when comparing ELL and non-ELL students?
2. When examining performance on R-CBM for ELL students, does initial status or rate of improvement vary based on demographic factors such as level of language proficiency, ethnicity, and home language?
3. To what extent are fall and winter R-CBM scores predictive of performance on a state standards test for 4th through 6th grade students, administered in March, considering ELL status?

- a. How closely correlated are R-CBM scores and scores on a state standards test?
 - b. How well do R-CBM cut scores correctly classify students in terms of their performance on a state standards test?
4. When comparing R-CBM and MAP, which is the most significant predictor of performance on a state standards test for sixth grade ELL and non-ELL students?

CHAPTER FOUR

RESULTS

The current study was conducted to investigate the predictive validity of R-CBM for students from diverse backgrounds. Building on previous research that supports the use of reading fluency as a general outcome measure at all elementary grade levels, the utility of R-CBM as a measure of growth and relationship between performance on R-CBM and the Illinois state standards test (ISAT) were studied. The sample used in this study consisted of students in intermediate grades that come from a wide range of ethnic and linguistic backgrounds. Analysis using Hierarchical Linear Modeling (HLM) was conducted to determine average initial scores and rates of growth in reading Curriculum-Based Measurement (R-CBM) for students in grades four, five, and six and whether there is significant variability in initial scores and rates of growth, depending on English Language Learner (ELL) status. The next question addressed through HLM was whether demographic factors such as level of language proficiency, ethnicity, and home language accounted for significant variability in R-CBM scores or rate of improvement. In addition, analysis of correlation between R-CBM and ISAT and diagnostic accuracy of R-CBM cut scores were investigated to determine predictive validity for ELL and non-ELL students. Finally, correlations between R-CBM and ISAT were compared to correlations between R-CBM and Measures of Academic Progress (MAP), another widely used reading assessment for students in 6th grade.

Descriptive Statistics

Description of the Sample

The following tables provide specific demographic information about the sample in the current study. Data collected includes gender, racial and ethnic categories, percentage of students who are eligible for free and reduced lunch, percentage of students who receive special education services, and the percentage of students who are English Language Learners. At each grade level, approximately one-third of the sample was identified as white, non-Hispanic. Approximately one-fourth of the sample identified as Hispanic. Slightly more than one third of the sample identified as Asian. Less than ten percent of the sample at each grade level identified as African American, American Indian, or two or more races. These categories were combined into a category of “Other”.

Table 3. Overall Demographic Data for the Current Sample

	Fourth grade	Fifth grade	Sixth grade
Total N	362	339	355
Race/Ethnicity			
White	34.3%	33.3%	34.4%
Hispanic	23.5%	22.1%	23.1%
Asian	35.6%	38.6%	34.6%
Other	6.6%	5.9%	8.2%
Other Demographics			
ELL	25.1%	20.4%	17.5%
IEP	11.0%	9.7%	11.5%
FRL	53.9%	49.9%	52.1%

At the fourth grade level, 25.1% of students were identified as ELL. In fifth grade and sixth grade, the number of students identified as ELL was slightly less (20.4% and 17.5%). Approximately 10% of students at each grade level were identified as receiving special education services (IEP). Approximately half of students at each grade level were identified as eligible for free or reduced lunch.

For ELL students, the same demographic information was collected regarding race and ethnicity, the percentage of students who receive special education services, and the percentage of students eligible for free or reduced lunch. The percentage of ELL students at each grade level who identified themselves as white ranged from 17.7% in sixth grade to 28.6% in fourth grade. Hispanic students accounted for between 35.3% of ELL students (fifth grade) and 46.8% of ELL students (sixth grade). Asian students accounted for between 31.9% (fourth grade) and 39.7% (fifth grade) of all ELL students. Students of other races accounted for between 0 and 3.2% of ELL students.

Table 4. Demographic Data for ELL Students

	Fourth grade	Fifth grade	Sixth grade
Total N	91	68	62
Race/Ethnicity			
White	28.6%	22.1%	17.7%
Hispanic	39.6%	35.3%	46.8%
Asian	31.9%	39.7%	32.3%
Other	0%	2.9%	3.2%
Other Demographics			
IEP	24.2%	27.9%	30.6%
FRL	74.7%	72.1%	72.6%

The percentage of students receiving special education services and eligible for free and reduced lunch was higher for ELL students than the overall student population described above. Depending on the grade level, between 24.2% and 30.6% of students identified as ELL also received special education services. Students eligible for free or reduced lunch made up between 72.1% and 74.7% of all students identified as ELL. Home languages and language proficiency were also investigated. Over 30 different home languages were reported across grade levels. Due to the large number of home languages, only the top five most reported languages at each grade level are included in the following table.

As shown in the table of home languages below, the most common home language spoken among ELL students was Spanish (39.6% of fourth graders, 35.3% of fifth graders, 46.8% of sixth graders). Spanish was by far the most common language at each grade level, and was more than three times more common than the second most common language. Gujarati was the second most frequently spoken home language (12.1% of fourth graders, 10.3% of fifth graders 12.9% of sixth graders). Among fourth graders, the third most commonly spoken home language was Urdu (8.8%), while Malayalam was the third most commonly spoken language with fifth (10.3%) and sixth (8.1%) graders. Assyrian was the fourth most commonly spoken language among fourth graders (6.6%). Among fifth graders, Urdu and Polish were equally common and represented the fourth and fifth most frequently spoken home languages (7.4%). Malayalam was the fifth most commonly spoken home language among fourth graders (4.4%). Among sixth grade ELL students, Pilipino (also known as Tagalog) and Assyrian

were equally common and represented the fourth and fifth most common home languages (6.5%).

Table 5. Home Languages for ELL Students

Fourth Grade		Fifth Grade		Sixth Grade	
Spanish	39.6%	Spanish	35.3%	Spanish	46.8%
Gujarati	12.1%	Gujarati	10.3%	Gujarati	12.9%
Urdu	8.8%	Malayalam	10.3%	Malayalam	8.1%
Assyrian	6.6%	Urdu	7.4%	Pilipino	6.5%
Malayalam	4.4%	Polish	7.4%	Assyrian	6.5%

ACCESS for ELLs test scores are reported below. For each grade level, the mean scores and standard deviations are reported. At the fourth grade level, the mean scale score was 346.65 (SD= 15.74). This places the average fourth grade ELL student in this district at proficiency level 3 to 4 on a scale of 1 to 6, which is interpreted as “developing” to “expanding” (WIDA Consortium, 2011). At fifth grade the mean scale score was 353.43 (SD = 12.97). This is also proficiency level 3 to 4 (WIDA Consortium, 2011). For sixth grade ELL students, the mean scale score was 355.19 (SD= 15.33). This is equivalent to level 2 to 3, which is interpreted as “beginning” to “developing” (WIDA Consortium, 2011). Therefore, although sixth grade ELL students demonstrate higher average scale scores than their fourth and fifth grade peers, their levels of language proficiency are relatively lower based on grade level expectancies.

Table 6. ACCESS for ELLs Scores

	N	Mean scale score	SD
Fourth Grade	91	346.65	15.74
Fifth Grade	68	353.43	12.97
Sixth Grade	62	355.19	15.33

Standardized Test Scores

The following tables display descriptive statistics for all students in grades four through six in the current study. The mean scores and standard deviations for all students in each grade are shown for the all students at each grade level. Test scores include R-CBM at fall, winter, and spring; ISAT scale scores, and MAP scores for students in grade six in fall and winter.

In fourth grade, the mean R-CBM score in fall was 109.73 with a standard deviation of 35.57. In winter, the mean R-CBM score was 125.24 with a standard deviation of 36.47. The spring mean R-CBM score was 141.93 with a standard deviation of 38.38. Based on AIMSweb cut scores, these scores are all above the cut score indicating “Tier 1” or low risk of failure on state standards tests (AIMSweb, 2011). The mean ISAT scale score in fourth grade was 218.80 with a standard deviation of 23.93. This score is within the category of “Meets Standards” indicating that students at this level meet state standards for overall reading (Illinois State Board of Education, 2011).

Table 7. Standardized Test Scores for Fourth Grade

	N	Mean score	SD
Fall R-CBM	360	109.73	36.57
Winter R-CBM	362	125.24	36.47
Spring R-CBM	358	141.93	38.38
ISAT	362	218.80	23.93

In fifth grade, the overall mean fall R-CBM score was 128.35 words read correctly with a standard deviation of 39.12. In winter, the mean R-CBM score was 144.86 with a standard deviation of 39.47 words read correctly. The spring mean R-CBM score was 160.72 with a standard deviation of 41.33 words read correctly. Based on AIMSweb cut scores, these scores are all above the cut score indicating “Tier 1” or low risk of failure on state standards tests (AIMSweb, 2011). The mean ISAT scale score was 231.51 with a standard deviation of 24.79. This score is within the category of “Meets Standards” indicating that students at this level meet state standards for overall reading (Illinois State Board of Education, 2011).

Table 8. Standardized Test Scores for Fifth Grade

	N	Mean score	SD
Fall R-CBM	337	128.35	39.12
Winter R-CBM	337	144.86	39.47
Spring R-CBM	336	160.72	41.33
ISAT	336	231.51	24.79

In sixth grade, the mean fall R-CBM score was 142.83 words read correctly with a standard deviation of 42.80. The winter mean R-CBM score was 158.24 with a standard deviation of 44.73. In spring, the mean R-CBM score was 174.52 words read correctly with a standard deviation of 45.90. Based on AIMSweb cut scores, these scores are all above the cut score indicating “Tier 1” or low risk of failure on state standards tests (AIMSweb, 2011). The mean ISAT scale score for sixth grade students was 243.88 points with a standard deviation of 33.79. This score is within the category of “Meets Standards” indicating that students at this level meet state standards for overall reading (Illinois State Board of Education, 2011). The mean fall MAP scale score was 211.83 points with a standard deviation of 16.57. This is very close to the mean score based on national norms of 211.6 (NWEA, 2011). The mean winter MAP scale score was 216.28 points with a standard deviation of 14.20. This is just above the mean score based on national norms of 213.8.

Table 9. Standardized Test Scores for Sixth Grade

	N	Mean score	SD
Fall R-CBM	357	142.83	42.80
Winter R-CBM	353	158.24	44.73
Spring R-CBM	352	174.52	45.90
ISAT	354	243.88	33.79
MAP fall	354	211.83	16.57
MAP winter	303	216.28	14.20

Research Question 1

In order to determine average initial scores, rates of growth, and variability, Hierarchical Linear Modeling (HLM) was conducted using the HLM 7 program from Scientific Software International. All students with R-CBM data were included in the analysis, although some individuals did not have three R-CBM data points. Therefore, cases with missing data were deleted when running the Level Two analyses.

Level One

The initial model for each grade level is represented by the equation:

$$WRC_{ti} = \pi_{0i} + \pi_{1i}WEEKS_{ti} + e_{ti}$$

In this equation, WRC_{ti} is the outcome variable representing R-CBM words read correctly, and π_{0i} is the initial score for person i . The variable for the rate of growth is π_{1i} , and the time variable is $WEEKS_{ti}$. In this case, the time variable represents the number of weeks elapsed since the initial score. The residual is represented by e_{ti} .

The initial model addresses the question of what is the average initial score and average rate of growth in R-CBM for students in fourth, fifth, and sixth grades. Separate analyses were run for each grade level. The estimated intercept terms are 109.69 for fourth grade, 128.33 for fifth grade, and 142.72 for sixth grade. The average growth rates are comparable at each grade level, with increases of 1.06, 1.08, and 1.07 words read correctly per week. The p-values for each grade level indicate that initial scores and rates of growth for each grade level are significantly different from zero.

Table 10. Fixed Effects for All Students

	Coefficient	S.E.	t-ratio	d.f.	p-value
Grade 4					
N= 362					
Initial Score	109.69	1.90	57.62	361	<0.01
Growth Rate	1.06	0.03	37.76	361	<0.01
Grade 5					
N= 339					
Initial Score	128.33	2.11	60.68	338	<0.01
Growth Rate	1.08	0.03	31.90	338	<0.01
Grade 6					
N= 354					
Initial Score	142.72	2.26	63.05	353	<0.01
Growth Rate	1.07	0.03	36.93	353	<0.01

The estimation of variance components details the variance in both initial status and slope across students at each grade level. The table below demonstrates that there is significant variance both in initial R-CBM scores and in rates of growth across students. Therefore, it is appropriate to investigate factors that make account for this variability in initial scores and growth rates.

Table 11. Estimation of Variance Components for All Students

Random Effect	Std. Dev.	Variance Comp.	d.f.	χ^2	p-value
Grade 4					
Initial Score	35.27	1243.92	361	6969.38	<0.01
Growth Rate	0.32	0.10	361	561.12	<0.01
Level 1	9.00	81.09			
Grade 5					
Initial Score	38.14	1454.76	338	8346.28	<0.01
Growth Rate	0.47	0.22	338	799.70	<0.01
Level 1	8.54	72.93			

Grade 6					
Initial Score	41.82	1749.33	353	9870.87	<0.01
Growth Rate	0.36	0.13	353	9870.87	<0.01
Level 1	8.78	77.02			

Level two

Next, the predictor of ELL status was added to the model. This was done to determine whether significant variance in initial scores or growth rates could be accounted for by ELL status. The equations that represent the level two models are as follows:

$$\pi_{0i} = \beta_{00} + \beta_{01}ELL_i + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}ELL_i + r_{1i}$$

In this equation, the variable ELL represents the presence or absence of ELL status (1= ELL, 0= non-ELL). The following table shows the average initial scores and rates of growth for non-ELL students, and the differences (positive for negative) in average scores for ELL students.

At each grade level, the coefficients for ELL students' initial status demonstrated that ELL students had significantly lower initial scores. In fourth grade, non-ELL students had predicted average scores of 119.77, while ELL students scored an average of 40.14 points lower. Fifth grade non-ELL students had predicted average initial scores of 137.28, with average scores 44.01 points lower for ELL students. In sixth grade, ELL students scored an average 56.82 points lower than their non-ELL peers' predicted scores of 152.66. Differences in slope were not significant at any grade level (p-values ranging

from 0.20 to 0.67), indicating that ELL status does not account for variance in rates of growth in R-CBM.

Table 12. Level Two All Students

Fixed Effects	Coefficient	S.E.	t-ratio	d.f.	p-value
Fourth Grade					
Initial Score					
Non-ELL	119.77	1.93	62.02	360	<0.01
ELL	-40.14	3.85	-10.42	360	<0.01
Growth Rate					
Non-ELL	1.08	0.03	33.31	360	<0.01
ELL	-0.08	0.06	-1.27	360	0.20
Fifth Grade					
Initial Score					
Non-ELL	137.28	2.11	64.98	337	<0.01
ELL	-44.01	4.69	-9.39	337	<0.01
Growth Rate					
Non-ELL	1.09	0.04	28.68	337	<0.01
ELL	-0.04	0.08	-0.04	337	0.67
Sixth Grade					
Initial Score					
Non-ELL	152.66	2.15	70.99	352	<0.01
ELL	-56.82	5.14	-11.05	352	<0.01
Growth Rate					
Non-ELL	1.08	0.03	34.09	352	<0.01
ELL	-0.09	0.08	-1.16	352	0.25

At each grade level, the model including ELL status as a predictor accounted for a relatively small amount of the variance in initial R-CBM scores and growth over time. The p-values for the model at each grade level remained <0.01 , indicating that a significant amount of variance remained. The percentage of variation in initial R-CBM scores accounted for by ELL status was calculated by subtracting the variance component associated with intercept in this model from the variance associated with the intercept in the level one model, and dividing that amount by the variance from the level one model. Results indicate that ELL status accounts for 24.21% of the variance in initial R-CBM scores at the fourth grade level, 21.36% of the variance at the 5th grade level, and 26.47% at the sixth grade level.

Research Question 2

Level One

An additional series of HLM analyses were conducted with only the ELL students at each grade level. The table below includes the estimated intercept terms and growth rates for students at each grade level. The estimated intercept term for ELL students in fourth grade was 79.64 words read correctly. The mean rate of growth for fourth grade ELL students was 1.00 words per week. At fifth grade, the estimated intercept term for ELL students was 93.27 words read correctly with a growth rate of 1.05 words per week. Sixth grade ELL students achieved an estimated intercept term of 95.83 points, with a rate of improvement of 1.00 words per week.

Table 13. Fixed Effects for ELL Students

	Coefficient	S.E.	t-ratio	d.f.	p-value
Grade 4					
N= 91					
Initial Score	79.64	3.63	21.96	90	<0.01
Growth Rate	1.00	0.05	19.80	90	<0.01
Grade 5					
N= 69					
Initial Score	93.27	4.14	22.51	68	<0.01
Growth Rate	1.05	0.08	13.95	68	<0.01
Grade 6					
N= 62					
Initial Score	95.83	4.28	22.41	61	<0.01
Growth Rate	1.00	0.08	13.18	61	<0.01

The estimation of variance components details the variance in both initial status and slope across ELL students at each grade level. The table below demonstrates that there is significant variance both in initial R-CBM scores and in rates of growth across students. Therefore, it is appropriate to investigate factors that make account for this variability in initial scores and growth rates.

Table 14. Estimation of Variance Components for ELL Students

Random effect	Std. Dev.	Variance Comp.	d.f.	χ^2	p-value
Grade 4					
Initial Score	33.77	1140.65	90	1936.92	<0.01
Growth Rate	0.29	0.08	90	139.77	<0.01
Level 1	8.13	66.08			
Grade 5					
Initial Score	33.69	1134.74	68	1656.64	<0.01
Growth Rate	0.51	0.26	68	203.77	<0.01
Level 1	7.59	57.61			

Grade 6					
Initial Score	32.42	1050.83	61	861.64	<0.01
Growth Rate	0.37	0.14	61	100.02	<0.01
Level 1	9.74	94.86			

Level Two

Predictors were added to the model to determine whether individual student variables would account for significant variance in initial R-CBM scores or growth rates for ELL students at each grade level. The predictors added to the model include ACCSCL, a measure of English language proficiency consisting of scale scores from the ACCESS for ELLs test, and the racial and ethnic categories of ASIAN and WHITE. Because the majority of the ELL students at each grade level were classified as Hispanic, that was used as the comparison group. Initial analyses were attempted using variables for home language, but these were removed from the model due to problems with multicollinearity. The level-2 model can be represented with the following equations:

$$\pi_{0i} = \beta_{00} + \beta_{01}ACCSCL_i + \beta_{02}ASIAN_i + \beta_{03}WHITE_i + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}ACCSCL_i + \beta_{12}ASIAN_i + \beta_{13}WHITE_i + r_{1i}$$

The first equation models the initial status, and the second models the slope, or rate of growth.

In fourth grade, a student in the Hispanic comparison group whose English proficiency score was at the average level would be expected to have an initial R-CBM score of 70.81 points and to improve at a rate of 1.00 words per week. For each point higher than the mean that a student achieved on the ACCESS test, they would be

predicted to have an initial R-CBM score that would be 1.33 words higher than the intercept store. Students who are Asian are predicted to perform 15.48 points higher on initial R-CBM scores than the Hispanic comparison group, and those who are white are predicted to perform 13.63 points higher. For each of these variables, p-values are below 0.05, indicating that the predictors are significant. Fourth grade students who are part of the Hispanic group and have average ACCESS scores are expected to improve at a rate of 1.00 words per week. None of the three predictors were significant in accounting for the variance in student growth rates for fourth grade ELL students.

At the fifth grade level, a student with an average language proficiency score and who is in the Hispanic comparison group would be expected to have an initial R-CBM score of 91.41 words read correctly. For each point higher students perform on the ACCESS test, they would be expected to have an R-CBM score 1.22 points higher than the intercept value. No significant differences were seen in the initial scores of students in the Asian or white ethnic groups. When examining growth rates, the expected value for students in the comparison group, with average ACCESS scores and in the Hispanic ethnic group, is 0.76 words read correctly per week. A significant positive relationship with regard to growth rate and language proficiency was seen at the fifth grade level. An ACCESS score higher than average would be predicted to be related to a rate of growth 0.01 words per week higher for each point on the ACCESS scale. Students coded as white are predicted to improve at a rate 0.39 words per week higher than those in the comparison group, and students in the Asian group are predicted to improve at a rate 0.71 words per week higher than students in the comparison group.

Table 15. Level Two for ELL Students

Fixed effects	Coefficient	S.E.	t-ratio	d.f.	p-value
Fourth Grade					
Intercept	70.81	4.30	16.49	87	<0.01
ACCSCL	1.33	0.18	7.44	87	<0.01
ASIAN	15.48	6.69	2.31	87	0.02
WHITE	13.63	6.51	2.09	87	0.04
Growth Rate	1.00	0.08	11.83	87	<0.01
ACCSCL	0.00	0.00	-0.10	87	0.92
ASIAN	-0.01	0.13	-0.08	87	0.94
WHITE	0.00	0.13	0.03	87	0.98
Fifth Grade					
Intercept	91.41	5.81	15.72	65	<0.01
ACCSCL	1.22	0.24	5.02	65	<0.01
ASIAN	3.40	8.22	0.41	65	0.68
WHITE	2.13	9.41	0.23	65	0.82
Growth Rate	0.76	0.11	7.23	65	<0.01
ACCSCL	0.01	0.00	3.24	65	<0.01
ASIAN	0.52	0.15	3.48	65	<0.01
WHITE	0.39	0.17	2.31	65	0.02
Sixth Grade					
Initial Score	88.85	4.85	18.32	58	<0.01
ACCSCL	1.24	0.23	5.63	58	<0.01
ASIAN	15.26	7.74	1.97	58	0.05
WHITE	11.48	9.60	1.20	58	0.24
Growth Rate	1.01	0.11	9.25	58	<0.01
ACCSCL	0.00	0.01	0.73	58	0.47
ASIAN	-0.10	0.17	-0.57	58	0.57
WHITE	0.13	0.22	0.65	58	0.52

At the sixth grade level, a student in the Hispanic comparison group with an ACCESS score at the mean level for the group would be expected to achieve an initial R-CBM score of 88.85 words read correctly. For each point higher than the mean ACCESS score a student achieved, they would be predicted to achieve an R-CBM score 1.24 points higher than the intercept value. Students in the Asian group are predicted to perform 15.26 words read correctly higher than the intercept score. Significant differences are not predicted for students in the white group. The growth rate for students in the comparison group in sixth grade is 1.01 words read correctly per week. None of the three predictors accounted for a significant amount of variance in the rates of growth.

At each grade level, a significant amount of variance remained. The percentage of variation in initial R-CBM scores in ELL students accounted for by language proficiency and racial and ethnic categories was calculated by subtracting the variance component associated with intercept in this model from the variance associated with the intercept in the level one model, and dividing that amount by the variance from the level one model. Results indicate that language proficiency and racial and ethnic status accounts for 51.19% of the variance in initial R-CBM scores at the fourth grade level, 27.11% of the variance at the 5th grade level, and 39.50% at the sixth grade level. At the fourth and sixth grade levels, none of the predictors significantly accounted for variance in growth rate. At the fifth grade level, the model included two significant predictors for growth rate and accounted for 39.44% of the variance in the rate of growth in R-CBM.

Research Question 3

Correlations

At each grade level, correlations were calculated between both fall and winter AIMSweb scores and the Illinois Standards Achievement Test (ISAT), which was administered in March. Preliminary analyses determined that there were no significant outliers and no violations of the assumptions of linearity, normality, and homoscedasticity. At the sixth grade level, correlations between the fall and winter Measures of Academic Progress (MAP) assessment and the ISAT were calculated as well. Correlations were calculated for the overall group at each grade level, the group of ELL students, and the group of non-ELL students.

At the fourth grade level, Pearson correlations were significant at the 0.01 level (two-tailed) for all groups of students in both fall and winter. There were strong correlations between both fall and winter R-CBM scores and ISAT scores for all three groups. For the group of ELL students, there was a correlation of 0.73 between fall AIMSweb scores and ISAT, and 0.74 between winter AIMSweb scores and ISAT. Non-ELL students had a correlation of 0.58 between both fall AIMSweb and winter AIMSweb scores and ISAT. The overall group of fourth grade student demonstrated correlations of 0.71 and 0.72 between fall and winter AIMSweb and ISAT, respectively.

At the fifth grade level, moderate correlations were seen between fall and winter R-CBM scores and ISAT for ELL students, and strong correlations were seen for non-ELL students and the overall fifth grade group. For ELL students, the correlation between fall R-CBM scores and ISAT was 0.44, and 0.49 for winter R-CBM and ISAT. For non-

ELL students, the correlations were 0.56 and 0.58 for fall and winter, respectively. For the overall fifth grade group, correlations of 0.64 and 0.65 were seen for fall and winter.

Table 16. Pearson Correlations

	Fall R-CBM * ISAT	Winter R-CBM * ISAT	Fall MAP * ISAT	Winter MAP * ISAT
Fourth Grade				
ELL	0.73**	0.74**		
N	90	91		
Non-ELL	0.58**	0.58**		
N	270	271		
All	0.71**	0.72**		
N	360	362		
Fifth Grade				
ELL	0.44**	0.49**		
N	66	64		
Non-ELL	0.56**	0.58**		
N	269	270		
All	0.64**	0.65**		
N	335	334		
Sixth Grade				
ELL	0.68**	0.68**	0.70**	0.70**
N	59	59	60	48
Non-ELL	0.61**	0.62**	0.69**	0.78**
N	291	290	287	252
All	0.72**	0.73**	0.78**	0.83**
N	350	349	347	300

** Correlation is significant at the 0.01 level (2-tailed).

At the sixth grade level, there were strong correlations between all four screening measures (fall and winter R-CBM, fall and winter MAP) and ISAT. For ELL students, correlations between fall and winter R-CBM and ISAT were both 0.68, and correlations between fall and winter MAP and ISAT were both 0.70. Non-ELL sixth grade students had correlations of 0.61 for fall R-CBM and ISAT, and 0.62 for winter R-CBM and ISAT. Non-ELL sixth grade students had correlations of 0.69 (fall) and 0.70 (winter) between MAP and ISAT. Overall, sixth grade students had correlations of 0.72 for fall R-CBM and ISAT and 0.73 for winter R-CBM and ISAT. When analyzing MAP and ISAT, the overall sixth grade group demonstrated correlations of 0.78 in the fall and 0.83 in the winter.

Cross Tabulations

One of the research questions in this study addressed the utility of cut scores association with the R-CBM measure in predicting success on a state standards test. To address this question, frequently used cut-scores for R-CBM were used to sort students into categories based on their scores in fall and winter. These cut scores were used to determine whether students were classified as Tier 1 (Low risk of failure) or Tier 2/3 (Some risk/ At-Risk). Students' classification categories based on ISAT scores were used in this analysis as well. Students who received scores that placed them within the Academic Warning or Below Standards classifications were coded as Below Standards, and those who received scores in the Meets Standards or Exceeds Standards were coded as Meets Standards. A cross tabulation analysis was conducted using classifications from fall and winter R-CBM scores to determine the sensitivity, specificity, negative predictive

power, positive predictive power, and overall correct classification for each group at each time period. The tables below show the numbers of students who were placed in each of the categories in both fall and winter, based on ISAT and R-CBM scores.

Table 17. Cross Tabulations

		Fall		Winter	
		Below Standards	Meets Standards	Below Standards	Meets Standards
Fourth Grade					
ELL	Tier 2/3	45	25	44	26
	Tier 1	1	20	2	19
Non-ELL	Tier 2/3	21	67	20	63
	Tier 1	6	177	7	181
All	Tier 2/3	66	92	64	89
	Tier 1	7	197	9	200
Fifth Grade					
ELL	Tier 2/3	32	14	33	12
	Tier 1	11	9	12	11
Non-ELL	Tier 2/3	18	48	18	42
	Tier 1	12	191	12	198
All	Tier 2/3	50	62	51	54
	Tier 1	23	200	20	209
Sixth Grade					
ELL	Tier 2/3	32*	20*	31*	20*
	Tier 1	0*	7*	1*	7*
Non-ELL	Tier 2/3	11	82	10	70
	Tier 1	1	197	2	208
All	Tier 2/3	43	102	41	90
	Tier 1	1	204	2	215

* For these groups, 50% of cells had an expected count of less than 5.

Values were calculated for sensitivity, specificity, negative predictive power, positive predictive power, and overall correct classification for each group of students in grades four, five, and six in fall and winter. The following table shows the interpretation of the values above that was used in these calculations.

Table 18. Cross Tabulation Interpretation

	Below Standards (Condition Positive)	Meets Standards (Condition Negative)
Test 2/3 (Test Positive)	True positive (a)	False positive (b)
Tier 1 (Test Negative)	False negative (c)	True negative (d)

Sensitivity was calculated with the formula $\frac{a}{a+c}$, indicating the probability of receiving a positive (Tier 2/3) score on the screening test, given a score below the target on the ISAT test. Specificity is a calculation of the likelihood of a negative score on the screening test (Tier 1), given a student who meets standards on the state test. The formula for specificity is $\frac{d}{b+d}$. Positive predictive power is the likelihood that someone who scores at the Tier 2/3 level (test positive) will receive a score in the below standards level on ISAT. The formula for positive predictive power is $\frac{a}{a+b}$. Negative predictive power is the likelihood that someone who scores at the Tier 1 level (test negative) on R-CBM will receive a score on ISAT in the meets standards level. The formula is $\frac{d}{c+d}$. Overall correct classification is the likelihood of correct classification, expressed by the total number of students with true positive and true negative classifications divided by the total number of

students in the group. The equation is $\frac{a+d}{N}$. The following table shows the five values described above.

Table 19. Diagnostic Accuracy

	Sensitivity	Specificity	Positive Predictive Power	Negative Predictive Power	Correct Classification
Fourth Grade					
Fall					
ELL	0.98	0.44	0.64	0.95	0.71
Non-ELL	0.78	0.73	0.24	0.97	0.73
All	0.91	0.68	0.42	0.97	0.72
Winter					
ELL	0.96	0.42	0.63	0.91	0.69
Non-ELL	0.74	0.74	0.24	0.96	0.74
All	0.88	0.69	0.42	0.96	0.73
Fifth Grade					
Fall					
ELL	0.74	0.39	0.45	0.70	0.62
Non-ELL	0.60	0.80	0.27	0.94	0.74
All	0.68	0.76	0.45	0.90	0.75
Winter					
ELL	0.80	0.48	0.58	0.73	0.69
Non-ELL	0.60	0.83	0.30	0.94	0.80
All	0.72	0.79	0.48	0.91	0.78
Sixth Grade					
Fall					
ELL	1.00	0.26	0.62	1.00	0.66
Non-ELL	0.92	0.71	0.12	0.99	0.71
All	0.98	0.67	0.30	0.99	0.71
Winter					
ELL	0.97	0.26	0.61	0.89	0.54
Non-ELL	0.83	0.75	0.13	0.99	0.75
All	0.93	0.71	0.31	0.99	0.73

Sensitivity ranged from 60% to 100%, depending on the group, time period, and grade level. This indicates that 60 to 100% percent of students who scored below standards on ISAT received a Tier 2/3 score on the screening measure. Specificity was lower, ranging from 26% to 83%, indicating that between 26% and 83% of students who met standards on ISAT scored within the Tier 1 level on R-CBM. Positive predictive power is the probability of scoring in the below standards range on ISAT, given a Tier 2/3 score on R-CBM, and results ranged from 12% to 64%. Negative predictive power indicated the probability of scoring within the “meets standards” rang on ISAT given a Tier 1 score on R-CBM, and those values varied from 70% to 100%. The overall correct classification ranged from 54% to 80%, which gives the likelihood that a student will be correctly classified by the R-CBM assessment in fall or winter.

Research Question 4

In order to determine whether R-CBM or MAP was a stronger predictor of performance on the ISAT test for sixth grade students, the correlations described above were compared. For ELL students in fall and winter, the Pearson correlation coefficient for R-CBM was 0.68, compared to 0.70 for MAP. For non-ELL students, the correlation coefficient for MAP was larger than that of R-CBM for both fall (0.61 and 0.69) and winter (0.62 and 0.78). Correlations for all students were slightly larger for MAP when compared to R-CBM in both fall (0.72 and 0.78) and winter (0.73 and 0.83). Additional calculations were conducted in order to determine whether or not these differences were significant.

First, the r values for the correlation coefficients were converted to z -values. The following equation was used to calculate the significance of the difference in correlation in each of these pairs:

$$z_{obs} = \frac{z_1 - z_2}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}}$$

The resulting z -scores for each of the pairs of scores are reported below. Z -scores between -1.96 and 1.96 may be due to chance and are not significant.

Table 20. Comparison of Correlations

	z_1 (R-CBM)	z_2 (MAP)	N_1	N_2	Z_{obs}
ELL Fall	0.83	0.87	59	60	-0.20
ELL Winter	0.83	0.87	59	48	-0.19
Non-ELL Fall	0.71	0.85	291	287	1.67
Non-ELL Winter	0.73	1.05	290	252	-3.69*
All Fall	0.91	1.05	350	347	-1.91
All Winter	0.93	1.19	349	300	-3.28*

* Significant at the 0.05 level

The only significant differences in correlation were found with non-ELL students and all students at the winter testing period. At that point, MAP scores had a significantly stronger relationship to ISAT scores than did R-CBM for non-ELL students and all students. Significant correlations were found in the relationship between MAP and ISAT and R-CBM and ISAT for all groups in fall and winter, but no significant differences were found in the strength of the correlation for the ELL only group at either time period.

Also, no significant differences were found in the strength of the correlations when comparing each group at the fall testing period.

CHAPTER FIVE

DISCUSSION

The purpose of this study was to examine the use of curriculum-based measurement in reading (R-CBM) as a tool for screening and measuring progress over time for English Language Learners (ELLs) at the intermediate elementary level. The sample consisted of 1,056 students in grades four through six in an elementary district in a suburb of a large, Midwestern city. Students at each grade level were administered R-CBM in fall and winter, and spring, and a state standards test in reading in the spring. Sixth grade students also were administered another formative reading assessment, the Measure of Academic Progress (MAP), in the fall and winter. Data analysis was conducted in order to determine average initial R-CBM scores and rates of progress for all students at each grade level and whether there were significant differences based on ELL status. For ELL students, the current study also investigated whether initial scores or rate of progress varied based on student demographic factors. Predictive validity of R-CBM with relation to a state standards test was investigated for students at each grade level, considering ELL status. Finally, for sixth grade students, the predictive validity of R-CBM and MAP was compared with relation to the state standards test for ELL and non-ELL students.

Research Question 1

Hierarchical Linear Modeling was used to determine average initial (fall) R-CBM scores for all students at each grade level, and average rates of growth. For fourth through sixth grades, average initial R-CBM scores were 109.69, 128.33, and 142.72 words read correctly per minute respectively. The average growth rates are comparable at each grade level, with weekly increases of words read correctly of 1.06 for fourth grade, 1.08 for fifth grade, and 1.07 for sixth grade. Based on AIMSweb default cut scores (AIMSweb, 2011), the average initial scores are within the Tier 1 category, which indicates an 80 percent likelihood of passing state standards tests in reading. Rates of growth are also relatively consistent with AIMSweb norms in fourth and fifth grade, and higher than average at the sixth grade level (AIMSweb, n.d.). For a student performing within the average range, a rate of improvement from fall to spring of 1.06 words read correctly per week at the fourth grade level would correspond to the 65th percentile. Similarly, a rate of improvement of 1.08 words read correctly per week would correspond to the 65th percentile at the fifth grade level. At the sixth grade level, a rate of improvement of 1.07 words read correctly per week would correspond to the 85th percentile. However, these norms are intended for use when interpreting individual student scores and are based on individual students initial fall scores. They may be less meaningful when interpreting mean rates of improvement for a group of students.

Further analysis determined that ELL status was a significant predictor of initial R-CBM scores in all three grades, accounting for between 21.36 and 26.47 percent of the variability in initial scores depending on grade level. ELL status was not determined to be

a significant predictor of variability in growth rates for any grade level. ELL students' initial R-CBM scores were significantly lower than non-ELL peers at each grade level. In fourth grade, ELL students achieved an average fall R-CBM score of 79.63, which is within the Tier 2 range based on AIMSweb default cut scores (AIMSweb, 2011), suggesting a 50 percent likelihood of passing the state standards test. Fourth grade non-ELL students achieved average scores of 119.77 words per minute, which is well within the Tier 1 category. In fifth grade, ELL students had average scores of 93.27 words per minute, which was within the Tier 2 classification. Non-ELL fifth graders had average scores of 137.28 words per minute, within Tier 1. In sixth grade, ELL students had average scores of 95.84 words per minute, which is within the Tier 3 classification, indicating less than 50 percent likelihood of passing a state standards test in reading. Non-ELL sixth graders achieved a mean score of 152.66 words per minute, within Tier 1.

These discrepancies between ELL and non-ELL students are consistent with findings from previous research. ELL students have consistently been found to perform at lower levels on measures of overall reading than non-ELL students (August & Shanahan, 2006; Kieffer, 2008). Results regarding rates of improvement have been inconsistent. Farmer, Swanlund, and Pluymert (2010) found inconsistent results when comparing rates of improvement depending on grade level. While ELL students in first grade had higher rates of improvement than non-ELL peers, second grade ELL students had lower rates of improvement, and third grade students did not demonstrate significant differences based on ELL status. Dominguez de Ramirez and Shapiro (2006) found rates of improvement were lower among ELL students in the primary grades when compared to the overall

population. Stokes (2010) found that differences in rates of improvement in sixth grade students were not significantly related to ELL status. The findings related to rate of improvement are consistent with the current study, and may suggest that that at higher elementary grade levels, rates of growth are not related to ELL status. Fuchs, Fuchs, Hamlett, Walz & Germann (1993) found that for all students, growth in R-CBM slows throughout the elementary grades, and as Stage and Jacobsen (2001) also note, may asymptote around fourth grade. It may be that because growth is generally lower in the upper elementary grades, differences in rates of growth across student groups become less significant.

Research Question 2

The second research question concerned initial R-CBM scores and rates of growth for ELL students in fourth through sixth grade. HLM analysis was used to determine whether demographic factors such as ethnicity and level of language proficiency would account for a significant amount of variance in R-CBM scores and rates of growth. Home language was intended to be used as a predictor, but that factor was eliminated due to problems with multicollinearity. The variables for home language were highly correlated with the variable for race and ethnicity. Within the sample, there was one hundred percent consistency between the ethnic category of Hispanic and the home language of Spanish. Therefore, only racial and ethnic categories were used. In the case of the Hispanic comparison group, this group can be interpreted to stand in for the home language category of Spanish. The Asian and white groups represented a variety of home

languages and it is difficult to draw conclusions about home languages in those racial and ethnic categories.

Initial R-CBM Scores

At all three grade levels, language proficiency accounted for a significant amount of variability in initial R-CBM scores. Overall, ELL students with higher language proficiency scores are predicted to have higher initial R-CBM scores. At the fourth grade level, the racial and ethnic categories of “White” and “Asian” accounted for a significant amount of variance in initial scores and those students were predicted to have higher initial scores than those in the comparison group (primarily Hispanic). Among fifth grade ELL students, no significant differences were found when comparing the initial R-CBM scores across ethnic groups. At the sixth grade level, the predictor “Asian” accounted for significant variability in initial scores and students who were identified as Asian had higher overall initial R-CBM scores. Looking at the results across grade levels, the only consistent finding is that ELL students with higher language proficiency perform at higher levels on their initial R-CBM assessment in the fall. This finding is consistent with previous research (Farmer, Swanlund, & Pluymert, 2010) that found scores on the ACCESS for ELLs test of language proficiency to be highly correlated with R-CBM scores.

The relationship between racial and ethnic categories and initial R-CBM scores is inconsistent across grade levels, but in two out of three grade levels, results indicate that Asian students would be predicted to have higher initial scores than those in the comparison group. In this sample, the majority of students in the comparison group are

Hispanic, and Spanish is the home language spoken by these students. The Asian ethnic group in this sample consists of students from a variety of home language groups, and varies across grade levels. Languages based in the Indian subcontinent comprise three of the top five home languages among ELL students (Gujarati, Malayalam, and Urdu). This suggests that ELL students who identify as Asian, particularly those whose families are from the Indian subcontinent, may have higher initial R-CBM scores than those who identify as Hispanic. However, this relationship was not consistent across grade levels.

Previous studies examined the predictive validity of R-CBM for students from different racial and ethnic backgrounds. Kranzler, Miller, and Jordan (1999) found predictive bias when comparing white and African American students' performance in grades five and six. However, Hintze, Callahan, et al. (2002) did not find significant predictive bias when comparing ethnic groups given a similar sample. Pearce and Gayle (2009) found that, when comparing American Indian and white students, a higher percentage of students performed adequately on R-CBM and did not meet standards on a state test. These studies, and the current findings, suggest that there may be differences within the ELL subgroup, depending on race.

The current sample does not allow for the use of home language as a predictor, but this is a possibility for further investigation. In a related study, Muyskens, Betts, Marston, and Lau (2009) analyzed predictive validity of a CBM reading assessment with respect to a state standards test for ELL students in fifth grade, and did not find significant differences across home language with a large sample that included students whose home languages were Spanish, Somali, and Hmong. While the current study did

not examine predictive validity with respect to home languages, the inconsistent relationship between ethnic groups in initial R-CBM scores may suggest an area for future research. Another significant study found that when comparing groups of ELL students, no significant differences were found when comparing language groups (Spanish and Somali), but that the length of time a student's family had been in the United States accounted for a significant amount of variability in reading achievement scores (Betts, Bolt, Decker, Muyskens, & Marston, 2009). Length of time in the US was not examined in the current study, but should be included in future modeling of the predictive validity of CBM among students from a variety of ethnic groups.

Rates of Growth

At the fourth and sixth grade level, neither language proficiency nor ethnicity significantly accounted for variability in rates of growth in R-CBM scores. However, at the fifth grade level, language proficiency and the categories of White and Asian accounted for significant variability in the rates of growth. There was a positive relationship between language proficiency scores and rates of growth in R-CBM, indicating that those students with higher scores are predicted demonstrate more growth in R-CBM over the course of a school year. Also, there was a positive relationship between both racial and ethnic categories and rate of growth, indicating that those fifth grade students in the White and Asian categories had overall higher rates of growth in R-CBM than those in the comparison (Hispanic) group.

The relationships between language proficiency, racial and ethnic categories, and rates of growth in R-CBM scores were inconsistent across grade levels. While at the fifth

grade level, these predictors account for a significant amount of variance in rates of growth in R-CBM, the same relationships were not seen at the fourth and sixth grade levels. Therefore, it is difficult to draw any conclusions about whether or not different rates of growth should be expected from ELL students depending on these demographic variables.

Research Question 3

The next research question addressed the predictive validity of R-CBM with respect to a state standards test in reading for students in grades four through six, while considering ELL status. In order to answer this question, correlations between R-CBM and the ISAT assessment were investigated for each group at each grade level. In addition, cut scores provided by AIMSweb were examined with respect to passing status on the ISAT and diagnostic accuracy of these cut scores was examined.

Correlations

Across groups and grade levels, moderate to strong correlations were found between fall and winter R-CBM scores and ISAT reading scores from the March administration, for ELL and non-ELL students and each grade level overall. Correlations ranged from 0.44 to 0.74. This is consistent with previous research in which moderate to strong correlations have been found between R-CBM and state standards tests in reading (Good, Simmons, and Kame'enui, 2001; Hintze & Silbergitt, 2005; McGlinchey & Hixon, 2004, Sibley, Biwer, & Hesch, 2001; Stage & Jacobson, 2001; Wiley & Deno, 2005). This finding supports the use of R-CBM as a screening tool when administered in

fall or winter with ELL and non-ELL students in grades four through six. Given the diversity of the sample in the current study, this is an important finding.

Diagnostic Accuracy

In order to further examine the predictive validity of R-CBM with respect to a state standards test of reading, AIMSweb default cut scores were used to sort students into two groups. Students were classified as either Tier 1 or Tier 2/3 based on these cut scores at each testing period (fall and winter). Students were also classified on the basis of their performance on the ISAT as either meeting standards or below standards. Based on these categories, sensitivity, specificity, negative predictive power, positive predictive power, and overall correct classification were calculated for ELL students, non-ELL students, and all students at each grade level.

Curriculum-based measurement, used in this context, is meant to be a screening measure. As such, students who are identified as Tier 2/3 based on fall or winter R-CBM scores are often provided with additional intervention in reading. In the current study, the question of whether or not students received intervention and the type of intervention that was provided was not available. However, it is reasonable to expect that some number of students who are identified as Tier 2/3 in fall or winter did, in fact, receive intervention. For some of these students, the intervention may make it possible for students who were previously at-risk to meet standards on the state test in the spring. For this reason, the current study will focus primarily on two values: sensitivity and negative predictive power. Lower values for positive predictive power are acceptable, as this may reflect that

some of the students initially identified as Tier 2 or Tier 3 made improvement throughout the school year.

Values for sensitivity are significant, as it is important to find a measure that will accurately identify most or all of the students who are at risk of failure on the ISAT. Looking at each grade level overall, values for sensitivity indicated that the majority of students who did not meet standards on the ISAT were identified. However, among non-ELL students at the fifth grade level in fall and winter, the probability of having a Tier 2/3 score given not meeting standards on the ISAT was only 60%. That means that 40% of students in this group who do not meet standards on the ISAT may not have been identified as at-risk based on their fall or winter R-CBM scores.

Another important indicator when looking at diagnostic accuracy is negative predictive power. It is significant to look at the likelihood that someone who scores at the Tier 1 level on R-CBM will meet standards on the state test. It is important that students at the Tier 1 level do in fact pass the state standards test, because as stated above, these students may not receive additional intervention. Values at most grade levels and time periods were at 89% and above, indicating that for most groups, the vast majority of students identified as Tier 1 did, in fact, pass the state standards tests in reading. However, for ELL students in fifth grade, only 70-73% of students who were identified as Tier 1 met standards on the state test. Again, this indicates that there may have been a significant percentage of students within this group who were not identified as at-risk and therefore may not have received appropriate interventions in reading.

A number of previous studies have investigated the diagnostic accuracy of cut scores for R-CBM with relation to other standardized reading tests. Stage and Jacobsen (2001) conducted an analysis of growth curves and determined cut scores based on these growth curves for fourth grade students. They found that the positive predictive power (indicating that lower fall scores predicted failure on the state test) was 0.41 and that the negative predictive power (that higher fall scores predicted success on the state test) was 0.90. This is relatively consistent with results in the current study at the fourth and sixth grade levels. In another study, Hintze and Silbergliitt (2005) used four different methods to determine cut scores for R-CBM for students in third grades and found that each method yielded cut scores that met their criteria for diagnostic accuracy. They noted that all methods yielded higher levels for specificity and negative predictive power than for sensitivity and positive predictive power, in an effort to maximize the number of “true negatives” (i.e. those who would be correctly predicted to pass the state test). Values for sensitivity were, therefore, lower than those found in the current study (ranging from 0.63 to 0.79). Values for negative predictive power ranged from 0.52 to 0.88, and therefore were somewhat lower than those found in the current study as well. Cabrera (2006) examined the predictive validity of R-CBM with regard to a state standards test for ELL students in sixth grade, and found values for sensitivity (0.83) and negative predictive power (0.74) that are consistent with the current study.

Overall, results for diagnostic accuracy are relatively consistent with previous research and indicate that published AIMSweb cut scores are generally adequate for use as a screening tool for students in grades four through six regardless of ELL status.

Again, this is a significant finding due to the diversity of the sample in the current study. At the same time, within the fifth grade group, some of the findings were inconsistent and suggest that significant numbers of at-risk students might not be correctly identified.

Research Question 4

The final research question was intended to compare the predictive validity of two screening measures of reading. R-CBM and MAP are both measures of reading that are administered to screen students in sixth grade in fall and winter. The two measures were compared with regard to their correlations with the ISAT test administered in March. In the fall, no significant differences were found in the correlations of MAP with ISAT and R-CBM with ISAT. In the winter, no significant differences were found in the correlations of MAP with ISAT and R-CBM with ISAT for ELL students in sixth grade. For non-ELL students and sixth graders overall in winter, MAP had a significantly stronger correlation to ISAT than R-CBM. Given the focus in the current study on the ELL population, this suggests that R-CBM has equivalent predictive validity with regard to ISAT when compared to MAP.

Andren (2010) compared the predictive validity of R-CBM and MAP for students in third grade with relation to a state standards test in reading using regression analyses and found that MAP was the best predictor of state scores for the general population and for students who were receiving reading support. This is consistent with the findings in the current study for the overall population at the winter test period. The sample used in Andren's study was 96% Caucasian and no data is provided regarding the number of students who are English Language Learners in the sample. While performance on the

MAP test appears to be significantly correlated to ISAT for students in all groups, data does not suggest that it is a better predictor of performance than R-CBM for ELL students.

Implications for Practice

The current study adds support to existing research indicating that ELL students perform at levels significantly below their peers on standardized reading assessments (Keiffer, 2008; August & Shanahan, 2006). The ELL students in this study had initial scores on screening measures that were significantly below that of same age peers, and did not demonstrate significant differences in rates of growth. This indicates that ELL students continue to lag behind peers. There is no data in this study to indicate that ELL students are making significant enough gains to catch up to their non-ELL peers. Previous research regarding ELL students indicates that word reading skills initially lag behind peers but catch up by the upper elementary years, but that comprehension skills continue to lag (August & Shanahan, 2006). The fact that R-CBM scores continue to lag behind peers is an indication that R-CBM continues to function as a measure of automaticity, or a general outcome measure, and is not simply a measure of word reading (Shinn & Shinn, 2002a). Because ELL students lag behind non-ELL peers in overall reading, intervention is needed in order to help ELL students make significant gains in reading skills at the intermediate elementary level.

Overall, results support the use of R-CBM for screening purposes and for progress monitoring. All groups at each grade level demonstrated significant progress over the course of the year. Rates of progress did not vary significantly based on ELL

status. This suggests that while initial scores would be expected to vary depending on ELL status and other factors, teachers could reasonably expect students to make similar rates of growth, regardless of ELL status. This is important to consider when setting goals for individual students or groups of students who are receiving intervention in reading.

The current study also supports the use of R-CBM as a screening tool for identifying ELL students who are in need of additional intervention. School psychologists are often called upon to participate in school-wide benchmark assessment as part of a three-tier problem solving process under Response to Intervention (RTI) (Shinn, 2008). Based on the results of this study, R-CBM appears to be moderately to strongly correlated to state standards tests, and published cut scores have good diagnostic accuracy for most groups of students in grades four through six. It is important to note, however, that there were some groups in this study where significant numbers of students who were at risk of failure might have been overlooked. This supports the idea that R-CBM should always be one data point used in decision-making, and was not designed to make predictions or diagnoses (Shinn, 2008). Additional data should always be used in making decisions for students so that students who are at risk of failure are not overlooked, and also so that students who perform adequately do not receive unnecessary interventions that use resources.

School psychologists also frequently use R-CBM data as part of the RTI process when they are asked to identify students with specific learning disabilities. When using R-CBM data for identification of students for special education, the results of this study were somewhat inconclusive due to questions raised by the make-up of the sample in the

study. In general, when making decisions about special education identification for students who are suspected of having a specific learning disability, it is important to compare a student to comparable peers. The Illinois State Board of Education (2012) recommends using disaggregated norms when assessing students who are ELLs, and being sure that the norming sample includes “like peers” (p. 5), which is defined as students “students from the same linguistic and socio-cultural background with similar exposure to the curriculum” (p.5). Given this recommendation, it would appear that it would be beneficial to compare ELL students suspected of a disability to the group of ELL students, rather than to the overall student body. However, in this sample, there is an overrepresentation of students who receive special education services within the ELL group. This could be due to overidentification of ELL students as disabled, or there could be another reason for the larger percentages of students with disabilities within the ELL population. Because of these discrepancies, it is important to stress the importance of using multiple sources of data in the problem-solving process rather than relying on R-CBM data alone, particularly with ELL students.

Limitations

In this study, existing data was used for analysis. R-CBM measures were administered by classroom teachers as part of the regular assessment schedule for the district. The 2010-2011 school year was the third year that this assessment was used consistently across the district. Individual schools in the district were responsible for training teachers in R-CBM administration, scoring, and interpretation. No data is available regarding the accuracy of implementation of R-CBM measures or whether

inter-rater reliability was established as part of the training for teachers who were assessing their students.

Another limitation of the current study was that the sample did not allow for the analysis of home language among ELL students. Home language was initially included as a predictor in the HLM analysis, but was removed due to problems with multicollinearity. For this reason, comparisons could not be made between students based on their home languages. The sample in the current study has a great deal of linguistic diversity, which makes it unique when compared to other studies relating to ELL students and R-CBM. However, the linguistic diversity made it difficult to make direct comparisons across home languages.

In this sample, the percentage of students within the ELL group who receive special education services is significantly higher than in the overall sample at each grade level. This could be due to over-identification of students who are ELLs for special education services. In Illinois, students must receive a certain score on the ACCESS for ELLs test of language proficiency in order to be considered English proficient and dismissed from ELL services. It is possible that students who receive special education services may have difficulty achieving the required score on the ACCESS assessment and may continue to qualify as ELL students under state policies. It is possible that the larger number of special education students in the sample used in this study had an impact on the results.

The current study did not look at whether or not students who received lower scores on R-CBM received reading intervention, and if so, the type of intervention,

frequency, and intensity of the intervention. It is reasonable to expect that some, if not all, students who were identified as at-risk based on fall or winter R-CBM scores received some sort of intervention. Students who received intervention may have made more progress than peers who did not receive the same. It is also unknown what type of intervention students received, and whether ELL students received comparable interventions to their non-ELL peers.

Other factors that may have contributed to initial scores and rates of growth for English Language Learners include length of time in the country, type of ELL services received (bilingual or monolingual English support) and other demographic factors such as socioeconomic status and parent education level. This data was not available in the current study, and was therefore not included in the analyses. Previous research indicates that socioeconomic status and length of time in the country may be significant variables that relate to reading performance (Betts, Bolt, Decker, et al., 2009; Baker, Smolkowski, Katz, et al, 2008).

Implications for Future Research

The current study used a sample of ELL students that was diverse, and included students from many ethnic and linguistic backgrounds. This is a strength of the current study, and provides support that previous research can be generalized across racial, ethnic, and linguistic groups. Previous research on ELL students and R-CBM has been relatively limited to samples that were primarily Spanish speaking, although limited studies also included Hmong and Somali-speaking students. Additional research including students from a wide variety of linguistic and ethnic backgrounds, using larger

sample sizes, would be beneficial. In this study, there was some indication that Asian ELL students have lower initial scores than the comparison group (primarily Hispanic). Additional research could be conducted to determine whether or not there are differences in predictive validity across racial and ethnic groups and across linguistic groups. Also, additional research could investigate other variable that may have an impact on initial R-CBM scores, rates of growth, and predictive validity of R-CBM with respect to state standards tests. These factors include length of time in the country, socioeconomic status, parent education level, intervention received, and type of ELL service received.

Additional research could also be conducted with a similar sample to develop cut scores that may provide better diagnostic accuracy. Although overall, R-CBM was found to provide adequate predictive validity with respect to a state standards test, there were some discrepancies across groups and grade levels. The current study was focused on one academic year in one school district, and was not concerned with developing local norms or cut scores based on local norms. Future research with a similar sample could be conducted using multiple years of data and could involve investigate the development of cut scores that would better meet the needs of the particular sample.

Summary

Teachers, administrators, and school psychologists are being called upon to meet the needs of all students in their schools even as diversity increases. One promising approach to meeting the needs of all students involves using a problem-solving model to screen and identify students at risk of academic failure and monitor progress over time using R-CBM. The current study adds support to previous research indicating that R-

CBM is a reliable and valid tool for screening and monitoring progress for ELL students in the intermediate grades. While results support previous research indicating that ELL students are performing at lower levels in overall reading, the current study also reinforces the idea that ELL students cannot be thought of as a unitary group. There are differences in language proficiency and race and ethnicity that may impact students' initial reading performance and rates of growth over time. Results varied across grade levels, making it difficult to draw conclusions about the relationship between language proficiency, race and ethnicity and reading performance over time.

R-CBM is also used as a screening tool to help identify students who are at-risk of failure on state standards tests. The results of this study provide additional support indicating that R-CBM has adequate predictive validity with relation to high stakes tests, even given a very diverse sample with a significant percentage of ELL students. However, some discrepancies between groups and grade levels reinforces the need to use additional data in decision-making rather than relying solely on scores from any single test when making decisions for individual students.

REFERENCES

- Adams, M. J. (1990). *Beginning to Read: Thinking and Learning about Print*. Cambridge, MA: MIT Press.
- AIMSweb (2011). *AIMSweb default cut scores explained*. Pearson Education, Inc. Retrieved April 15, 2012 from https://aimsweb.pearson.com/downloads/AIMSweb_Default_Cut_Score_Guide.pdf
- AIMSweb (n.d.). *AIMSweb ROI Growth Norms*. Retrieved January 15, 2013 from http://www.aimsweb.com/wp-content/uploads/ROI_Norm_Tables1.pdf
- Anderson, K. (2000). The reading wars: Understanding the debate over how best to teach children to read. *Los Angeles Times*, June 18, 2000.
- Andren, K.J. (2010). *An analysis of the concurrent and predictive validity of curriculum-based measures (CBM), the measures of academic progress (MAP), and the New England Common Assessment Program (NECAP) for reading* (Doctoral Dissertation). Retrieve from ProQuest. (UMI number 3422255).
- August, D. & Shanahan, T. (2006). Introduction and methodology. In D. August and T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth* (pp. 1-42). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Baker, Plasencia-Peinado & Lezcano-Lytle (1998). The use of curriculum-based measurement with language-minority students. In M. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp.175-213). New York, NY: Guilford Press.
- Baker, S.K., Smolkowski, K., Katz, R., Fien, H., Seeley, R., Kame'enui, E.J., & Beck, C.T. (2008). Reading fluency as a predictor of reading proficiency in low-performing, high-poverty schools. *School Psychology Review*, 37(1), 18-37.
- Batsche, G., Elliott, J., Graden, J.L, Grimes, J., Kovaleski, J.F., Prasse, D., Reschly, D.J., Schrag, J., & Tilly, W.D. (2005). *Response to Intervention: Policy considerations and implementation*. Alexandria, VA: National Association of State Directors of Special Education, Inc.

- Betts, J., Bolt, S., Decker, D., Muyskens, P., & Marston, D. (2009). Examining the role of time and language type in reading development for English Language Learners. *Journal of School Psychology, 47*, 143-166.
- Betts, J., Reschly, A., Pickart, M., Heistad, D., Sheran, C., Marston, D. (2008). An examination of predictive bias for second grade reading outcomes from measures of early literacy skills in kindergarten with respect to English-language learners and ethnic subgroups. *School Psychology Quarterly, 23*(4), 553-570
- Cabrera, C. (2008). *Curriculum-based measurement as a predictor for California state standards tests* (Master's thesis). Retrieved from ProQuest. (UMI Number 1460337).
- Capps, R., Fix, M., Murray, J., Ost, J., Passel, J.S., & Herwanto, S. (2005). *The new demography of America's schools: Immigration and the No Child Left Behind Act*. Washington, D.C.: The Urban Institute.
- Caywood-Rukas, D. (2010). *The predictive validity of early reading curriculum-based measurement of high-stakes tests in grades 3, 4, and 5* (Doctoral Dissertation). Retrieved from ProQuest. UMI number 3367100.
- 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*, 5-10.
- Deno, S.L. (2003). Curriculum-based measures: Development and perspectives. *Assessment for Effective Intervention, 28*(3), 3-12.
- Ditkowsky, n.d. *Cut Scores to Illinois Standards Achievement Test based on 2010 scores*. Retrieved July 10, 2011 from http://measuredeffects.com/measuredeffects.com/UserFiles/modules/file_upload_library/Full%202010%20Cut%20Scores.pdf
- Ditkowsky, B. (2009). *2008 cut scores for R-CBM to ISAT*. Retrieved May 20, 2012 from http://measuredeffects.com/measuredeffects.com/UserFiles/modules/file_upload_library/Recc%20CBM%20CUTS%2008a.pdf
- Ditkowsky, B. & Koonce, D.A. (2010). Predicting performance on high-stakes assessment for proficient students and students at risk with oral reading fluency growth. *Assessment for Effective Intervention, 35*, (159-167).

- Dominguez de Ramirez, R. & Shapiro, E.S. (2006). Curriculum-based measurement and the evaluation of reading skills of Spanish-speaking English language learners in bilingual education classrooms. *School Psychology Review*, 35(3), 356-369.
- Farmer, E., Swanlund, L., & Pluymert, K. (2010). *Curriculum based measurement and language proficiency in English language learners*. Paper presented at the annual meeting of the National Association of School Psychologists, Chicago, IL.
- Federal Interagency Forum on Child and Family Statistics (2011). *America's Children: Key National Indicators of Well-Being, 2011*. Washington, DC: U.S. Government Printing Office. Retrieved July 17, 2011 from http://childstats.gov/pdf/ac2011/ac_11.pdf
- Fien, H., Baker, S.K., Smolkowski, K., Mercier Smith, J.L., Kame'enui, E.J. & Beck, C.T. (2008). Using nonsense word fluency to predict reading proficiency in kindergarten through second grade for English learners and native English speakers. *School Psychology Review*, 37(3), 391-408.
- Fuchs, L.S. & Fuchs, D. (1999). Monitoring student progress toward the development of reading competence: A review of three forms of classroom based assessment. *School Psychology Review*, 28, 659-671.
- Fuchs, L.S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review*, 33, 188-192.
- Fuchs, L.S. & Deno, S. (1994). Must instructionally useful performance assessment be based in curriculum? *Exceptional Children*, 61, 15-24.
- Fuchs, L.S., & Fuchs, D. (2006). A framework for building capacity for Responsiveness to Intervention. *School Psychology Review*, 35, 621-626.
- Fuchs, L.S., Fuchs, D., (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review* 22(1), 27-48.
- Fuchs, L.S., Fuchs, D., Hosp, M., & Jenkins, J. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, 5, 239-256.
- Geva, E. (2006). Second-language oral proficiency and second language literacy. In D. August and T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth* (pp. 123-140). Mahwah, N.J.: Lawrence Erlbaum Associates.

- 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(3), 257-288.
- Hamilton, C. & Shinn, M. E. (2003). Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skills. *School Psychology Review*, 32, 228-240.
- Hemphill, F.C. & Vanneman, A. (2011). *Achievement Gaps: How Hispanic and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress* (NCES 2011-459). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Retrieved July 17, 2011 from <http://nces.ed.gov/nationsreportcard/pdf/studies/2011459.pdf>
- Hintze, J.M., Callahan, J.E., Matthews, W.J., Williams, S.A.S. & Tobin, K.G. (2002). Oral reading fluency and prediction of comprehension in African American and Caucasian elementary school children. *School Psychology Review*, 31, 540-553.
- Hintze, J.M. & Silbergitt, 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.
- Hosp, J.L., Hosp, M.A., & Dole, J.K. (2011). Potential bias in predictive validity of universal screening measures across disaggregation subgroups. *School Psychology Review*, 40(1), 108-131.
- Howe, K.B. & Shinn, M.M. (2003). *Standard reading assessment passages (RAPs) for use in general outcome measurement: A manual describing development and technical features* [Technical Manual]. Eden Prairie, MN: Edformation.
- Illinois State Board of Education (n.d.). *English Language Learning Annual Assessment-ACCESS for ELLs. Frequently asked questions*. Retrieved December 28, 2011 from http://www.isbe.state.il.us/bilingual/htmls/access_faq.htm#1
- Illinois State Board of Education (2010a). *Interpretive Guide. Illinois Standards Achievement Test 2010*. Retrieved January 4, 2012 from http://www.isbe.net/assessment/pdfs/ISAT_Interpr_Guide_2010.pdf
- Illinois State Board of Education (2010b). *Illinois Standards Achievement Test 2010 Technical Manual*. Retrieved January 4, 2012 from http://www.isbe.net/assessment/pdfs/isat_tech_2010.pdf

- Individuals with Disabilities Education Act (2004). 20 U.S.C. § 1400 *et seq.*
- Jenkins, J.R. & Jewell, M. (1993). Examining the validity of two measures for formative teaching: Reading aloud and maze. *Exceptional Children*, 59, 421-432.
- KewalRamani, A., Gilbertson, L., Fox, M. A., & Provasnik, S. (2007). *Status and trends in the education of racial and ethnic minorities* (NCES 2007-039). Retrieved April 25, 2012 from <http://nces.ed.gov/pubs2007/minoritytrends/index.asp>
- Kieffer, M.J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology*, 100(4), 851-868.
- Kranzler, J.H., Miller, M.D. & Jordan, L. (1999). An examination of racial/ethnic and gender bias on curriculum-based measurement of reading. *School Psychology Quarterly*, 14, 327-342.
- Lesaux, N. (2006). Development of literacy. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (pp. 75–122). Mahwah, NJ: Erlbaum.
- Linklater, D.L. (2007). *Kindergarten literacy assessment of English only and English Language Learner students: Which measure(s) is most predictive of reading skills?* (Doctoral Dissertation). Retrieved from ProQuest. UMI number: 3298235
- Mancilla-Martinez, J. & Lesaux, N.K. (2011). The gap between Spanish speakers' word reading and word knowledge: A longitudinal study. *Child Development*, 82, 1544-1560.
- Marston, D. (1989). Curriculum-based measurement: What is it and why do it? In M. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp.18-78). New York: Guilford Press.
- Marston, D., Mirkin, P., & Deno, S. (2001). Curriculum-based measurement: An alternative to traditional screening, referral, and identification. *The Journal of Special Education*, 18(2), 109-117.
- McCardle, P. (2006). Forward. In D. August and T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth* (pp. 1-42). Mahwah, N.J.: Lawrence Erlbaum Associates.

- McGlinchey, M.T. & Hixson, M.D. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review*, 33(2), 193-203.
- Meisinger, E.B., Bradley, B.A., Schwanflugel, P.J., Kuhn, M.R., & Morris, R.D. (2009). Myth and reality of the word caller: The relation between teacher nominations and prevalence among elementary school children. *School Psychology Quarterly*, 24(3), 147-159.
- Muyskens, P, Betts, J., Lau, M.Y., & Marston, D. (2009). Predictive validity of curriculum-based measures in the reading assessment of students who are English language learners. *The California School Psychologist*, 14, 11-21.
- National Center for Culturally Responsive Educational Systems (2005). *Cultural considerations and challenges in Response-to-Intervention models. An NCCREST position statement*. Retrieved July 15, 2011 from http://www.nccrest.org/PDFs/rti.pdf?v_document_name=Culturally%20Responsive%20RTI
- National Clearinghouse for English Language Acquisition (n.d.). *The growing numbers of limited English proficient students*. Retrieved June 15, 2010 from http://www.ncela.gwu.edu/files/uploads/4/GrowingLEP_0506.pdf
- National Institute for Literacy (2007). *What content area teachers should know about adolescent literacy*. Retrieved August 5, 2011 from http://lincs.ed.gov/publications/pdf/adolescent_literacy07.pdf
- National Institute of Child Health and Human Development (2000). *Report of the national reading panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for instruction*. Retrieved September 2, 2010 from http://www.nichd.nih.gov/publications/nrp/upload/smallbook_pdf.pdf
- No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002).
- Northwest Evaluation Association (2011a). *MAP Basics Overview*. Retrieved January 5, 2012 from <http://www.nwea.org/sites/www.nwea.org/files/resources/MAPBasicsOverview.pdf>
- Northwest Evaluation Association (2011b). *Illinois Linking Study. A Study of the Alignment of the NWEA RIT Scale with the Illinois Standards Achievement Test*. Retrieved January 4, 2012 from

http://www.nwea.org/sites/www.nwea.org/files/resources/IL_Linking%20Study_Final.pdf

- Pearce, L.R. & Gayle, R. (2009). Oral reading fluency as a predictor of reading comprehension with American Indian and white elementary students. *School Psychology Review*, 38(3), 419-427.
- Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models, applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Reschly, A.L., Busch, T.W., Betts, J., Deno, S.L., Long, J.D. (2009). Curriculum based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology*, 47, 427-469.
- Shinn, M.M. & Shinn, M.R. (2002a). *AIMSweb training workbook. Administration and scoring of reading curriculum-based measurement (R-CBM) for use in general outcome measurement*. NCS Pearson. Retrieved May 12, 2012 from <https://aimsweb.pearson.com/downloads/AdminandScoringR-CBM09292005.pdf>
- Shinn, M.M. & Shinn, M.R. (2002a). *AIMSweb training workbook. Administration and scoring of reading Maze for use in general outcome measurement*. NCS Pearson. Retrieved May 12, 2012 from <http://www.aimsweb.com/wp-content/uploads/Maze-Manual.pdf>
- Shinn, M. R. (2008). Best practices in using curriculum-based measurement in a problem-solving model. In A. Thomas and J. Grimes (eds.) *Best practices in school psychology V* (pp. 671-697). Bethesda, MD: National Association of School Psychologists.
- Sibley, D., Biwer, D., & Hesch, A. (2001, April). *Establishing curriculum-based measurement oral reading fluency performance standards to predict success on local and state tests of reading achievement*. Paper presented at the annual meeting of the National Association of School Psychologists, Washington, DC.
- Snow, C.E., Burns, M.S., & Griffin, P. (Eds.) (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Stage, S.A. & Jacobsen, M.D. (2001). Predicting student success on a state mandated performance-based assessment using oral reading fluency. *School Psychology Review*, 30, 407-419.
- Stokes, N. O. (2010). *Examining the relationship among reading curriculum-based measures, level of language proficiency, and state accountability test scores with*

middle school Spanish-speaking English language learners (Doctoral dissertation). Retrieved from ProQuest. UMI Number: 3433188.

University of Oregon, Center on Teaching and Learning, (n.d.). *Big ideas in beginning reading*. Retrieved August, 5, 2011 from http://reading.uoregon.edu/big_ideas/

U. S. Department of Education (2004). *No Child Left Behind: A toolkit for teachers*. Retrieved August 12, 2011 from <http://www2.ed.gov/teachers/nclbguide/nclb-teachers-toolkit.pdf>

U.S. Department of Education. Institute of Education Sciences (2011). Retrieved June 26, 2011 from <http://nces.ed.gov/programs/coe/index.asp>

U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), *2007 Reading Assessment*, retrieved April 23, 2009, from the Main NAEP Data Explorer (<http://nces.ed.gov/nationsreportcard/naepdata/>).

Vanderwood, M.L., Linklater, D., & Healy, K. (2008). Predictive accuracy of nonsense word fluency for English language learners. *School Psychology Review*, 37(1), 5-17.

Vanderwood, M. L. & Nam, J. E. (2007). Response to Intervention for English Language Learners: Current development and future directions. In *Handbook of Response to Intervention*, Jimerson, S.R., Burns, M.K., & VanDerHeyden, A.M. eds (pp 408-417). New York, NY: Springer Science + Business Media.

WIDA Consortium (2011). *ACCESS for ELLs interpretive guide for score reports*. Board of Regents of the University of Wisconsin System. Retrieved December 28, 2011 from <http://www.wida.us/assessment/ACCESS/#scores>

Wiley, H.I. & Deno, S.L. (2005). Oral reading and maze measures as predictors of success for English learners on a state standards assessment. *Remedial and Special Education*, 26(4), 207-214.

Yanosky, T., Yen, S.J., Longuit, M., MacGregor, D., Zhang, Y., & Kenyon, D.M. (2011). *Annual technical report for ACCESS for ELLs English Language Proficiency Test, Series 201, 2009-2010 Administration*. WIDA Consortium. Retrieved January 4, 2012 from <http://www.wida.us/DownloadDocs/assessment/ACCESS/TechReports/1%20Tech%20Rpt%206%20Series%20201.pdf>

Zehr, M.A. (2009). Graduation rates on ELLs a mystery. *Education Week*, 29(3), 1, 20-21.

VITA

Ella was born and raised in Oregon. She received her undergraduate degree in anthropology with general honors from the University of Chicago. After graduation, she began working as an intern in an elementary Montessori classroom and subsequently completed a teacher training program through the Montessori Institute of Milwaukee. She also received a Master's degree in Education with a focus on Montessori education from Loyola College in Maryland and completed additional requirements to obtain teacher certification at the elementary level in Illinois. After several years of teaching, Ella entered the graduate program in school psychology at Loyola University Chicago. After completing her course requirements and internship, she received state certification and worked for three years as a school psychologist in a suburb of Chicago. She recently returned to teaching and is currently employed as a Montessori elementary teacher in a private school in Chicago.