



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

Nicole Osterman Stokes
Loyola University Chicago

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



Part of the [Educational Assessment, Evaluation, and Research Commons](#)

Recommended Citation

Stokes, Nicole Osterman, "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" (2010). *Dissertations*. 137.

https://ecommons.luc.edu/luc_diss/137

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 © 2010 Nicole Osterman Stokes

LOYOLA UNIVERSITY CHICAGO

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

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

PROGRAM IN SCHOOL PSYCHOLOGY

BY

NICOLE OSTERMAN STOKES

CHICAGO, ILLINOIS

DECEMBER 2010

Copyright by Nicole Osterman Stokes, 2010
All rights reserved.

ACKNOWLEDGEMENTS

This has been a long journey, but I am so pleased that I kept my focus on the final goal. I could not have made it through this journey without the continual guidance and support of Dr. David Prasse, my dissertation chair. I would like to express my deepest appreciation to him for helping me stay focused and not give up on my goal. His knowledge and expertise that he has so willingly shared with me throughout my doctoral pursuit has been vital to my success. I respect him greatly not only as my dissertation chair, but as a researcher and professor.

I would also like to thank the other members of my committee, Dr. Theresa Pigott and Dr. Diane Morrison. I am extremely grateful to Dr. Pigott for her guidance and support with a highly technical and difficult statistical analysis. I would also like to express my appreciation to Dr. Morrison for her support and encouragement throughout the dissertation process. I am also very grateful to Dr. Alex Yu from Arizona State University. Although he was not an official member of my dissertation committee, without his vast knowledge and expertise in statistical analysis, this study would not have been possible.

Many thanks also to the administrators of Cartwright Elementary School District in Phoenix, AZ, who took time out of their busy schedules to help facilitate my data collection and answer my questions. I would also like to thank Brian Owin of the

Litchfield Elementary School District who provided much needed assistance in getting my large data set organized.

Finally, I am deeply grateful to my family and friends for their tremendous amounts of encouragement and support during the dissertation process. I would especially like to thank my husband, Jeff, and daughter, Kendall for believing in me and allowing me to pursue my educational goals.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iii
LIST OF TABLES	vii
LIST OF FIGURES	ix
ABSTRACT	xi
CHAPTER I: INTRODUCTION	1
Older ELL Students with Reading Difficulties	3
CBM in Reading with English Language Learners	5
CBM and High-Stakes Assessment	7
Statement of the Problem	8
Relevance of the Study	10
CHAPTER II: REVIEW OF THE LITERATURE	13
Reading Difficulties in Older Students	14
Reading and English Language Learners	17
Curriculum-Based Measurement in Reading	22
Curriculum-Based Measurement with Older Students	24
CBM and English Language Learners	25
Curriculum-Based Measures and High-Stakes Assessments	34
Summary and Concluding Remarks	42
CHAPTER III: METHODOLOGY	44
Introduction	44
Subjects and Setting	44
Procedures	45
Description of Measures	46
Data Analysis	52
Summary	56
CHAPTER IV: RESULTS	58
Descriptive Statistics	61
CBM Models	65
AIMS Models	69
Prediction Analyses	88
CHAPTER V: DISCUSSION	104
R-CBM Initial Status and Growth: Results	105
R-CBM Initial Status and Growth Relative to ELL Status: Results	105
AIMS Performance and R-CBM Initial Status and Growth Rates: Results	106
Predictors for AIMS Performance: Results	109

Results in Relation to Previous Research	112
Implications for Instruction	120
Limitations for the Present Study	124
Suggestions for Future Research	126
Conclusion	129
REFERENCES	133
VITA	141

LIST OF TABLES

Table 1: Student Population Demographics	45
Table 2: Subjects and Data	46
Table 3: AZELLA Performance Level Descriptions of Language Proficiency for Sixth Grade	48
Table 4: Arizona Reading Concepts and Strands	50
Table 5: AIMS Performance Level Descriptions	51
Table 6: Types of Data and Assessments	52
Table 7: Frequency and Percent of ELL Students in AZELLA Levels 1-5	59
Table 8: Means and Standard Deviations of Fall, Winter, and Spring R-CBM Scores	61
Table 9: Means and Standard Deviations of Fall, Winter, and Spring R-CBM Scores for ELL and Non-ELL Students	62
Table 10: Means and Standard Deviations for Oral Reading Fluency Scores at Fall, Winter and Spring by AIMS Level	64
Table 11: R-CBM Unconditional Growth Model (from Fall to Spring)	66
Table 12: Final Estimation of Variance Components for the CBM Unconditional Growth Model	67
Table 13: Differences in Initial Status and Growth Rates Between ELL and Non-ELL Students	68
Table 14: Final Estimation of Variance Components for ELL Status	69
Table 15: ANOVA for AIMS, ELL Status, and Time on R-CBM Initial Status	70
Table 16: Differences in Initial Status and Growth Rates by AIMS Reading Level	71
Table 17: Final Estimation of Variance for AIMS Level	72

Table 18: ELL Status and AIMS Reading Level Three-way Interaction Effect on R-CBM Initial Status and Growth Rate	73
Table 19: Final Estimation of Variance for AIMS Level and ELL Status	74
Table 20: Correlation Matrix of Predictor and Outcome Variables	88
Table 21: ANOVA: Model and AIMS Reading Scores	92
Table 22: Parameter Estimates of Predictor Variables on AIMS Outcome	93
Table 23: ANOVA: Non-ELL Model and AIMS Reading Scores	96
Table 24: Parameter Estimates of Predictor Variables on AIMS Outcome for Non-ELL Students	96
Table 25: ANOVA: ELL Model and AIMS Reading Scores	96
Table 26: Parameter Estimates of Predictor Variables on AIMS Outcome for ELL Students	97

LIST OF FIGURES

Figure 1: R-CBM data for all students that Fell Far Below the standard on the AIMS assessment. The outlier was removed from the data analysis.	60
Figure 2: Parallel coordinate of individual student's growth from Fall to Spring in R-CBM data by AIMS level 1 (Falls far below).	75
Figure 3: Parallel coordinate of individual student's growth from Fall to Spring in R-CBM data by AIMS level 2 (Approaches the standard).	76
Figure 4: Parallel coordinate of individual student's growth from Fall to Spring in R-CBM data by AIMS level 3 (Meets the standard).	77
Figure 5: Parallel coordinate of individual student's growth from Fall to Spring in R-CBM data by AIMS level 4 (Exceeds the standard).	78
Figure 6: Parallel coordinate of non-ELL students' growth on R-CBM from Fall to Spring.	79
Figure 7: Parallel coordinate of ELL students' growth on R-CBM from Fall to Spring.	80
Figure 8: Parallel coordinate of ELL students' growth on R-CBM from Fall to Spring at AIMS level 1 (Falls far below).	82
Figure 9: Parallel coordinate of non-ELL students' growth on R-CBM from Fall to Spring at AIMS level 1 (Falls far below).	82
Figure 10: Parallel coordinate of ELL students' growth on R-CBM from Fall to Spring at AIMS level 2 (Approaches the standard).	84
Figure 11: Parallel coordinate of Non-ELL students' growth on R-CBM from Fall to Spring at AIMS level 2 (Approaches the standard).	85
Figure 12: Parallel coordinate of ELL students' growth on R-CBM from Fall to Spring at AIMS level 3 (Meets the standard).	86
Figure 13: Parallel coordinate non-ELL students' growth on R-CBM from Fall to Spring at AIMS level 3 (Meets the standard).	87

Figure 14: Homoscedasticity and independence of residuals.	90
Figure 15: Normality of residuals.	90
Figure 16: Scatterplot of the normality of residuals.	91
Figure 17: Scatterplot of Fall R-CBM by AIMS Scaled Score.	93
Figure 18: Scatterplot of ELL status by AIMS Scaled Score (ELL = 1, non-ELL = 0)	94
Figure 19: Scatterplot of Percentage of Growth on R-CBM from Fall to Spring and AIMS Scaled Score.	94
Figure 20: Scatterplot of Maze scores and AIMS Scaled Score.	95
Figure 21: Scatterplot of R-CBM by AIMS Scaled Score for non-ELL students.	98
Figure 22: Scatterplot of R-CBM percentage of growth by AIMS Scaled Score for non-ELL students.	99
Figure 23: Scatterplot of Maze by AIMS Scaled Score for non-ELL students.	99
Figure 24: Scatterplot of Fall R-CBM b AIMS Scaled Score for ELL students.	100
Figure 25: Scatterplot of R-CBM growth by AIMS Scaled Score for ELL students.	100
Figure 26: Scatterplot of Maze by AIMS Scaled Score for ELL students.	101
Figure 27: AIMS Scaled Score versus growth by ELL status.	102
Figure 28: AIMS Scaled Score versus growth by Maze.	102
Figure 29: AIMS Scaled Score versus initial status (Fall R-CBM) by ELL status.	103
Figure 30: AIMS Scaled Score versus initial status (Fall R-CBM) by Maze.	103

ABSTRACT

The purpose of the present study was to examine the predictive ability of oral reading fluency (R-CBM) on a sixth grade high-stakes assessment with ELL and non-ELL students, as well as determine the average rate of growth on R-CBM and how that relates to level of English Proficiency. The participants in the current study included 350 sixth grade students from a middle school located in west Phoenix, AZ. Ninety of the 350 students were English Language Learners at varying levels of language proficiency in English. Archival data was used for the study. Each participant completed R-CBM three times throughout the 2006-2007 school year (fall, winter, and spring benchmarks), the Maze assessment (administered once in early spring), and the state accountability exam (AIMS) given in the spring. Of the 350 total students in the sample, 90 also were administered a Language proficiency exam (Arizona English Language Learner Assessment), due to their ELL status. In an attempt to avoid statistical and interpretive problems with the data analysis, HLM, multiple regression, and data visualization was used. This study was conducted for the purpose of examining the typical rate of growth on R-CBM over one school year for ELL and non-ELL students. It also assisted in determining whether reading screening measures (R-CBM, Maze) are effective in predicting future reading success or failure.

Results indicated that the average initial status as measured by the Fall R-CBM was 103.51 words read correctly per minute with an average growth rate (slope) of 10.36

words read correct from Fall to Winter, and from Winter to Spring. Thus, the overall growth was about 20.72 words over one school year. The results indicated that there was significant variation between ELL and non-ELL students in their initial status as measured by the Fall R-CBM score. However, while ELL students score significantly lower in the Fall on R-CBM, they did not vary in their growth trajectories from non-ELL students throughout the school year. Students varied significantly based on their AIMS level in their initial status as measured by the fall R-CBM score, but did not vary significantly on their growth rates relative to their performance level on AIMS. The regression analysis for the overall student population showed that Fall RCBM (initial status), percentage of growth on RCB from Fall to Spring, and ELL status were all significant predictors of student performance on AIMS. The Maze assessment was not a significant predictor. A student's initial status, as measured by the Fall R-CBM measure, was the most important factor in predicting whether that student will meet the state standards on the AIMS assessment in the spring for both the ELL and non-ELL populations. Implications for instruction and suggestions for future research are discussed.

CHAPTER I

INTRODUCTION

The importance of the acquisition of early literacy skills is well known. Students who have difficulty learning to read often continue to have significant difficulty throughout their educational careers. Often referred to as the “Matthew Effect” described by Stanovich (1986), students who acquire early reading skills are equipped with the tools to exponentially grow in their knowledge and skills while students who fail to develop early literacy skills continue to fall further and further behind. Evident in recent national concern and legislation, many children are failing to develop early literacy skills which lead to poor academic and social success (Haager & Windmueller, 2001). With the introduction of the No Child Left Behind Act (NCLB, 2001), states are required to test students on an annual basis and show that at least 95% of students are meeting annual measurable objectives in reading and math. In addition to the general student population, individual subgroups, including English Language Learners and students with disabilities, must also meet the 95% goal. Thus, it is essential for schools to identify students at the start of the school year who are at risk of not meeting state standards, as well as to improve overall reading outcomes for these students through appropriate intervention. Universal screening using Curriculum Based Measurement is one way of quickly assessing and identifying at-risk students early on in the school year. Curriculum Based Measurement (CBM) oral reading has been shown as a reliable assessment tool for

predicting student performance on state tests (Hintze & Silbergliitt, 2005; McGlinchey & Hixson, 2004; Silbergliitt et al., 2006; Stage & Jacobsen, 2001; Wiley & Deno, 2005). However, very few studies have addressed the use of CBM oral reading with English Language Learners, and more specifically whether CBM is a reliable predictor of ELL performance on high stakes assessments. The CBM Maze task is another general outcome measure of reading that has been explored in the literature (Ardoin et al., 2004; Brown-Chidsey, Johnson, & Fernstrom, 2005; Shinn, Deno, & Espin, 2000; Twyman & Tindal, 2007; Wiley & Deno, 2005). However, research examining the predictive validity of the Maze task on state high-stakes assessments in reading is minimal, particularly with English Language Learners.

English language learners (ELLs) frequently have difficulties developing early reading skills. For example, Hispanic students are almost twice as likely as non-Hispanic Whites to be reading below the expected level for their age (Snow, Burns & Griffin, 1998 as cited in Gunn et al., 2000). What's more, many Hispanic students are getting as far as middle school with reading skills that still fall significantly below their non-Hispanic White peers. While there is a growing body of research on ELL students and reading in the elementary grades, there has been minimal research done with middle school students (i.e., grades 6-8). According to Denton et al. (2008), the research on effective intervention for older readers with reading difficulties is lacking, particularly with students who are English Language Learners (ELLs). The importance of early intervention is well established. However, many districts are facing high numbers of transition ELL students moving in and out of their schools, which poses a significant

problem for effectively meeting these student's needs. Often, students are starting middle school (sixth grade) with reading skills that are falling significantly below grade level expectations. More research is needed on how to effectively identify, intervene, and monitor these students in order to improve reading outcomes for this population. The current study attempts to further research in the area of assessment and prediction with older ELL students by examining the use of Curriculum Based Measurement (CBM) in reading, both oral reading and Maze, and how these general outcome measures relate to ELL student performance on high stakes tests.

Older ELL Students with Reading Difficulties

According to Kamil (2003), more than 25% of middle school students are unable to read well enough to sufficiently identify the main idea of reading passages. There have been three possible explanations for the low reading skills of older students identified by Denton et al. (2008). These include: 1) older students with reading deficits do not have adequate knowledge of the alphabetic principle and word reading; 2) older students with reading difficulties do not possess an adequate understanding of word meaning and adequate reasoning skills necessary for comprehension of text; and 3) older students with reading difficulties do not have the task orientation toward reading to gain adequate reading proficiency. For older students who are also English Language learners, the challenge of becoming an adequate reader becomes even more complex. Not only do these students struggle with word reading skills and vocabulary knowledge, but many of them also present with limited background knowledge as well as various

contextual factors (e.g., high mobility rates, low socioeconomic status, etc.) that may affect their future reading success.

Although there has been research conducted on elementary ELL students struggling with reading, studies examining older ELL students with reading difficulties are limited. One study by Denton et al. (2008) examined the effectiveness of a multicomponent reading intervention with middle-school Spanish-speaking ELL students. Results showed that the treatment group did not demonstrate higher outcomes than the non-treatment group in terms of word recognition, comprehension, or fluency. The authors concluded that ELL students with severe reading difficulties may require considerably more intensive interventions over considerably longer amounts of time in order to improve reading outcomes. The results from this study may contribute to the notion that middle school students with reading difficulties are too old for remediation. However, there have been several studies conducted with adolescents with reading difficulties which indicate that older students are generally responsive to reading interventions (Edmonds et al., 2009; Ehren, Lenz, & Deshler, 2004; Scammacca et al., 2007). Although these studies are promising, there is more research needed in the use of reading assessment procedures in order to quickly and accurately identify middle-school students in need of intervention, particularly with ELL students.

CBM in Reading with English Language Learners

Assessment is a key component of any effective reading program. According to Haager and Windmueller (2001), assessment serves three critical purposes in developing a reading intervention program, which include the identification of students in need of supplemental instruction, guiding instructional planning, and ongoing monitoring of student progress. As with any intervention, students need to be screened in some way to identify a need for assistance. For reading intervention, there are numerous screening tools that can be used for identifying students at risk for reading failure. Curriculum-based measurement (CBM) is one such screening tool. CBM has been established as a systematic, standardized, and reliable assessment tool for examining a student's progress in reading. Deno (2003) described multiple uses of CBM including norm development, identification of students academically at risk, and prediction of performance on important criteria. CBM also has the potential to be a method for assessing both the level and growth of student performance in reading. While the research on the typical rate of growth in reading for monolingual English speaking students has grown over the past decade, studies investigating the typical growth rate of ELL students are limited (Dominguez de Ramirez & Shapiro, 2006). Moreover, there is even less research examining growth patterns in reading of middle school students, specifically students who are English language learners. By developing growth standards in reading specifically for Spanish-speaking ELLs, including those students in the higher-elementary grades and middle school, practitioners could be more efficiently and effectively identifying students who are having difficulty (Dominguez de Ramirez &

Shapiro, 2006). Once identified, these students could be provided with intervention and their progress monitored frequently. Their growth trajectory could be compared with that of the typical Spanish-speaking ELL student, which would also assist school teams in answering that difficult question of whether their lack of academic progress is due to second language acquisition or an underlying disability.

The underlying assumption that fluency and comprehension are related is often challenged by ELL teachers who claim that these students can decode words without comprehending what they read (Dominguez de Ramirez & Shapiro, 2006). As such, additional research is needed that looks more carefully at the link between ORF (oral reading fluency) and comprehension for ELLs. One measure that is currently being used to assess reading comprehension at the higher grade levels is the Maze assessment. The Maze assessment consists of one standardized reading passage of grade-level difficulty (Howe & Shinn, 2002). In each Maze passage, the first sentence is left intact, with each subsequent seventh word being replaced with three words to choose from (e.g., When **(red/she/told)** was a little girl...). The Maze task has been found to be a reliable and valid measure of reading comprehension for students in elementary, middle, and high school (Brown-Chisdey, Davis, & Maya, 2003; Shin, Deno, & Espin, 2000). However, very few studies have examined the reliability and validity of the Maze task with the ELL population. One study by Wiley and Deno (2005) found that while the Maze task was a better predictor than oral reading fluency for fifth grade non ELL students, it was less predictive than oral reading for the ELL population. Thus, more research is needed on

whether the Maze task would be an appropriate assessment tool for use with the ELL population in predicting reading success.

CBM and High-Stakes Assessment

Another benefit of using CBM data is to assist in predicting whether students will meet state standards as assessed by the state tests mandated by the No Child Left Behind Act of 2001. There have been a few studies examining the relationship between curriculum-based measurement for reading (R-CBM) and state accountability test scores (Hintze & Silbergitt, 2005; McGlinchey & Hixson, 2004; Silbergitt et al., 2006; Stage & Jacobsen, 2001). One study found that while there was a relationship between CBM and state test scores, the magnitude of the relationship declined significantly as grade level increased (Silbergitt et al., 2006), yet this decline was established with a primarily non-ELL population. As found in previous research, the growth trajectory on reading fluency measures for ELL students differs from that of mono-lingual English speaking students. In one particular study (Dominguez de Ramirez & Shapiro, 2006), while Spanish-speaking ELLs rate of growth in English was significantly slower than their non ELL peers, they demonstrated substantial improvement by fifth grade. As such, the predictability of reading CBM on state test performance may differ for the ELL population. In other words, if ELL student's oral reading fluency continues to increase at a high rate into the later elementary-early middle school years, the magnitude of the relationship between CBM and state test performance may remain consistent or even increase as the grade level increases. In the current studies, the sample populations participating in the studies were primarily White, not of Hispanic Origin. While one of

the current studies “non-Caucasian” populations made up 52% of the total sample, the language proficiency of these students was not revealed (McGlinchey & Hixson, 2004). As such, it is difficult to generalize the current findings of the relationship between CBM and performance on state tests to the ELL student population. Another study by Wiley and Deno (2005) examined oral reading and the Maze measures to predict performance on the state test with both non-ELL and ELL students in third and fifth grade. Results showed moderate to moderately strong correlations between the state standards test and the two CBM measures for all students. However, when the two measures were combined, the predictive power only increased for the non-ELL students. Further research is needed that examines the potential of oral reading and maze for use with the ELL population in terms of assessment of reading proficiency.

Statement of the Problem

Further research is needed to examine the relative validity of CBM reading assessment with older ELL students for the purpose of developing an assessment system that is able to quickly and accurately target students in need of intervention. In addition, this assessment system could be utilized to change student outcomes in terms of future high-stakes assessments. The current research examined the average rate of growth over one school year in CBM oral reading fluency, the relationship of this rate of growth to students’ level of English Language Proficiency, and whether reading fluency and MAZE measures are predictive of later high-stakes success in reading for older ELL students.

Dominguez de Ramirez and Shapiro (2006), Wiley and Deno (2005), Twyman and Tindal (2007), Sibley et al. (2001), Hintze and Silbergliitt (2005), McGlinchey and

Hixson (2004) have indicated the need for more research in the areas being addressed by the current research. The present study adds to previous findings of the predictiveness of R-CBM assessments to high-stakes tests with the use of both oral reading fluency and Maze measures. Unlike previous studies, however, this research examines the predictive power of R-CBM measures on future high-stakes success with middle-school students who are also English Language Learners. Level of language proficiency will also be examined in terms of its relationship with the rate of growth on reading fluency measures, as well as whether it has an effect on the magnitude of R-CBM and Maze as they compare to high-stakes test scores. Due to the limited research on R-CBM with the ELL population, an expected rate of growth in terms of reading fluency for this population is lacking. In a study by Dominguez de Ramirez and Shapiro (2006), although progress was slower than their non-ELL peers, Spanish-speaking ELL students showed growth on CBM measures over time, demonstrating that CBM can be sensitive to examining the course of literacy development of this population. As such, there is a need for more research to assist in developing an expected rate of growth in order to determine norms and/or benchmarks for this population. In other words, what is the expected rate of growth from the fall to winter benchmark assessment for students who are learning English? In addition to rate of growth, it is also crucial to be able to identify early (i.e., on the fall benchmark assessment) which students are at risk of reading failure in order to provide intervention quickly. As such, there needs to be more research on establishing appropriate “cut-off” points for accurately identifying ELL students in need of more intensive intervention. The current study utilized R-CBM in fall, winter, and spring to

examine the typical rate of growth for this population as well as examining the oral reading cut-off scores in the fall in order to predict success or failure on state tests on the spring.

Relevance of the Study

The number of ELL students enrolled in public schools in the U.S. has increased significantly over the past few years. According to the U.S. Census Bureau, it is estimated that the foreign-born population in the United States increased by 57% from 1990 to 2000 (from 19.8 million to 31.1 million, respectively), and that over half of the foreign-born population were from Latin America (U.S. Census Bureau, 2000). According to a report by Macias (2000), it was estimated that 78% of the limited English proficient students in the U.S. in 1997-1998 were Spanish-speaking. With this growing number of Spanish-speaking ELL students in the schools, educators have become more concerned with how to appropriately meet the educational needs of this population, particularly in reading. The educational difficulty of these students has become a national concern due to low reading achievement that increases in severity in later grades, which in turn has led to high rates of grade retention and dropout for these students (Rueda & Windmueller, 2006).

Currently, there is a political and educational emphasis on accountability, high-stakes assessment, and student outcomes. Districts are being held accountable for the outcomes of all students based on the results of standards-based, high stake assessments. Schools not only have to show that 95% of their students are proficient on state tests, but individual subgroups, including English Language Learners, must also meet the 95% goal

(Wiley & Deno, 2005). This puts schools in a difficult position, because by the time this assessment is given (usually in the spring), it is too late to identify students at risk as well as to determine whether reading instruction was effective in improving student outcomes. For this reason, many districts have been using CBM as a way of identifying early in the school year students who are at risk of reading failure, as well as measuring student progress over time. In addition, recent research has been promising in the use of oral reading fluency measures as a predictor of whether students will meet standards on state tests. While CBM has been well established as an effective assessment tool for English speaking students, there has been limited research on the use of CBM with ELL students. According to McCardle, Mele-McCarthy and Leos (2005), there is a need for accurate and user-friendly assessment tools that schools can utilize for screening and progress monitoring with the ELL population.

The purpose of the current study is to examine the following questions:

- 1) What was the average rate of growth of the current population over one school year on the reading fluency R-CBM measures?
- 2) What is the relationship among the rate of growth on R-CBM and level of English Language Proficiency?
- 3) If students vary on their growth rates (slopes) on R-CBM and/or intercept (performance on state AIMS reading assessment), is this variation related to ELL status?
- 4) Is R-CBM initial status, R-CBM rate of growth, and Maze predictive of reading achievement on the spring AIMS reading assessment relative to ELL

status? Which of these factors is a more potent predictor of reading achievement on the AIMS reading assessment?

CHAPTER II

REVIEW OF THE LITERATURE

With the introduction of the No Child Left Behind Act (2001), schools are being required to test all of their students on an annual basis in order to establish that the majority of students are making adequate yearly progress in both reading and math. In other words, at least 95% of the student population, which includes individual subgroups (e.g., English Language Learners) must score proficient on some sort of high-stakes assessment administered in the spring. Because of this 95% goal, schools, teachers, and support personnel are feeling tremendous amounts of pressure to identify early on those students who are at-risk of not meeting state standards, particularly in reading. As such, there is a well-established need for simple yet reliable assessments that have the ability to predict academic outcomes. One such assessment is Curriculum-Based Measurement (CBM) (Deno, 1985), which is a brief set of measures that serve as indicators, or general outcome measures, of one specific academic area such as reading (Stecker, Lembke, & Foegen, 2008). CBM is widely used and well-researched at the elementary level in order to universally screen for students who are at risk of reading failure, as well as progress monitor those students who have been targeted as in need of additional instruction/intervention. However, there is limited research on the use of CBM with older students, and whether it is an effective assessment tool for predicting outcomes and monitoring progress over time. Yet another challenge is the use of CBM with English Language

learners (ELLs), particularly middle and secondary students. While the research on the use of CBM with the ELL population is emerging for the elementary level, there is very limited research on the whether CBM is a technically adequate measure of student performance for older ELL students. Due to CBMs strong empirical history as reliable and valid assessment procedures (Fuchs, 2004), it would seem that CBM has the potential to be a valid and reliable assessment tool for older students who also happen to be English Language Learners. However, more research is needed in this area that will further examine the use of CBM with older students and English Language Learners in terms of its effectiveness with screening/identification, predicting future reading success, and monitoring student progress.

Reading Difficulties in Older Students

It has been documented that reading failure has a negative impact on academic achievement, peer relations, extracurricular activity, vocabulary development, and acquisition of knowledge (Stanovich, 1986). Moreover, many adolescents who are homeless and who have committed suicide also were experiencing reading disabilities (Barwick & Siegel, 1996; McBride & Siegel, 1997). It is well-documented that early identification of reading difficulties is crucial for changing the reading trajectories of students at risk for reading failure. However, many middle school and secondary school students continue to struggle with reading, yet not much emphasis is placed on identification and intervention with older, struggling readers. By the time students reach the middle school level, they are expected to have acquired the basic reading skills in order to comprehend high level reading material.

One study by Leach, Scarborough, and Rescorla (2003) examined late-identified reading disabilities with fourth and fifth grade students. Results of this study revealed that 35% of the students had word-level processing deficits with adequate comprehension skills, 32% of the students had poor comprehension skills with strong word-level skills, and 32% of the students had difficulty with both word-level processing and comprehension. The authors noted that the reading difficulties of students with deficits in both areas (i.e., comprehension and word-level) not only stemmed from poor word recognition, but also from several other factors, including limited vocabulary, lack of background knowledge, and poor inferential abilities. Another study examining older students with reading difficulties (i.e., fourth grade students) also found that for 36% of their sample, below-average word reading skills did not emerge until the students were in fourth grade (Lipka, Lesaux, & Siegel, 2006). The authors found that these students showed a sharp decline in age-appropriate word reading skills in grade 4. It was concluded that these students were reading primarily by sight in the earlier grades and only began to have difficulty as the word reading required for grade 4 increased in complexity.

Older students may struggle with reading for numerous reasons, including a lack of effective early intervention, or if they have received effective intervention early, they begin to struggle as text and knowledge demands increase (Leach, Scarborough, & Rescorla, 2003; Lipka, Lesaux, & Siegel, 2006). Two other possible explanations for poor reading skills in older students is that they have difficulty with word-level reading and/or they struggle with word meaning and have poor reasoning abilities (Denton et al.,

2008). Often by the time student get to middle-school, they are expected to have the ability to decode words efficiently and read fluently. The expectation of many middle-school teachers is that by the time students reach middle-school, these students should be reading to learn, not learning to read. However, many older students who are struggling with reading may still need instruction at the word-reading level, as they may have not yet mastered the basic reading skills (i.e., decoding and fluency) that are required to benefit from instruction focused on reading for meaning (Leach et al., 2003). In a study by Fletcher (2007) that examined the word reading, fluency, and comprehension skills of struggling middle-school readers, results showed that over half of the students demonstrated significant difficulties with word reading. Older students may also lack adequate vocabulary and may not possess adequate reasoning abilities, which are both essential in comprehending text.

Results of intervention studies with older students who are struggling with basic reading deficits indicate that these students benefit from instruction focused on word reading, regardless of their age (NICHD, 2000). Two recent meta-analyses examined the effectiveness of reading interventions for adolescent readers (Edmonds et al., 2009; Scammacca et al., 2007). In the Edmonds et al. study, intervention research on decoding, fluency, vocabulary, and comprehension and its influence on reading comprehension outcomes for students in grade 6 through 12 was reviewed. Of the 29 total intervention studies identified, 13 were used in the meta-analysis to examine the overall effect on reading comprehension. Results revealed a mean weighted effect size of .89. Specifically, results indicated that when explicitly taught reading comprehension

strategies, older struggling readers can improve their comprehension. Word-level interventions were associated with small to moderate effects on reading comprehension. In terms of interventions focused specifically on reading fluency, it was found that an increase in oral reading fluency did not lead to an improvement in reading comprehension.

In the Scammacca et al. (2007) synthesis, they analyzed 31 intervention studies, many of which were included in the Edmonds et al. (2009) meta-analysis. However, they included more recent studies and expanded the grade levels to 4 through 12. Similar to Edmonds et al., they also found moderate to large effects when comparing treatment versus comparison groups in terms reading comprehension outcomes with an overall effect size of .95. Results showed the largest effect sizes for interventions focusing on reading comprehension strategies and vocabulary, moderate effect sizes with word study, and the lowest effect sizes with interventions focused only on fluency. While these findings suggest that interventions with older, struggling readers can be effective in improving overall reading comprehension, very few of the studies included in the meta-analyses focused on ELL students.

Reading and English Language Learners

According to Klinger, Artiles, and Barletta (2006), a question commonly raised regarding ELL students and reading is whether ELLs struggle to develop literacy because of their limited proficiency in English or because of other factors, such as learning disabilities. It has been documented in research that second language acquisition is a complicated process, which is impacted by numerous factors, including language

proficiency in the first language, socio-cultural factors and personality (August & Hakuta, 1997). According to McLaughlin (1992), acquiring a second language is not automatic and often requires a lengthy amount of time for the development of social and academic language skills. ELLs are not only expected to learn academic content, but they are also required to navigate and comprehend a new system of language and communication (Scribner, 2002). In fact, while basic interpersonal conversation skills (BICS) often develop without formal instruction, cognitive academic language proficiency (CALP) takes between five and seven years to develop and requires formal schooling (Cummins, 1984 as cited in Vanderwood & Nam, 2009). For example, one study found that it took ELLs longer to attain non-ELL norms in reading and language arts than it took them to attain non-ELL norms in math (Collier, 1987). Acculturation has also been discussed as having an impact on academic achievement on ELLs. In one study by Lopez, Ehly, and Garcia-Vazquez (2002), acculturation pattern and amount of social support received by Mexican-American students were examined in terms of these student's academic achievement. Results showed that acculturation patterns that integrated aspects of both the old and new cultures were correlated with higher levels of academic achievement. Research has also shown that time in the U.S. is a significant predictor of reading outcomes for ELLs (Betts et al., 2009).

As a result of the various factors discussed above, English language learners frequently have difficulties developing early reading skills. For example, Hispanic students are almost twice as likely as non-Hispanic Whites to be reading below the expected level for their age (Snow, Burns & Griffin, 1998 as cited in Gunn et al., 2000).

One longitudinal study by Kieffer (2008) examined reading growth rates for students who entered kindergarten either proficient or not proficient in English and their reading trajectories through fifth grade. Results showed that the students who entered kindergarten with limited English proficiency had reading growth trajectories that grew at a substantially slower pace than their non-ELL peers throughout their elementary school years. Moreover, the group of students in this study who technically acquired oral English proficiency by the end of first grade (as indicated by passing Pre-K language proficiency assessment) continued to have slow rates of growth in English reading across the elementary school years (Kieffer, 2008).

Due to these gaps in the acquisition of early reading skills, the implementation of effective strategies to teach these early skills while also supporting English-language learning has become a principal concern. However, addressing this concern presents a very complicated issue. There is little agreement in the literature on how to effectively teach ELL students how to read (Gersten & Baker, 2000 are cited in Haager & Windmueller, 2001). However, the latest research suggests that both English language learners and native speakers of English develop prereading skills (e.g., phonological awareness) in similar ways (Gersten & Geva, 2003). Moreover, new research indicates that with appropriate instruction, English language learners are able to learn phonological awareness and decoding skills in English as quickly as their native English speaking peers (Geva, 2000 as cited in Gerstan & Geva, 2000; Thompson, Vaughn, Hickman-Davis, & Kouzekanani, 2003). However, the research on the effectiveness of reading instruction with English language learners remains limited. Therefore, there is a

significant need for research to further examine effective intervention strategies for English language learners at-risk for reading difficulties.

In examining the literature in the area of reading instruction with English language learners with reading difficulties, several themes emerge. The first theme deals with what constitutes an effective reading program, particularly for English language learners. The subtopics that fall under this broad theme include the type of reading program being used (i.e., whether it is a specific program, like Reading Mastery, or just supplemental instruction incorporating the critical elements of reading instruction), the amount of time the students receive the intervention/supplemental instruction, the primary language of instruction, and fidelity of the treatment. A second emerging theme is how the students are evaluated/assessed before, during, and after receiving the supplemental reading instruction. Thirdly, experience and training, including ongoing support (i.e., coaching) of the teachers/staff provide the supplemental instruction surfaces as a theme throughout the literature. The majority of the studies on the effectiveness of supplemental reading instruction for English language learners seem to focus mainly on quantitative data in assessing student outcome, although few studies do incorporate qualitative data regarding teacher feedback and teacher strategies.

The research on effective reading instruction and intervention with elementary-aged ELLs is growing and providing more evidence that the reading outcomes for these students can be improved. While the research is promising on the effectiveness of improving the reading outcomes for older, struggling non-ELL readers through targeted intervention, very little research has examined the effectiveness with older ELL students.

These students are presented with unique challenges when it comes to reading, as they often struggle with both word reading as well as word meaning (Denton et al., 2008). In one of the few studies focusing on older ELL students and reading, Denton et al. examined the effectiveness of a reading intervention with Spanish-speaking ELL middle school students with severe reading difficulties. The sample consisted of 64 students in grades 6 through 8, who were screened using R-CBM. If the students read less than 80 correct words per minute on the fifth grade passage, they were included in the intervention study. A total of 41 students qualified for the study, and were randomly assigned to the treatment or typical instruction condition. The students in the treatment group received daily intervention for 13 weeks total, which resulted in a total of 29 hours of supplemental instruction. The intervention consisted of small-group phonics-based remedial instruction that included ELL strategies, vocabulary instruction, fluency, and comprehension strategies. Results revealed no differences among the treatment and comparison group on any of the pre and post reading measures (i.e., Woodcock Johnson III Basic Reading, Passage Comprehension, and Spelling subtests; Test of Word Reading Efficiency Sight Words; and Oral Reading Fluency). What's more, neither the treatment nor comparison group demonstrated significant standard score growth over the course of the study. The authors identify several hypotheses as to why these students did not respond to the intervention, which included "within-student" factors such as self-efficacy beliefs, language development, and low oral vocabulary, as well as instructional factors, such as the intensity and duration of the intervention.

Curriculum-Based Measurement in Reading

Curriculum-based measurement (CBM) has been described by Deno (1985) as a set of simple, standardized, short-duration fluency measures that can be used to monitor student growth in important areas of literacy. CBM is considered a dynamic indicator of basic skills, in that the measures are designed to be sensitive to change in student performance as a result of instruction (Shinn, 2009). According to Stecker, Fuchs, and Fuchs (2005), CBM possesses three distinctive characteristics. The first unique characteristic is that CBM assesses student progress toward long-term goals, and is considered an evaluation of general outcomes. Second, CBM uses frequent monitoring (usually once or twice per week) as a predictor of whether students are making adequate rates of progress toward long-term goals. The third feature of CBM that is critical to its use is its well-documented technical adequacy. Marston (1989) summarized results of several psychometric studies on the development of CBM, which validated its use for progress monitoring and instructional decision-making. Further research on the reliability of CBM has examined median level of standard error of measurement (SEM) and has begun to establish median levels of SEM across grades (Christ & Silberglitt, 2007; Poncy, Skinner, & Axtell, 2005).

The use of CBM with elementary students has been extensively documented in the literature. This research has suggested that for the majority of elementary-aged students, the total number of words read correctly on a one-minute reading probe is a valid indicator of general reading proficiency (Shinn, 2009). Common uses of CBM as outlined by Deno (2003) include improving individual instructional programs, enhancing

teacher instructional planning, predicting performance on important criteria, developing norms, increasing the ease and effectiveness of communication, screening to identify students academically at risk, evaluating students' response to instruction/intervention, reducing bias in assessment, and providing alternative procedures for special education identification. Research has recently begun to also examine the use of CBM in measuring reading growth in middle and secondary school environments, in assessing and monitoring English Language Learners, and in predicting student performance on state accountability measures (Deno, 2003).

It is well known that by assessing students in early elementary (kindergarten and first grade) in phonemic awareness, those students who are in need of supplemental instruction can be easily identified. As students move into the upper elementary grades, measures of oral reading fluency can be used to predict future reading success, as well as monitor students' ongoing progress in reading. In middle and secondary schools, however, technically adequate measurement systems are lacking (Twyman & Tindal, 2007). According to Deno (2003), research in the use of CBM methods to assess middle and secondary students' reading growth are promising, yet tentative. While it has been established that CBM measures do correlate with important criteria for older students, such as test scores and grade point average, these correlations have not been as strong as they are for elementary school students (Deno, 2003).

Curriculum-Based Measurement with Older Students

Currently, very little is known about effective intervention with older ELL students in terms of improving reading outcomes. It is clear that more research in this area is needed in order to improve the reading trajectories for older students who also happen to be English Language Learners. Before these students can receive intervention, however, schools need quick and efficient ways of identifying those students in need of such intervention. While screening for early reading difficulties has been well-established in the research with elementary-aged students, very little is known about valid and reliable screening tools for use with older students (i.e., middle and secondary schools). Currently, middle schools have limited technically adequate measurement systems to assist in both the identification and ongoing progress monitoring of students with reading difficulties. Curriculum-based measurement has the potential to become that technically adequate tool for evaluating the reading skills of older students, although the research is still emerging in this area. Most of the research with CBM in middle school has focused on its use in evaluating content-area reading comprehension skills. The two assessment tools that have been used for this purpose are vocabulary matching and reading comprehension mazes (Espin & Tindal, 1998). In one study by Twymann and Tindal (2007), the reliability of a concept maze task was investigated for use with middle-school students. The study examined the performance of 240 sixth grade Caucasian students on a traditional maze task, a concept maze task with examples, and a concept maze task with attributes. Results revealed that the traditional maze task showed the least alternate form reliability, but the most consistency with readability. The

attribute maze task was found to be more difficult and more stable when compared to the other two types of maze tasks. However, due to moderate correlations and low internal consistency, the authors concluded that the attribute maze was not very reliable.

Results from the above study suggest that all three mazes are different from each other, and that for purposes of assessing reading comprehension skill, the traditional maze is an effective tool. However, authors indicate that the traditional maze is not sufficient for measuring conceptual knowledge as measured by the attribute maze. The question then becomes for what purpose is the CBM maze task being used. For identification and screening purposes, it would seem that the traditional maze task would prove a valid tool in targeting those students who have poor reading skill and who are at risk for poor reading outcomes. For those targeted students, the use of the traditional maze task may also then be used to monitor student progress in terms of reading comprehension skills. However, if a middle-school teacher is more interested in a tool for monitoring the growth of students in terms of the specific content area, then it would seem the attribute maze would be more appropriate. Whatever the purpose, it is evident that further research is needed in examining the use of CBM maze tasks with middle and secondary students.

CBM and English Language Learners

Research has shown that many English Language Learners are not developing the skills necessary to become good readers, and as a result, have become highly overrepresented in special education (Case & Taylor, 2005). What's more, the achievement gap between ELLs and their non-ELL peers continues to grow based on

national and state test results (Snow & Biancarosa, 2003 as cited in Vanderwood & Nam, 2009). Based on the overrepresentation of ELLs in special education, as well as the rising achievement gap between the ELL and non ELL population, there is an obvious need for reliable and valid assessment tools for the purpose of early identification of ELL students who may be struggling in reading. However, such assessment tools are lacking for this population (Wagner, Francis, & Morris, 2005). According to Good, Simmons, and Kame'enui (2001), there is a need for developing cut-off scores for the ELL population in order to differentiate those students who are at-risk versus those who are not at risk of reading difficulties. Moreover, in order to adequately assess ELL student's response to instruction, further research needs to be conducted which examines appropriate progress monitoring tools and the expected growth rates on these tools with ELL students (Vanderwood & Nam, 2009).

According to Scribner (2002), there are several factors to keep in mind when assessing English Language Learners. First, it is typical for English skills such as vocabulary, grammar, and comprehension to be less developed than their non-ELL peers. Second, ELL students follow a predictable sequence of learning a second language (i.e., listening, speaking, reading, and writing) that is similar to young children learning a first language. Third, it is common for ELL students to switch back and forth between the two languages. Scribner (2002) also outlined five necessary components in the assessment of English Language Learners, which includes pre-referral documentation (e.g., dynamic assessment of skills), language proficiency assessment, assessment of literacy skills (e.g., oral and written skills in both native language and in English), formal

assessment (e.g., use of reliable and valid psychometric measures as well as non-psychometric measures), and instructional recommendations (e.g., use various modalities in teaching, pre-teach concepts/new vocabulary, etc.).

Due to its well-documented use with English-speaking students, curriculum-based measurement has the potential to become the primary assessment tool used to identify and monitor the progress of the ELL population, and research on the use of curriculum-based measures in reading with English Language Learners is emerging. One study by Baker and Good (1995) examined the use of curriculum-based measurement in reading (oral reading fluency) with bilingual Hispanic students. The study involved 76 second grade students, and these students were divided into two groups for the purpose of the study: bilingual and English-only. The measures used for the study were CBM English oral reading fluency, the Stanford Diagnostic Reading Test (SDRT), and the Teacher Rating of English Reading (TRR), which was specifically developed for the study. In addition to the reading measures, three language measures were utilized: English Language Fluency (ELF), Spanish Language Fluency (SLF), and the Language Assessment Scales (LAS). The ELF and SLF were developed by the researchers for the study, and the LAS is a published, norm-referenced test of English and Spanish language proficiency. Results indicated that R-CBM was a reliable and valid measure of English reading for both groups, with no significant differences between ELL and non-ELL students. Results also showed that R-CBM correlated higher with criterion measures of reading than with criterion measures of language. The authors also conclude that R-CBM may be a sensitive tool in measuring ELL students' reading progress over time.

However, the reliability of the slope of student progress over time was low for both ELL and non-ELL students. The authors indicate the need for further research examining the rate of progress, or slope, of R-CBM in terms of student outcomes.

Other studies have examined R-CBM as a measure of change in student progress over time with Spanish-speaking ELLs. In a study examining the effectiveness of supplemental reading instruction over a two-year period for Hispanic and non-Hispanic students, oral reading fluency was shown to be a sensitive measure of reading growth and students' response to intervention (Gunn et al., 2005). In fact, benefits of the intervention were still evident for two years following, as students in the intervention group continued to show significantly greater growth in oral reading fluency as compared to the control group. The results of this study provide support for the use of R-CBM as a sensitive assessment tool to monitor the English reading growth rates of ELL students over time. Another study by De Ramírez and Shapiro (2006) examined the oral reading fluency growth rates among 83 English-speaking students and 62 Spanish-speaking English Language Learners in grades 1-5. Students were assessed three times during the year using R-CBM passages. Results indicated that R-CBM can be used as a sensitive measure in assessing the rate of growth in reading for Spanish-speaking ELL students, although the rates of reading growth on English R-CBM measures were significantly slower for ELL students than for non-ELLs. However, the authors point out that rate of improvement, or slope, across grade levels substantially improved by fifth grade. They conclude that these results are reflective of the 5-7 year time frame that has been well-established in terms of how long it takes typical ELL students to reach proficiency in

English (Cummins, 1991 as cited in De Ramírez & Shapiro, 2006). A more recent study by Betts et al. (2009) found that the average R-CBM growth rate for third-grade ELLs met the third-grade expectations of one word correct per minute per week. The authors concluded that the oral reading fluency growth of ELLs in their study resembled that of non-ELL students. However, it was noted that the ELLs actual oral reading fluency rate was substantially lower than what would be typically expected for third grade students.

The research on R-CBM has indicated that the rate of growth in oral reading fluency beings to level off as students approach the upper elementary grades. However, the majority of this research has been conducted with English speaking students. The current study provides a starting point for examining the use of R-CBM in later grades with English Language Learners, particularly in examining their growth patterns over time.

Due to its growing use with the ELL population, research has also begun to examine potential ethnic, language, and socioeconomic bias in oral reading fluency (R-CBM) scores. According to Kranzler, Miller, and Jordan (1999), the validity of R-CBM with students from diverse racial/ethnic and linguistic backgrounds has not been thoroughly examined in the research. Since CBM is now being used for high-stakes decisions (e.g., selection for intervention programs and special education eligibility), it is essential to gain an understanding of potential bias and group differences (Klein & Jimerson, 2005). In one study by Kranzler, Miller, and Jordan (1999), the ethnic and gender bias with African American and Caucasian students was examined in terms of oral reading fluency probes. Results showed that at grades 4 and 5, African American's

students reading comprehension was over-predicted by CBM while Caucasian students' reading comprehension was under-predicted. However, in a follow-up study that utilized the same methodology as the Kranzler et al. study, results did not detect any evidence of bias in the use of R-CBM in predicting reading comprehension scores for African-American and Caucasian students (Kranzler & Taylor, 2000 as cited in Klein & Jimerson, 2005). Yet another similar study conducted by Hintze, Callahan, Matthews, Williams, and Tobin (2002) found that oral reading fluency probes were a stronger indicator of reading comprehension for African American students than for Caucasian students. While these studies did examine bias in terms of ethnicity, they did not examine bias in terms of linguistic differences.

One study by Klein and Jimerson (2005) did examine potential home-language and ethnicity bias in oral reading fluency scores predicting reading outcomes with Caucasian and Hispanic students. The outcome measure utilized in this study was the Stanford Achievement Test – Ninth Edition (SAT-9). When combining ethnicity and home language factors, results revealed that R-CBM over-predicted the reading proficiency (as measured by the SAT-9) of Hispanic students whose home language was Spanish while it under-predicted the reading proficiency of Caucasian students whose home language was English. However, bias was not detected for language or ethnicity alone when the other factor was held constant. The authors concluded that it is a combination of these two factors, as opposed to ethnicity or home language alone that significantly contributes to bias. Nonetheless, a close examination of the analyses indicated that the home language factor appeared to be the strongest factor influencing

the bias. The authors hypothesize that the ecological factors associated with home language, including CALP (cognitive/academic language proficiency) may have a different effect on the skills that are required by R-CBM and SAT-9. In other words, the SAT-9 is more of a direct measure of higher level comprehension skills, while oral reading fluency is only related to these skills. The overall correlations between R-CBM and SAT-9 scores did show oral reading fluency probes to be a good measure of reading proficiency for both Hispanic and Caucasian students. The results of this study as well as the other studies that detected potential linguistic and ethnic bias with R-CBM points to the need for further research examining the use of R-CBM with various populations, particularly students whose home language is not English.

Much of the research with CBM as a predictor of future reading performance has focused on oral reading fluency. While ORF has been well-established as a valid and reliable general outcome measure of reading for English speaking students, the research on the use of ORF with the ELL population is limited. The underlying assumption that fluency and comprehension are related is often challenged by ELL teachers who claim that these students can decode words without comprehending what they read (De Ramírez & Shapiro, 2006). As such, additional research is needed that looks more carefully at the link between ORF (oral reading fluency) and comprehension for ELLs. One measure that is currently being used to assess reading comprehension at the higher grade levels is the Maze assessment. The Maze assessment consists of one standardized reading passage of grade-level difficulty (Howe & Shinn, 2002). In each Maze passage, the first sentence is left intact, with each subsequent seventh word being replaced with

three words to choose from (e.g., When (**red/she/told**) was a little girl...). The Maze task has been found to be a reliable and valid measure of reading comprehension for students in elementary, middle, and high school (Brown-Chisdey, Davis, & Maya, 2003; Shin, Deno, & Espin, 2000). In a study by Shin, Deno, and Espin, 43 English speaking second graders' reading performance was monitored on a monthly basis using the CBM Maze assessment over one school year. Results from the study revealed the Maze to be a reliable assessment of student growth in reading over time, and that student performance on the Maze task sensitively reflects individual and group improvements in reading. Results also showed that the rate of growth on the Maze task was predictive of student performance on a state accountability assessment. Brown-Chisdey, Davis, and Maya (2003) also examined the use of curriculum-based measurement silent reading (CBM-SR or Maze) with fifth through eighth grade students. Results indicated that CBM-SR was able to accurately identify students whose reading comprehension skills fell below their peers. The results of this study provide initial evidence for the use of the Maze task as a screening tool for poor reading comprehension in older students. However, in a study by Ardoin et al. (2004), it was found that administering a Maze in addition to a R-CBM probe did not provide more information about the reading comprehension of second graders than did R-CBM alone. The differences among these studies in terms of the utility of the Maze task may be due in part to the varying grade levels of the participants. In other words, the Maze task may be a more reliable and valid assessment tool when used with older students versus younger students. Nonetheless, further research is needed

that examines the usefulness of the Maze task in predicting future reading performance as well as monitoring student reading growth over time.

Research has begun to examine the reliability and validity of the Maze task with the ELL population. Teachers often question the validity of oral reading fluency for ELL students, as many believe that ORF only measures the ability of ELL students to decode the words as opposed to measuring their reading comprehension skills (Wiley & Deno, 2005). As such, the Maze task may provide a more accurate measure of ELL students' overall reading proficiency versus oral reading fluency alone. One study by Wiley and Deno found that while the Maze task was a better predictor than oral reading fluency for fifth grade non ELL students, it was less predictive than oral reading for the ELL population. Furthermore, the Maze scores did not significantly contribute to the performance on a high-stakes reading assessment for ELL students, while it did significantly contribute for non-ELLs. In another study by Vanderwood, Linklater, and Healy (2008), the Maze was used as an outcome measure along with R-CBM and a norm-referenced, standardized achievement test with ELL students. Researchers were examining the relationship of another type of curriculum-based measure (Nonsense Word Fluency) with the various outcome measures along with level of language proficiency. Results showed Maze to have strong positive correlations with Nonsense Word Fluency scores in fall, winter, and spring (.50, .51, and .54, respectively), with R-CBM (.66), and with ELL Level (.56). Maze was moderately correlated with the standardized achievement test in reading (.27). It is evident from the varying results of these two

studies that further research needs to investigate whether the Maze assessment could be utilized as a valid and reliable measure of reading in English Language Learners.

Curriculum-Based Measures and High-Stakes Assessments

Early work on looking at the use of oral reading fluency and performance on standardized achievement tests began with a study by Deno, Mirkin and Chiang (1982). In this study, five CBM measures were used, and oral reading fluency emerged as the strongest measure, with correlations ranging from .71 to .91 with a published criterion measures. Another validity study by Fuchs, Fuchs, and Maxwell (1988) examined curriculum-based measures in reading with special education students in middle school. Results revealed that reading fluency, or number of correct words read within one minute, consistently had the highest correlation with standardized achievement tests among other simple measures of reading (e.g., answering questions, cloze tests, and recall). One of the most important findings from this study was that oral reading fluency had a higher correlation with the measure of reading comprehension (.92) than it did with the decoding criterion measure (.81). This supports the notion that oral reading fluency or reading curriculum-based measurement (R-CBM) can be considered a general outcome measure of reading achievement. Much of the criterion validity research on R-CBM was summarized by Marston (1989). He examined 14 separate studies which included several different criterion measures, but the majority were published norm-referenced tests. Correlations clustered around .80, with coefficients ranging from .63 to .90. In more recent research, studies have begun to examine the usefulness of curriculum-based measures in predicting future reading success in terms of state

accountability outcomes. By screening students in the fall using R-CBM, research has shown that student performance on the early screening measures can predict whether or not the students will meet state standards on the high-stakes assessments in the spring.

One study conducted in 2001 by Crawford, Tindal, and Steiber examined the use of R-CBM in predicting success on Oregon statewide achievement tests in both reading and math. The study examined a total of 51 students over two years of schooling (second and third grade), with 94% of the sample being White. In general, students made a 42 word gain in their oral reading rate from second to third grade. There was a moderate correlation (.60) found between R-CBM scores and the third grade high stakes reading assessment. A strong correlation was found between students' R-CBM score in second grade and their score in third grade. A nonparametric analysis conducted between reading rate (R-CBM) and state reading test performance revealed a significant relationship between the two variables. Furthermore, the study found that if students were reading at least 119 correct words within one minute, they were nearly guaranteed to pass the high stakes reading assessment. The authors identify the most critical finding of the study as a result of the cross-years analysis, which revealed that 100% of the students reading at least 72 correct words per minute in second grade passed the state accountability reading test in third grade.

Another study conducted in 2001 by Stage and Jacobsen examined the relationship between student's oral reading fluency and a high stakes reading assessment (Washington Assessment of Student Learning or WASL). They looked at 173 fourth grade students, with the majority of the sample (90%) being European Americans. They

were assessed three times during the school year (Fall, Winter and Spring) using CBM reading fluency benchmarks, and also participated in the state accountability reading test in the Spring. Unlike the other study in this area, the current study conducted two growth curve analyses using hierarchical linear modeling (HLM). Results indicated that the Fall R-CBM cut-scores increased the predictive power of whether students would pass or fail the WASL by 30% over base rate levels. The study also found that R-CBM scores in the Fall were a better predictor of the state test performance than growth in R-CBM throughout the school year. Consistent with previous research (Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993), this study also found that the oral reading fluency slope decreased to less than one word per week by fourth grade. The authors conclude that although oral reading fluency increases in later grades, it does not increase as much as in earlier grades, which helps to explain why the increase in slope did not account for as much variance as the level of performance at each assessment period (Fall, Winter, and Spring). The authors caution that further research needs to be conducted examining the slope in oral reading fluency and high-stakes reading assessments with more representative samples before ruling out the use of slope as a significant predictor of performance.

While the findings of Stage and Jacobsen (2001) provided a promising beginning in substantiating the use of R-CBM as an effective tool for predicting performance on high stakes assessments, the limited sample size did not allow for generalizability of the findings. Due to this limited generalizability, another study by McGlinchey and Hixson (2004) replicated the Stage and Jacobsen (2001) study, but examined a different state

fourth grade reading test across eight years, as well as using a larger, more diverse sample of students. The total student sample involved 1,362 fourth grade students, with about half of the student sample being Caucasian, half African American, and very small percentage of Hispanic and American Indian. The students were administered a R-CBM probe in the two weeks prior to the administration of the state assessment across a total of eight years. This study examined the diagnostic efficiency of the reading rate cut-score. For this study, 100 words read correct per minute were used as the cut-score based on previous research (Hasbrouck & Tindal, 1992; Stage & Jacobsen, 2001). Results indicated that 72% of students who read at least 100 word correct within one minute passed the state assessment. These results are similar to the previous study by Stage and Jacobsen (2001), in that R-CBM increased the prediction of performance on a fourth grade state accountability assessment; however, in the current study, the correlation between the state test scores and R-CBM was higher. The authors conclude this higher correlation may have been due to the differing state assessments and/or the differences between the populations used in each study.

Another study by Hintze and Silbergitt (2005) examined varying statistical and methodological approaches to using R-CBM as a predictor of success on state accountability tests. The authors conducted a longitudinal study that examined 1,766 students who had completed R-CBM and state tests over a period from first through third grades. Their overall sample consisted of mainly White not of Hispanic Origin (94%). The students were assessed using R-CBM eight times beginning in the Fall in Grade 1 and continuing through Spring of Grade 3, along with the state accountability test in the

Spring of Grade 3. Results of the study revealed that the predictive validity of R-CBM to the high stakes test was significant across all time periods that were assessed. In addition, the study examined the sensitivity, specificity, positive predictive power and negative predictive power of R-CBM on high stakes test outcomes using three statistical procedures (i.e., discriminant analysis, logistic regression, and ROC curves). Results revealed that each procedure set R-CBM cut scores with adequate levels of diagnostic accuracy and efficiency. The authors concluded that R-CBM is both an accurate and efficient way of predicting students who are likely to pass the reading portion on state accountability tests as far back as first grade. In further support of these findings, Wood (2006) also found that oral reading fluency cut-off scores for students in grades 3, 4 and 5 had adequate levels of diagnostic efficiency in predicting whether students would pass or fail the Colorado Student Assessment Program (CSAP).

Silberglitt, Burns, Madyun, and Lail (2006) examined the relationship between curriculum-based measurement in reading and high stakes assessments, and whether there were differences among grade levels in the magnitude of the relationship. Their data sample included 5,472 students in grades 3, 5, 7 and 8, with the majority of the sample being White, not of Hispanic Origin (94.3%). All students in the sample were administered R-CBM and a Maze assessment, along with a criterion-referenced, high-stakes standardized achievement test. Results revealed a significant relationship between R-CBM and state accountability test scores, with correlation coefficients ranging from .71 for third graders to .51 for eight graders. The study also found that the magnitude of the relationship between R-CBM and high-stakes test results declined significantly as

grade level increased. For this reason, the authors cautioned against using R-CBM as a predictor of test performance on state accountability for later grades. They recommended that future research examine potential reasons for the decline of predictive power of R-CBM in later grades. However, in a study by Wood (2006), results showed that oral reading fluency predicted performance on a state accountability assessment equally well for third, fourth, and fifth grade students. While this sample only included students through fifth grade, this research does call into question the hypothesis proposed by Fuchs et al. (2001) that the relationship between oral reading fluency and reading comprehension changes as students reach higher grades. Nonetheless, further research is needed examining the predictiveness of oral reading fluency on the performance of older students regarding high-stakes assessments.

In addition to R-CBM being used as a predictor of future reading success at a single point in time, few studies have begun to examine oral reading fluency as a direct measure of reading over time. While the study by Stage and Jacobsen (2001) found that oral reading fluency slope was not a significant predictor of performance on a state accountability test, there were several problems cited in the literature with their data analysis (Baker et al., 2008). According to Baker et al., the Stage and Jacobsen study computed fall, winter, and spring ORF estimates from the slope, and as a result the four variables used in the regression most likely led to multicollinearity. This, in turn, inflated standards and generated tests of statistical significance that were highly problematic (Cohen, Cohen, West, & Aiken, 2003 as cited in Baker et al., 2008). In one study by Fuchs, Fuchs, and Compton (2004), the CBM slope of two first grade reading measures

(word identification fluency and nonsense word fluency) were examined in terms of predictive validity on criterion-referenced measures of reading. Results revealed large and reliable differences in the magnitude of the correlations for full-year slopes, with coefficients ranging from .27 to .58 for nonsense word fluency slopes, and .50 to .85 for word identification fluency slopes. However, this study estimated the slope without considering the student's initial performance level (intercept). One more recent study by Baker et al. (2008) investigated whether the ORF slope of performance over time predicted performance on high stakes reading tests over and above the initial level of performance. This study is also the first to examine to what degree a change in oral reading fluency over time (i.e., slope) is related to better performance on state accountability tests. The study used a large sample size, with four cohorts of students in grades 1-3, with each cohort representing about 2,400 students. It is also important to note that 32% of their sample was English language learners, with 68% being Latino students. Results showed that in grades 1-3, oral reading fluency was associated with performance on the high stakes tests with correlations ranging from .58 to .82, which is consistent with previously cited research in this area. The most noteworthy finding from this study was that oral reading fluency slope went above information provided by levels of performance alone, and added to the accuracy of predicting student performance on state accountability tests.

As is evident from the reviewed research, R-CBM has been shown to be a significant predictor of student performance on state accountability tests. However, very few studies have looked at the use of R-CBM with English Language learners,

particularly whether they are useful in predicting future reading success on high stakes assessments. One study by Wiley and Deno (2005) examined both oral reading fluency and Maze measures with ELL and non-ELL students in terms of their ability to predict student performance on the *Minnesota Comprehensive Assessment* (MCA). The sample consisted of 36 third-grade students (21 non-ELL and 15 ELL), and 33 fifth-grade students (19 non-ELL and 14 ELL). The primary languages of the third grade ELL students was Hmong (80%), Somali (13%), and Spanish (7%). The primary language of the fifth grade students was Hmong (100%). The students were administered measures of oral reading fluency, maze and the state assessment (MCA). Results showed moderately strong correlations between ORF and the MCA for non ELLs (.71) and moderate for ELLs (.61). For fifth grade students, the correlations were moderate for both ELL (.57) and non ELL students (.69). Another study by Vanderwood, Linklater, and Healy (2008) examined the relationship between nonsense word fluency, also a curriculum-based measure in reading used mainly with first grade students, and performance of Spanish-speaking English Language Learners in reading. The sample consisted of 134 ELL students in a low-income school in southern California. All first grade students were administered Nonsense Word Fluency (NWF) three times per year, and two years later, were assessed with measures of oral reading fluency, Maze, and the California Achievement Test, Sixth Edition (CAT 6). Results revealed that NWF had a strong relationship with the third-grade measures which included Oral Reading Fluency (.65), Maze (.54), and had a moderate relationship with CAT 6 (.39). Another important finding was that NWF scores accurately identified close to 80% of the students above the

25th percentile on all three third-grade outcome measures (ORF, Maze, and CAT 6).

While these two studies found moderate to strong correlations between curriculum-based measures and ELL students performance on high-stakes assessments in reading, further research needs to be conducted that examines R-CBM and whether it is an accurate predictor for reading performance in the English Language Learner population.

Summary and Concluding Remarks

The current review of the literature was intended to provide a summary of the most relevant research related to the topic of reading assessment, particularly how it applies to older ELL students. Due to the increasing number of ELL students within the educational system, research has begun to examine how we are assessing this population of students for the purposes of identification and monitoring of academic progress, with a specific focus on the area of reading. A brief summary of the research on reading difficulties of both older students and ELL students was provided. This research points to the need for valid and reliable assessment tools to appropriately identify older ELL students in need of intervention. Curriculum-based measurement was described as one possible assessment tool that could be used for this purpose. However, as can be seen in the review of the research on CBM, most of the CBM studies have focused on younger, English speaking students. Research on the use of CBM with younger ELL students is emerging, and these studies have shown promising results in terms of its effectiveness in appropriate identification of ELL students at risk of reading failure, as well as its ability to predict future reading outcomes for this population. Very few studies have examined

the use of CBM with older students, particularly older students who also are English Language Learners.

This review of the literature has identified a gap in the research in terms of dynamic models of assessment and English-Language Learners, particularly older ELL students. The current study attempts to further research in this area by using curriculum-based measures in reading to predict rate of growth over time on later high stakes assessments in relation to level of language proficiency with a population of middle-school English Language Learners.

CHAPTER III
METHODOLOGY

Introduction

The purpose of the present study was to examine the predictive ability of oral reading fluency (R-CBM) on a sixth grade high-stakes assessment with ELL and non-ELL students, as well as determine the average rate of growth on R-CBM and how that relates to level of English Proficiency. The high-stakes test is a state-mandated achievement test based on Arizona state learning standards. Although the high-stakes assessment measures several academic skills, only the reading scores are used in this study.

Subjects and Setting

The participants in the current study included 350 sixth grade students from a middle school located in west Phoenix, AZ. Ninety of the 350 students were English Language Learners at varying levels of language proficiency in English. Some of the demographic data was obtained through the Arizona Department of Education School Report Card for the 2006-2007 school year, which is displayed in Table 1. Certain data, such as mobility rate and socioeconomic status, were not made available on the School Report Card.

Table 1 *Student Population Demographics*

Meets or exceeds expectations on AIMS in Reading 2007	49%
Attendance rate	94%
White	7.4%
Black	3.9%
Hispanic	86.3%
Asian/Pacific Islander	1.4%
Native American	.8%
Limited English Proficient	26%

Procedures

Archival data was used for the current study. All assessment procedures involved in this study would have occurred even if this study were not conducted. Confidentiality was maintained by coding all students.

Each participant completed R-CBM three times throughout the 2006-2007 school year (fall, winter, and spring benchmarks), the Maze assessment (administered once in early spring), and the state accountability exam (AIMS) given in the spring. Of the 350 total students in the sample, 90 also were administered a Language proficiency exam (Arizona English Language Learner Assessment), due to their ELL status (see Table 2).

Table 2 *Subjects and Data*

Assessments/Tests	Number of Subjects
Reading Curriculum-Based Measurement (Fall, Winter and Spring)	350
Curriculum-Based Measurement - Maze	350
Arizona Instrument to Measure Standards (AIMS)	350
Arizona English Language Learner Assessment (AZELLA)	90

Description of Measures

Curriculum-Based Measurement: Oral Reading Fluency

R-CBM is a measure of oral reading fluency that serves as a general outcome measure of reading performance. Correlations between R-CBM and published assessments of reading comprehension have ranged from .63 to .90, with the majority of correlations falling above .80 (Marston, 1989). CBM has many documented uses, including screening to identify at risk students, evaluating classroom interventions, developing norms, predicting performance on high-stakes tests, and assessing English Language Learners (Deno, 2003). The R-CBM assessment for this study consisted of three standardized reading passages written at the sixth grade level. These passages were developed by Edformation, Inc. and taken from AIMSweb (Howe & Shinn, 2002). The same three passages were used for each of the benchmark assessments (fall, winter, and spring). Students were required to read each passage aloud for one minute, and the median number of words read correctly across the three passages was calculated. The

median score served as the student's final score on the assessment. Students are considered at risk for poor reading outcomes if they perform in the lowest 25th percentile.

Curriculum-Based Measurement: Maze

The Maze assessment has been found to be a reliable, sensitive, and valid measure of reading comprehension for elementary through high school students (Brown-Chisdey, Davis, & Maya, 2003; Shin, Deno, & Espin, 2000). For the current study, students were administered one standardized Maze reading passage at the sixth grade level. The passage selected was also developed by Edformation and taken from the AIMSweb system. In each Maze passage, the first sentence is left intact, with each subsequent seventh word being replaced with three words to choose from (e.g., When (**red/she/told**) was a little girl...). The assessment was administered in a group setting, and the students were given three minutes to complete the passage. This assessment was given once during the 2006-2007 school year in the early spring.

Arizona English Language Learner Assessment (AZELLA)

The AZELLA is a state adopted test that was used to determine the English language proficiency of Arizona K-12 students whose primary home language is one other than English. The AZELLA includes a composite proficiency level score, which includes all of the subtest scores, and also provides separate subtest scores (i.e., Listening, Speaking, Reading, and Total Writing). The AZELLA also included an oral language score, which is a combination of the listening and speaking subtest scores, and a comprehension score, which combines the listening and reading subtest scores. The AZELLA rates performance levels as: (1) Pre-Emergent, (2) Emergent, (3) Basic, (4)

Intermediate, and (5) Proficient. Pre-Emergent is the category for students who have not yet or just begun to learn English. Proficient means that the student is proficient enough to participate in regular classes taught in English. Each student was required to complete the exam once during the school year if they were classified as an English language learner. For the purposes of this study, only the Overall Composite scores will be utilized as a measure of English language proficiency. See Table 3 for the specific descriptions of each performance level for the Composite scores. The reliability coefficients for the overall composite scale were in the high 90s for grades 3 through 12. Accuracy rates ranged from .87 to 1.00, and consistency rates ranged from .81 to 1.00 (AZELLA Technical Manual, 2007).

Table 3 *AZELLA Performance Level Descriptions of Language Proficiency for Sixth Grade*

Score Range	Performance Level	Description
300-553	Pre-Emergent	Student made very few or no responses. Student has very limited or no ability to speak in English. Student may try to communicate with gestures or in a language other than English.
554-574	Emergent	The student has limited ability to speak in English. Student is unable to speak using English grammatical structures and linguistic forms. Many errors and pronunciation difficulties hinder communication.
575-619	Basic	Student is able to speak using satisfactory control over below-grade English grammatical structures and linguistic forms. Student's vocabulary use is accurate but limited.

Table 3 (continued)

Score Range	Performance Level	Description
620-664	Intermediate	Student is able to speak using grade-level English grammatical structures and linguistic forms, although many errors or irregular forms impede communication. Student's vocabulary use is accurate, but ordinary and somewhat limited.
665-900	Proficient	Student is able to speak using grade-level English grammatical structures and linguistic forms, although habitual errors sometimes impede communication. Student's vocabulary is accurate, purposeful, and somewhat varied.

Arizona's Instrument to Measure Standards (AIMS)

The 2007 Spring AIMS assessments were administered in reading, writing, and math to all students in grades 3 through 8. Arizona's Instrument to Measure Standards (AIMS), a Standards-Based test, provided educators and the public with valuable information regarding the progress of Arizona's students toward mastering Arizona's reading, writing and mathematics Standards (Arizona Department of Education). Table 4 shows the Arizona content standards for reading. The AIMS tests were dual purpose assessments, in that both criterion and norm-referenced scores were given based on student performance on the tests. Some of the test items were written by Arizona teachers and items were taken from CTB/McGraw-Hill's norm-referenced test, *Terra Nova, The Second Edition*® (*Terra Nova*; CTB/McGraw-Hill, 2001). It should be noted that this study did not analyze student performance on each reading strand/concept; nor did it use the norm-referenced data. The current study only used the overall aggregate

reading score (criterion-referenced) to determine predictive validity. The AIMS generates a score range in order to determine whether the student exceeds, meets, approached, or falls far below the Arizona state standards. See Table 5 for the AIMS score and performance descriptions.

Table 4 *Arizona Reading Concepts and Strands*

Strand 1: Reading Process

Concept 1: Print Concepts

Concept 3: Phonics

Concept 4: Vocabulary

Concept 6: Comprehension Strategies

Strand 2: Comprehending Literary Text

Concept 1: Elements of Literature

Concept 2: Historical and Cultural Aspects

Strand 3: Comprehending Informational Text

Concept 1: Expository Text

Concept 2: Functional Text

Concept 3: Persuasive Text

Table 5 *AIMS Performance Level Descriptions*

Score Range	Performance Level	Description
571-690	Exceeds the Standard = ES	Students demonstrate strong analytical and inferential skills in comprehending more challenging and complex text. Students are able to identify various literary elements, describe how character traits influence actions, and analyze the movement toward climax.
478-570	Meets the Standard = MS	Students are able to identify the appropriate appropriate reference sources required for a specific purpose, use organizational features to locate information, and differentiate between important and unimportant details. Students are able to determine the author's position and recognize valid support for a given argument.
433-477	Approaches the Standard = AS	Students show a basic understanding of making predictions, identifying cause and effect, and drawing conclusions. Students can identify the main idea, analyze specific facts, and distinguish between fact and opinion. Some gaps in knowledge and skills are apparent and may require additional instruction to achieve a satisfactory level of understanding.
250-432	Falls Far Below the Standard = FFBS	Students have significant gaps and limited knowledge and skills necessary to meet the state's reading standard. Students will usually require a large amount of additional instruction and remediation to achieve a satisfactory level of understanding.

Table 6 *Types of Data and Assessments*

Formative Assessments	Summative Assessments	Norm-Referenced: Percentile Ranks	Continuous Data: Numerical Scores	Criterion-Referenced Risk Status or Performance Levels
R-CBM		√	√	√
Maze		√	√	√
	AIMS			√
	AZELLA			√

Data Analysis

Recent studies have begun to use an analysis tool that allows for explicit modeling of individual differences in growth (Silbergliitt & Hintze, 2007; Stage & Jacobsen, 2001; Shin, Deno, & Espin 2000). Hierarchical Linear Modeling (HLM) (Raudenbush & Bryk, 2002) uses individuals' slopes as the unit of analysis, as opposed to analysis of variance which examines group change over repeated measures. According to Shin et al. (2004), HLM allows researchers to investigate students' academic growth using more flexible and practical research designs as opposed to those available with Repeated Measures ANOVA or Structural Equation Modeling (SEM). Shin et al. (2004) cite several reasons for the use of HLM when examining academic skills development and its relationship correlates which include: 1) HLM does not require all students to be assessed at the same time; 2) HLM allows for both continuous and categorical predictors

to be used in examining the relationships between growth rates and correlates; and 3) In examining the relationships between growth rates and correlates, HLM allows differential weights to be used (Bryk & Raudenbush, 1992 as cited in Shin et al., 2004). The analysis of students' individual slopes allows for the trajectory of student's growth on oral reading fluency to be evaluated as opposed to the average change in explained variance between measurement intervals (Stage & Jacobsen, 2001). Raudenbush and Bryk (2002) also indicate that HLM more accurately takes into account under- and over-estimation of observed relationships between variables.

In order to examine student growth and examine the group difference (ELL versus non-ELL), multilevel modeling or HLM will be utilized. In this analysis, assessment scores over time (Fall CBM, Winter CBM, and Spring CBM) will be treated as the within-subject factor and the grouping variable (e.g., ELL vs. non-ELL) will be treated as the between-subject factor. Traditionally, ANOVA repeated measures are employed for this type of trend-based data. However, as discussed earlier, ANOVA repeated measures have several limitations, particularly by its assumption of a single covariance structure, specifically compound symmetry. Conversely, multi-level modeling permits various covariance matrix structures, including Toeplitz, heterogeneous AR, auto-regressive, heterogeneous compound symmetry, and several others. Moreover, mixed-level modeling employs maximum likelihood estimation, which is more accurate than the sum of squares approach in a generalized linear model. In addition, the multiple fitness criteria such as Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), AAIC, etc. can be used to evaluate the goodness of a multi-level model.

In order to examine the first three research questions posed in this study, a growth curve analysis using HLM will be conducted using SAS (2009). The analysis will model individual student slopes in R-CBM across fall, winter and spring, generating an overall average slope in R-CBM for the student population. The data will be represented in a two-level hierarchical model. In the first level of analysis, the goal will be to generate individual slope and intercept estimates for each student. At level 1, each participant's reading skill development will be represented by their initial status on R-CBM and their growth trajectories. The level 2 analysis will examine whether person level characteristics (ELL vs. non-ELL) will provide any information that could predict differences in the individual slopes and intercepts. Finally, in order to examine the growth model of RCBM conditioned by AIMS and ELL, a multilevel model will be employed examining the fixed effects of students' AIMS level (Level 1= Falls Far Below, Level 2 = Approaches the Standard, Level 3 = Meets the Standard, and Level 4 = Exceeds the Standard), students' rate of growth on R-CBM from Fall to Spring, and their ELL status (ELL versus non-ELL). Displaying assessment scores by group in histograms would be inadequate to reveal the interaction among several variables, as well as the varying slopes across different groups and different points in time. In order to remedy this issue, multi-dimensional visualization will be utilized to examine the complexity of these relationships among all of the variables. Parallel coordinates will be created using Datadesk (2008).

In order to address the last research question, multiple regression will be utilized using SAS and TIBCO Spotfire Miner (2009). Since the AIMS Reading scaled score is

a cross-sectional measure (one-time measure) and there are no nested data in between-subject factors (e.g., school district, school, teacher ...etc), regression instead of HLM is employed to examine how RCBM, MAZE, and ELL could predict AIMS Reading scaled score. A regression equation is developed for use with scores on a set of predictors to predict performance on a criterion. The predictors in this study are R-CBM Fall score (initial status), growth on R-CBM from Fall to Spring, ELL status, and Maze. The criteria of which performance is to be predicted on are the spring reading score of the Arizona Instrument to Measure Standards (AIMS). In other words, the regression analysis of the dataset included utilizing to evaluate overall predictive ability of R-CBM initial status, R-CBM growth, ELL status and Maze on student performance on the high-stakes reading test (AIMS) in the spring.

In order to apply a valid regression model, several assumptions for the data should be met. According to Berry and Feldman (1985), most regression assumptions are concerned with residuals. The current data analysis will examine the whether the residuals are random and normally distributed in this population. Multi-collinearity, or co-dependence of variables, will also be examined. In regression when several predictors are highly correlated, it will cause multi-collinearity. In a regression model, the higher the variance explained (R-square), the better the model. However, if multi-collinearity exists, most likely the variance, standard error, and parameter estimates are inflated. Variance Inflation Factor (VIF) can be used for detecting multi-collinearity. Generally, the VIF should not exceed 10 (Belsey, Kuh, & Welsch, 1980).

As such, the above predictors will be analyzed using HLM, multiple regression, and data visualization in attempts to address the following:

1. What was the average rate of growth of the current population over one school year on the reading fluency R-CBM measures?
2. What is the relationship among the rate of growth on R-CBM and level of language proficiency?
3. If students vary on their growth rates (slopes) on R-CBM and/or intercept (performance on state AIMS reading assessment), is this variation related to ELL status?
4. Is R-CBM initial status, R-CBM rate of growth, and Maze predictive of reading achievement on the spring AIMS reading assessment relative to ELL status? Which of these factors is a more potent predictor of reading achievement on the AIMS reading assessment?

Summary

The current study will examine data for ELL and non-ELL sixth grade students in a school district in west Phoenix, Arizona to determine the predictive validity of R-CBM oral reading fluency and Maze assessments with a high-stakes measure, AIMS. This is a non-experimental study using archival data from the 2006-2007 school year. The students' level of language proficiency and how it relates to reading growth and performance will also be examined. In an attempt to avoid statistical and interpretive problems with the data analysis, HLM, multiple regression, and data visualization will be used. This study will be conducted for the purpose of examining the typical rate of

growth on R-CBM over one school year for ELL and non-ELL students. It will also assist in determining whether reading screening measures (R-CBM, Maze) are effective in predicting future reading success or failure for this population of students.

CHAPTER IV

RESULTS

The purpose of the present study was to examine the predictive ability of oral reading fluency (R-CBM) on a sixth grade high-stakes assessment with ELL and non-ELL students, as well as determine the average rate of growth on R-CBM and how that relates to level of English Proficiency. Hierarchical Linear Modeling (HLM) was used in this study to determine if students varied significantly in initial status (Fall R-CBM score) and in growth rate (slope) from Fall to Spring on the fluency measure of Curriculum-Based Measurement, and whether the students' level of language proficiency affected that initial status or growth. In addition, if there was a significant variation among students in their initial status (Fall R-CBM) and growth rate on R-CBM, further analysis would determine if this variation was affected by ELL status and/or level of AIMS performance. The final analysis used multiple regression to examine the predictive validity of R-CBM initial status, R-CBM rate of growth from Fall to Spring, ELL status, and the Maze assessment on student achievement as measured by the AIMS reading score. Further regression analyses also examined the predictive ability of R-CBM initial status, R-CBM rate of growth, and Maze on the outcome measure AIMS for ELL versus non-ELL students.

The current study planned to examine a student's level of language proficiency as measured by the AZELLA and how that affected a student's initial status or growth as

measured by R-CBM. Table 7 shows the frequencies for the students in each of the five levels of AZELLA: (1) Pre-Emergent, (2) Emergent, (3) Basic, (4) Intermediate, and (5) Proficient. There are no students in Level 2, Level 1 only has 2 students, Level 3 has 8 students, and Level 4 has 80 students. As such, the level of language proficiency was not utilized in this analysis. Students were grouped into two categories, ELL versus non-ELL. If students are classified ELL upon entering school per parent report, they are required by the state to take the AZELLA, and there were a total of 90 ELL students in the current sample.

Table 7 *Frequency and Percent of ELL Students in AZELLA Levels 1-5*

AZELLA Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1 (Pre-Emergent)	2	.6	2	2.2
2 (Emergent)	0			
3 (Basic)	8	2.3	10	11.1
4 (Intermediate)	80	22.9	90	100.0
5 (Proficient)	0			

In examining the R-CBM scores through data visualization, an outlier was detected. Figure 1 shows the outlier. This subject was removed from the data analysis.

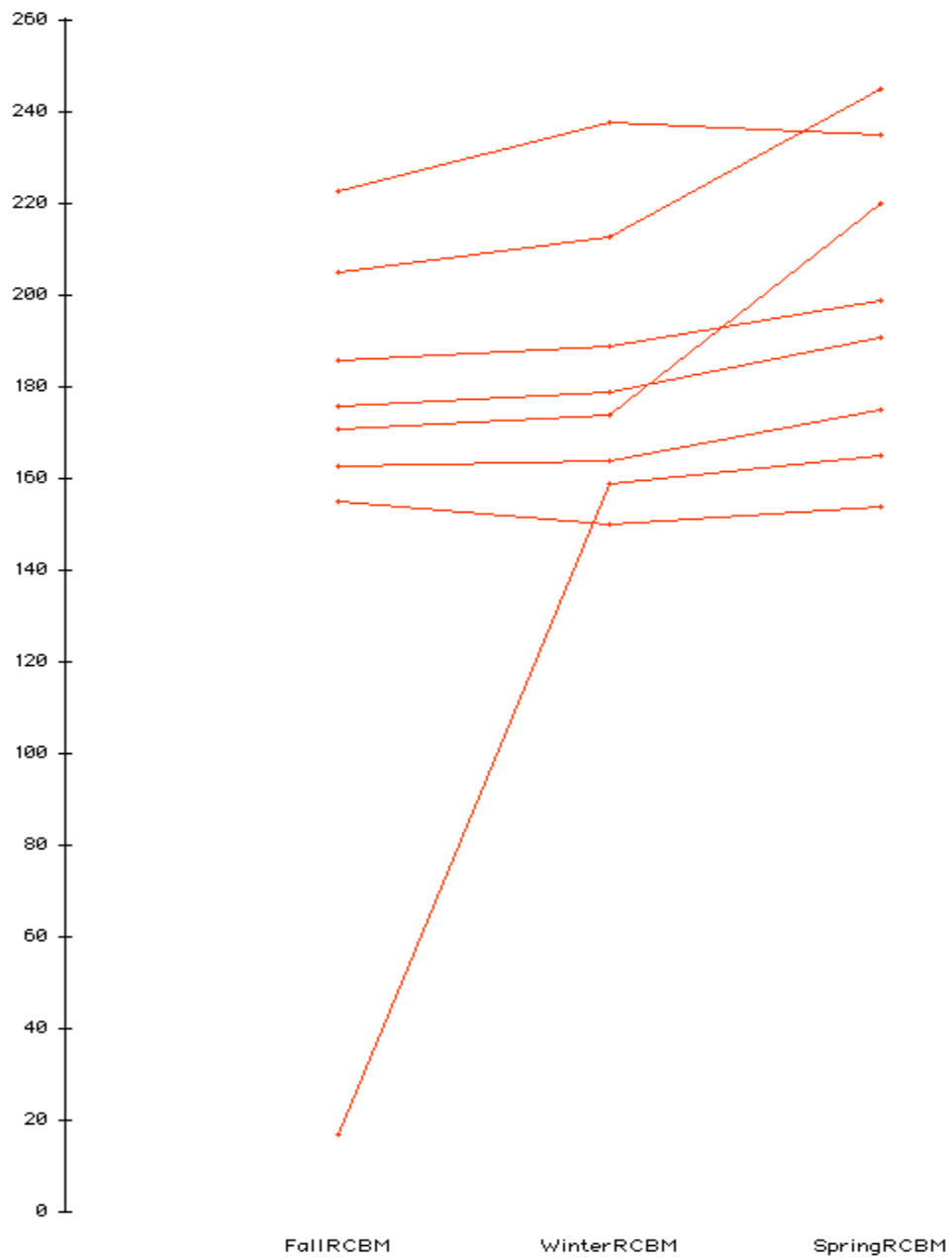


Figure 1. R-CBM data for all students that Fell Far Below the standard on the AIMS assessment. The outlier was removed from the data analysis.

Descriptive Statistics

Table 8 provides descriptive statistics of the continuous data which includes raw scores/benchmarks and percentile rankings for R-CBM Oral Reading Fluency. Of the 349 R-CBM scores within the sample, the mean score for Fall was 113.21, with a standard deviation of 38.71. The mean score for R-CBM Winter was 125.76, with a standard deviation of 38.70. The mean score for R-CBM Spring was 134.00, with a standard deviation of 38.08. The mean score for the Maze assessment was 19.23, with a standard deviation of 5.86.

Table 8 Means and Standard Deviations of Fall, Winter, and Spring R-CBM Scores

Measure	# of Students	Mean	SD
Fall R-CBM	349	113.21	38.71
Winter R-CBM	349	125.76	38.70
Spring R-CBM	349	134.00	38.08
MAZE	349	19.23	5.86

Table 9 provides descriptive statistics of the continuous data which includes raw scores/benchmarks for R-CBM Oral Reading Fluency for ELL and non-ELL students. Of the R-CBM scores within the ELL sample, the mean score for Fall was 84.78, with a standard deviation of 35.29. The mean score for R-CBM Winter was 97.79, with a standard deviation of 32.57. The mean score for R-CBM Spring was 108.47, with a standard deviation of 35.15. The mean score for the Maze assessment was 16.52 with a

standard deviation of 4.21. Of the R-CBM scores within the non-ELL sample, the mean score for Fall was 123.09 with a standard deviation of 34.81. The mean score for Winter was 135.47 with a standard deviation of 35.86. The mean score for Spring was 142.88 with a standard deviation of 34.97. The mean score for the Maze assessment was 20.17 with a standard deviation of 6.07.

Table 9 Means and Standard Deviations of Fall, Winter, and Spring R-CBM Scores for ELL and Non-ELL Students

Variable	Label	Non-ELL Students		
		N	Mean	SD
FallRCBM	FallRCBM	259	123.09	34.81
WinterCBM	WinterCBM	259	135.47	35.86
SpringRCBM	SpringRCBM	259	142.88	34.97
MAZE	MAZE	259	20.17	6.07
Variable	Label	ELL Students		
		N	Mean	SD
FallRCBM	FallRCBM	90	84.78	35.29
WinterRCBM	WinterRCBM	90	97.79	32.57
SpringRCBM	SpringRCBM	90	108.47	35.16
MAZE	MAZE	90	16.52	4.21

The means and standard deviations for each R-CBM measurement and MAZE by AIMS level scores are shown in Table 10. For students at AIMS Level 1, the Fall mean score was 67.16 with a standard deviation of 32.99, the Winter mean was 79.91 with a standard deviation of 30.77, the Spring mean was 86.28 with a standard deviation of 33.54, and the mean Maze score was 15.19 with a standard deviation of 3.78. For students at AIMS Level 2, the Fall mean score was 94.49 with a standard deviation of 31.59, the Winter mean was 107 with a standard deviation of 30.17, the Spring mean was 116.09 with a standard deviation of 29.76, and the mean Maze score was 17.26 with a standard deviation of 4.62. For students at AIMS Level 3, the Fall mean score was 129.31 with a standard deviation of 30.29, the Winter mean was 142.15 with a standard deviation of 31.82, the Spring mean was 149.93 with a standard deviation of 29.60, and the mean Maze score was 20.62 with a standard deviation of 5.76. For students at AIMS Level 4, the Fall mean score was 182.71 with a standard deviation of 24.05, the Winter mean was 186.71 with a standard deviation of 30.03, the Spring mean was 202.71 with a standard deviation of 32.71, and the mean Maze score was 31.43 with a standard deviation of 6.56.

Table 10 Means and Standard Deviations for Oral Reading Fluency Scores at Fall, Winter and Spring by AIMS Level

Variable	AIMS Level 1 = Falls Far Below			
	Label	N	Mean	SD
FallRCBM	FallRCBM	32	67.16	32.99
WinterRCBM	WinterRCBM	32	79.91	30.77
SpringRCBM	SpringRCBM	32	86.28	33.54
MAZE	MAZE	32	15.19	3.78
Variable	AIMS Level 2 = Approaches			
	Label	N	Mean	SD
FallRCBM	FallRCBM	115	94.49	31.59
WinterRCBM	WinterRCBM	115	107.00	30.17
SpringRCBM	SpringRCBM	115	116.09	29.76
MAZE	MAZE	115	17.26	4.62
Variable	AIMS Level 3 = Meets			
	Label	N	Mean	SD
FallRCBM	FallRCBM	195	129.31	30.29
WinterRCBM	WinterRCBM	195	142.15	31.82
SpringRCBM	SpringRCBM	195	149.93	29.60
MAZE	MAZE	195	20.62	5.76

Table 10 (continued)

Variable	Label	AIMS Level 4 = Exceeds		
		N	Mean	SD
FallRCBM	FallRCBM	7	182.71	24.05
WinterRCBM	WinterRCBM	7	186.71	30.03
SpringRCBM	SpringRCBM	7	202.71	32.71
MAZE	MAZE	7	31.43	6.56

CBM Models

HLM allows for individual growth curves to be modeled as a function of time using the repeated measures obtained from each student. These individual growth curves are compared to see if they differ significantly from one another in initial status (Fall R-CBM score) and in growth rate. If there are differences among children in their R-CBM Fall score or in their rate of growth, then student's characteristics can be examined that explain or predict the differences among children in their initial status and growth rates. The predictor included: Level of Language Proficiency (ELL vs. non-ELL).

The CBM Model #0 is an unconditional means model with no predictors at level 2. This model answers the first research question: "What was the average rate of growth of the current population over one school year on the reading fluency R-CBM measures? It is the average growth model for all of the subjects for the CBM from Fall to Spring. Table 11 below of the final estimation of fixed effects specifies that the mean CBM score for students in the fall (initial status) is 103.51 words read correctly per minute. The

coefficient of 10.36 indicates that the change in slope was 10.36 words per minute across the measurement times. Therefore, on average, the students read 103.51 words per minute in Fall, 113.87 words per minute in Winter ($103.51 + 10.36$), and 124.23 words per minute in Spring ($113.87 + 10.36$). These coefficients are both significantly different from zero.

Table 11 *R-CBM Unconditional Growth Model (from Fall to Spring)*

Fixed Effect	Coefficient	SE	DF	t ratio	p value
Initial Status	103.51	1.45	347	71.51	0.00
Growth Rate	10.36	0.86	698	12.10	0.00

Table 12 gives the estimates of variance for the intercept and slope before the predictors were added. The variance components indicate the estimates of the variation in initial status and slope across the individually estimated growth curves for the students. The data indicates that initial status is significantly different between the students. In addition, the growth rates of the students are also significantly different. The students, therefore, grow at significantly different rates as well as have significantly different starting points on the fall CBM assessment.

Table 12 *Final Estimation of Variance Components for the CBM Unconditional Growth Model*

Random Effect	Standard Error	Variance Component	df	Chi-square	P-value
Initial Status	37.56	1410.92	347	6138.31	0.00
Growth Rate	2.96	8.74	347	406.38	0.02
Level-1	10.08	101.63			

The CBM model #1 is a model with a Level 2 predictor, ELL versus non-ELL. This model answers the second question, “what is the relationship among the rate of growth on R-CBM and level of English Language Proficiency?” It compares both the initial status (Fall R-CBM) and the growth rates of ELL versus non-ELL students for the CBM from Fall to Spring. Table 13 below of the final estimation of fixed effects specifies that the mean CBM score for non-ELL students in the fall (initial status) is 114.01 words read correctly per minute (wrc) with an average growth rate (slope) of 9.87 wrc from Fall to Spring. It shows the mean initial status for non-ELL students was higher than that for ELL students by 40.68 words read correct on R-CBM. The mean difference of initial status between the two groups (ELL versus non-ELL) was statistically significant. ELL students’ growth rate from Fall to Spring on R-CBM, on average, was 1.98 correct words more than non-ELL students. The mean difference in growth rate between the two groups (ELL versus non-ELL) was not significant. Thus,

while ELL students score significantly lower in the Fall on R-CBM, they do not vary in their growth trajectories from non-ELL students.

Table 13 *Differences in Initial Status and Growth Rates Between ELL and Non-ELL Students*

Fixed Effect	Coefficient	SE	<i>t</i> ratio	DF	<i>p</i> value
Model for initial status					
Intercept	114.01	4.75	4.04	346	0.00
ELL	-40.68	3.06	-13.31	351	0.00
Model for growth rate					
Intercept	9.87	0.91	10.83	346	0.00
ELL	1.98	1.79	1.10	351	0.27

The final estimation of the variance components for CBM Model #1 are displayed below in Table 14. R-CBM initial status intercepts and rate of growth continue to be significantly different between the students with a variance component. The corresponding χ^2 test statistic for the hypothesis that there are no individual differences among ELL and non-ELL students initial status and growth rates was 848.92 (df = 1, $p < .0001$). It can be concluded that there is significant variation in students' initial status and growth rates. The percent of variance accounted for by the predictor (ELL status) for initial status and growth rate is 0.18 (126.30-103.20 divided by 126.30).

Table 14 *Final Estimation of Variance Components for ELL Status*

Random Effect	Standard Error	Variance Component	df	Chi-square	P-value
Initial Status/ Growth Rate	8.42	103.20	1	848.92	0.0001
Level-1	11.93	223.07			

AIMS Models

The AIMS models attempt to examine the relationship among the four levels of AIMS reading and initial status and growth rates on R-CBM. Students fall in to one of four performance categories for the AIMS assessment. Students at AIMS level 1 fall far below the standard, students at AIMS level 2 approach the standard, students at AIMS level 3 meet the standard, and students at AIMS level 4 exceed the standard. Table 15 shows the analysis of variance on R-CBM initial status by AIMS level, ELL status, and time or rate of growth. While AIMS, ELL and rate of growth all had a significant effect on R-CBM, none of the interaction effects (e.g., ELL by time, etc) were significant.

Table 15 ANOVA for AIMS, ELL Status, and Time on R-CBM Initial Status

Effect	DF	F Value	Pr>F
AIMS	3	90.73	.0001
ELL	1	26.92	.0001
Time	2	4.57	.0106
ELL*Time	2	.02	.9758
AIMS*Time	6	.09	.9974
AIMS*ELL	3	1.57	.1952
AIMS*ELL*Time	6	.15	.9890

Table 16 shows initial status and growth rates by AIMS level. Results show that, overall, students vary significantly based on their AIMS level in their initial status as measured by the fall R-CBM score. Students' initial status (Fall R-CBM) at AIMS level 1 (Falls Far Below) was 63.21 words read correct. The average amount of words read correct varied by 33.62 at each AIMS level. However, students do not vary significantly on their growth rates relative to their performance level on AIMS. The average rate of growth for students' at AIMS level 1 was 20.7 words read correct from Fall to Spring. At each AIMS level (1-4), students' rate of growth varied, on average, .06 words read correct from Fall to Spring.

Table 16 *Differences in Initial Status and Growth Rates by AIMS Reading Level*

Fixed Effect	Coefficient	SE	<i>t</i> ratio	DF	<i>p</i> value
Model for initial status					
Intercept	63.21	3.99	15.82	346	0.00
AIMS	33.62	2.41	13.93	346	0.00
Model for growth rate					
Intercept	10.35	0.99	10.39	346	0.00
AIMS	0.03	0.60	0.047	346	0.96

The final estimation of the variance components for AIMS Model #1 are displayed below in Table 17. R-CBM initial status intercepts and rate of growth continue to be significantly different between the students after accounting for the AIMS level of the students. The corresponding χ^2 test statistic for the hypothesis that there are no individual differences among initial status for students by AIMS level was 3927.47 (df = 346, $p < .0001$). The corresponding χ^2 test statistic for the hypothesis that there are no individual differences among growth rates for students by AIMS level was 406.37 (df=346, $p < .05$). It can be concluded that there is significant variation in students' initial status and growth rates remaining after accounting for the AIMS level of the students. The percent of variance accounted for by the predictor (AIMS level) for initial status is .38 (1410.91-876.17 divided by 1410.92). The percent of variance accounted for by the predictor (AIMS level) for growth rate is -.02 (8.74 -8.91 divided by 8.74).

Table 17 *Final Estimation of Variance for AIMS Level*

Random Effect	Standard Error	Variance Component	df	Chi-square	P-value
Initial Status	29.60	876.17	346	3927.47	0.00
Growth Rate	2.99	8.91	346	406.37	0.01
Level-1	10.08	101.64			

Table 18 shows the overall fixed effects for both within-subject and between subject factors. Each fixed effect is labeled by the variable to which it refers, and is accompanied by a *t*-test, degrees of freedom, and *p*-value. Both ELL status and AIMS level related significantly to initial status as measured by R-CBM. However, only ELL status was significantly related to rate of growth on R-CBM from Fall to Spring. AIMS level did not relate significantly to students' rate of growth.

Table 18 *ELL Status and AIMS Reading Level Three-way Interaction Effect on R-CBM Initial Status and Growth Rate*

Fixed Effect	Coefficient	SE	<i>t</i> ratio	DF	<i>p</i> value
Model for initial status					
Intercept	76.64	4.61	16.63	345	0.00
AIMS	28.32	2.53	11.19	345	0.00
ELL	-21.05	3.98	-5.29	345	0.00
Model for growth rate					
Intercept	8.86	1.19	7.47	345	0.00
AIMS	0.62	0.65	0.95	345	0.35
ELL	2.34	1.02	2.28	345	0.02

The final estimation of the variance components for AIMS Model #2 are displayed below in Table 19. R-CBM initial status intercepts and rate of growth continue to be significantly different between the students with a variance component. The corresponding χ^2 test statistic for the hypothesis that there are no individual differences among initial status for students by AIMS level and ELL status was 3632.18 (df=345, $p < .00$). It can be concluded that there is significant variation in students' initial status. The corresponding χ^2 test statistic for the hypothesis that there are no individual differences among rate of growth for students by AIMS level and ELL status was 400.32 (df=345, $p < .02$). It can be concluded that there is significant variation in students' growth rates. The percent of variance accounted for by the predictors for the intercept is 0.43 (1410.92-

806.64 divided by 1410.92). CBM slopes continue to be significantly different between the students. The percent of variance accounted for by the predictors for the slope is 0.06 (8.74-8.19 divided by 8.74).

Table 19 *Final Estimation of Variance for AIMS Level and ELL Status*

Random Effect	Standard Error	Variance Component	df	Chi-square	P-value
Initial Status	28.40	806.64	345	3632.18	0.00
Growth Rate	2.86	8.19	345	400.32	0.02
Level-1	10.08	101.63			

In order to further examine the complex relationships among the variables, data visualization was utilized. Figures 2 through 5 display the parallel coordinates of individual students' rates of progress on R-CBM from Fall to Spring relative to their AIMS performance level. It appears that students at the lowest level of performance on AIMS (Level 1 Falls Far Below) and students at the highest level of performance on AIMS (Level 4 Exceeds) exhibited slower rates of linear growth from Fall to Spring on R-CBM. Thus, for students who are exceeding standards on AIMS, their level of reading fluency may have hit a ceiling. For students who are falling far below standards on AIMS, their R-CBM scores in Fall are significantly lower and they do not appear to make much progress through the spring. The two middle levels on AIMS (Level 2 Approaches and Level 3 Meets) showed the most consistent linear growth rates from Fall to Spring on R-CBM.

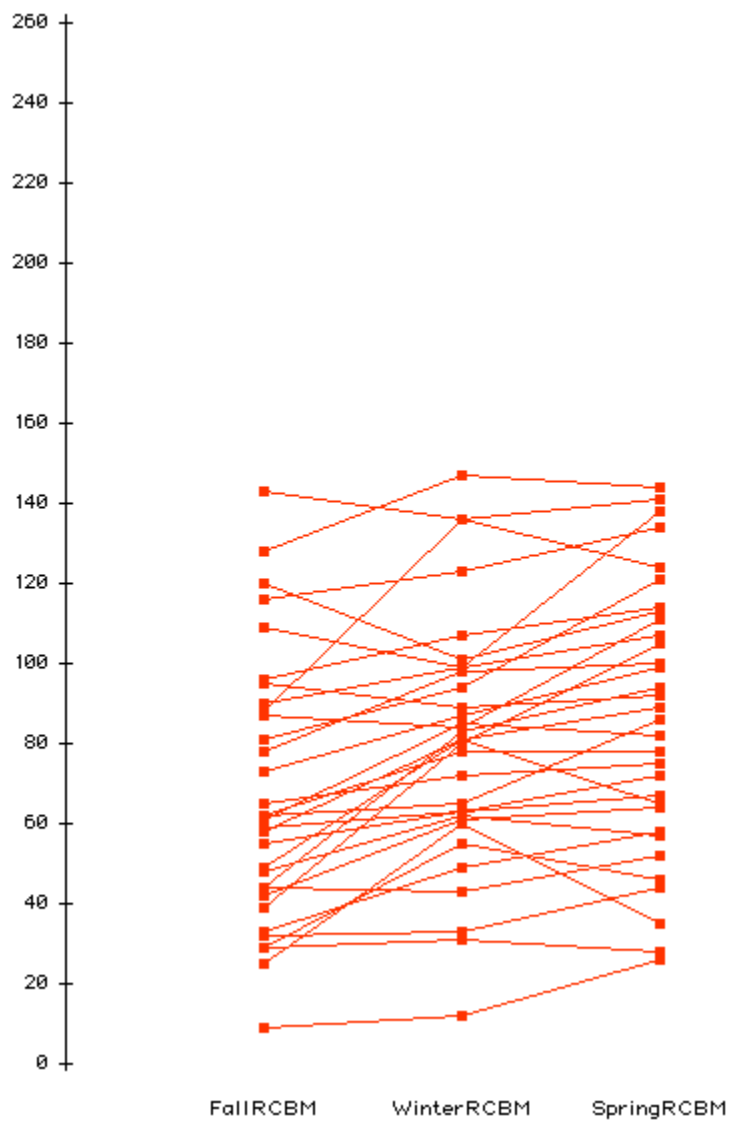


Figure 2. Parallel coordinate of individual student's growth from Fall to Spring in R-CBM data by AIMS level 1 (Falls far below).

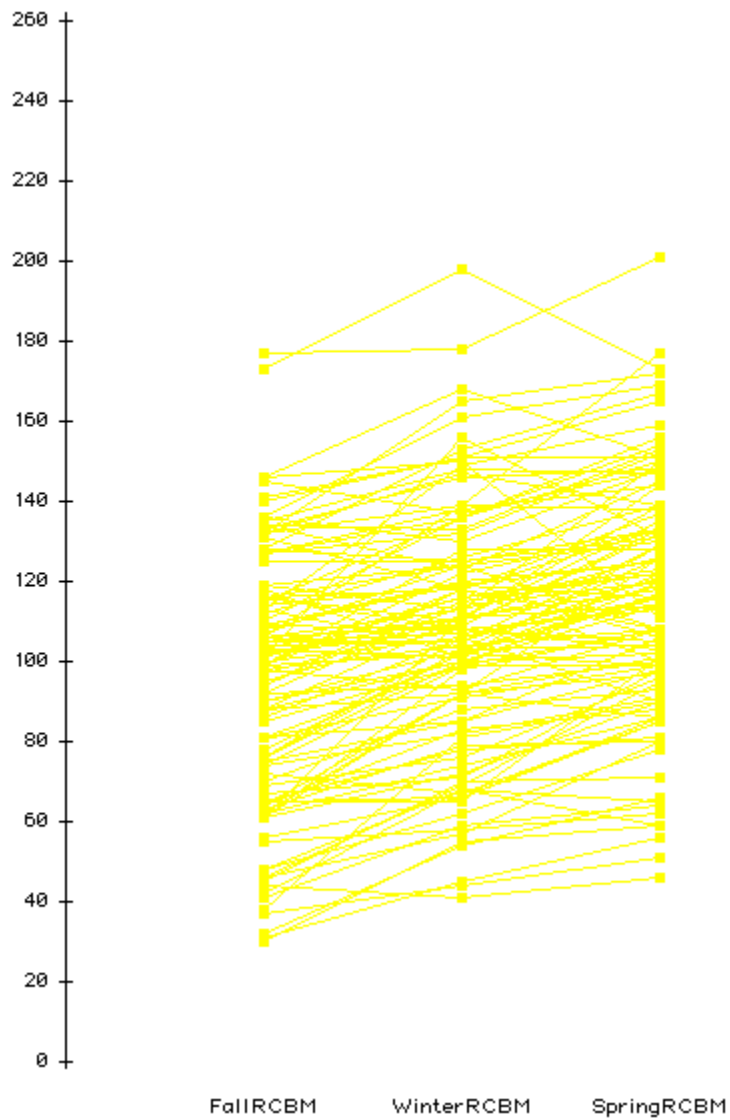


Figure 3. Parallel coordinate of individual student's growth from Fall to Spring in R-CBM data by AIMS level 2 (Approaches the standard).

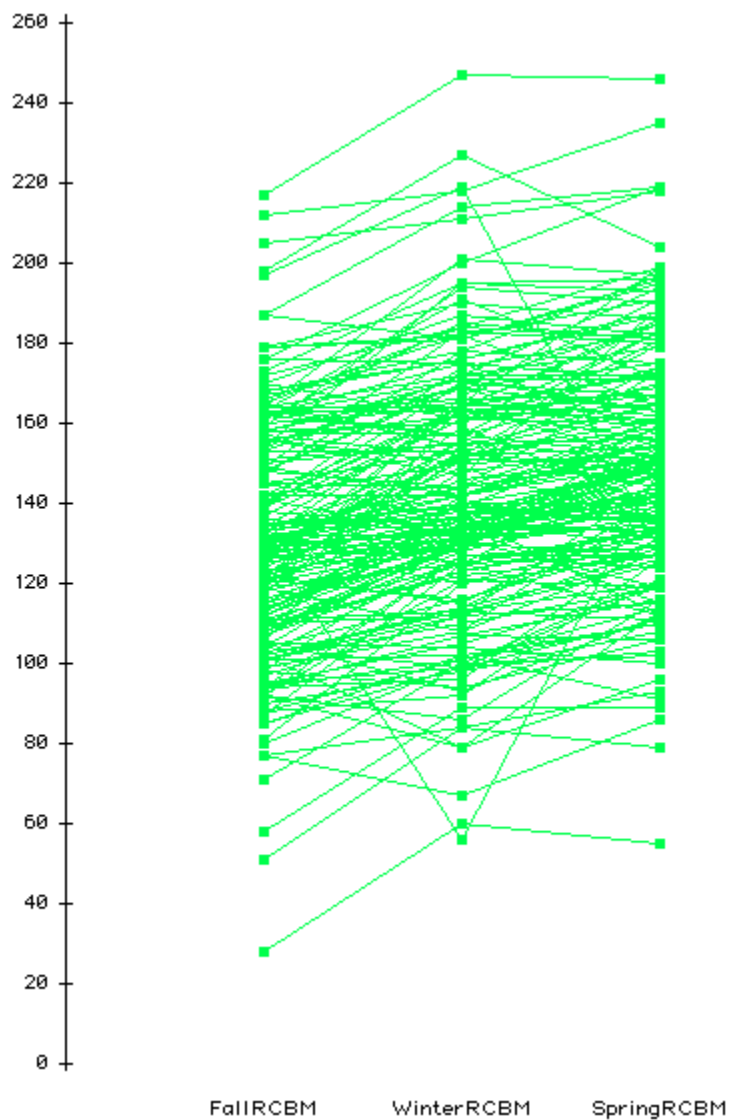


Figure 4. Parallel coordinate of individual student's growth from Fall to Spring in R-CBM data by AIMS level 3 (Meets the standard).

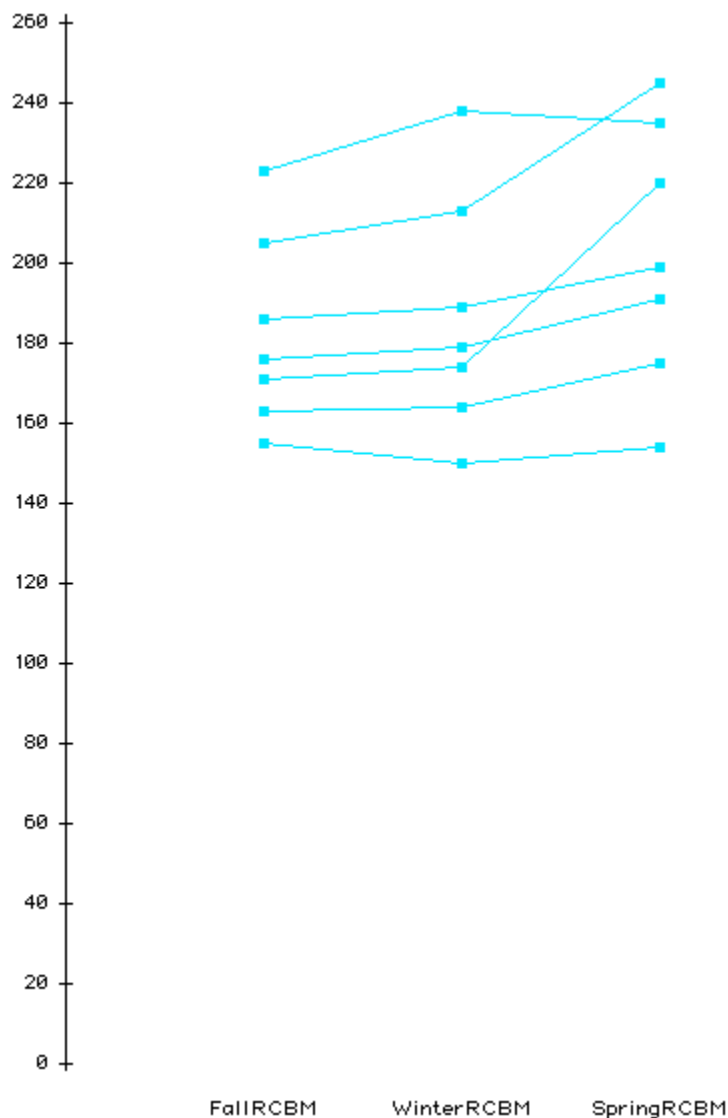


Figure 5. Parallel coordinate of individual student's growth from Fall to Spring in R-CBM data by AIMS level 4 (Exceeds the standard).

Figures 6 and 7 display individual students' rates of progress on R-CBM from Fall to Spring relative to their ELL status. It is evident from the data visualization that ELL students' R-CBM scores in general are lower than their non-ELL peers, and their

initial status (Fall R-CBM) scores are lower. However, both groups (ELL and non-ELL) appear to have similar growth trajectories on R-CBM from Fall to Spring.

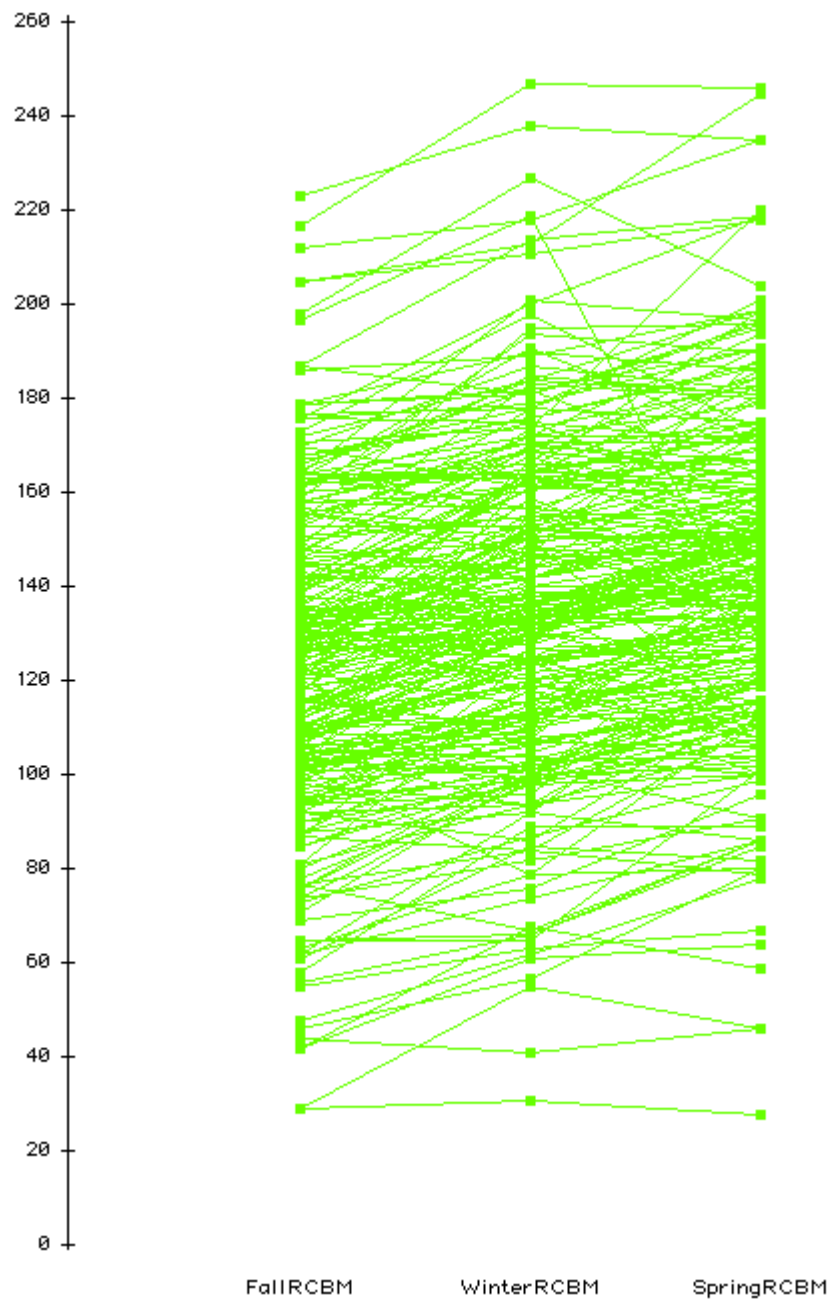


Figure 6. Parallel coordinate of non-ELL students' growth on R-CBM from Fall to Spring.

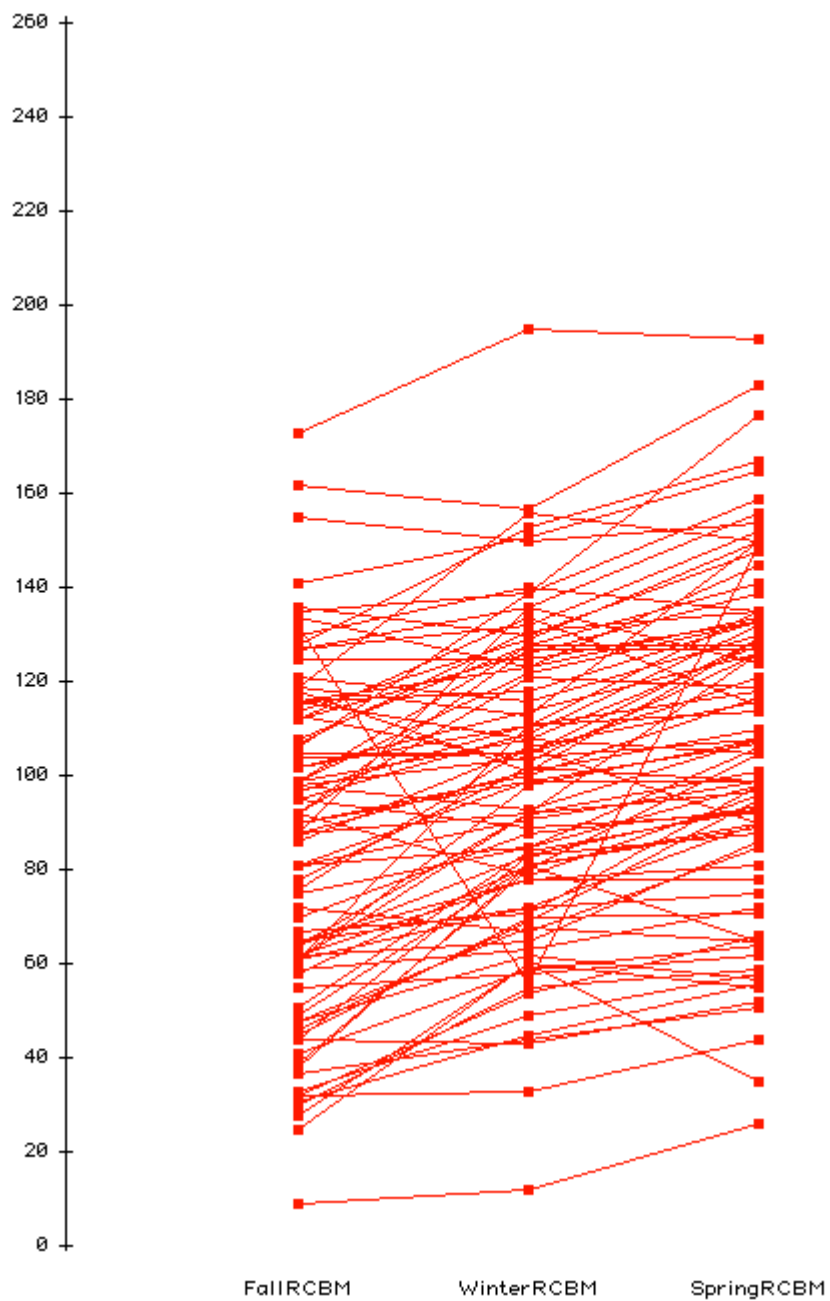


Figure 7. Parallel coordinate of ELL students' growth on R-CBM from Fall to Spring.

Figures 8 through 13 display individual students' rates of progress on R-CBM from Fall to Spring by AIMS level relative to their ELL status. Only the graphics that

were meaningful were reported. Figure 8 shows ELL students' growth from Fall to Spring that were at AIMS level 1, or falls far below the standard. It is apparent from the data visualization that most ELL students at AIMS level 1 showed slow growth from Fall to Spring, but there were a few students who made substantial growth from Fall to Winter but then dropped down from Winter to Spring. Figure 9 shows the non-ELL students at Level 1 and their growth from Fall to Spring on R-CBM. For these students in general, their growth trajectories were flat, in that most students did not make much growth from Fall to Spring. Unlike the ELL students at Level 1, the non-ELL students did not appear to drop substantially from Winter to Spring. For both ELL and non-ELL students at AIMS Level 2 (approaches the standard), there seems to be a general upward trend from Fall to Spring on R-CBM as shown in Figures 10 and 11. For ELL students at AIMS level 3 (meets the standard), a consistent pattern in growth from Fall to Spring does not emerge, which would not be expected given these students met the standard on the AIMS reading test (see Figure 12). However, for non-ELL students at AIMS Level 3, a consistent pattern of growth emerged, where most students grew consistently from Fall to Spring on R-CBM (see Figure 13). This pattern would be expected given that these students met the standard on the AIMS state assessment in the spring.

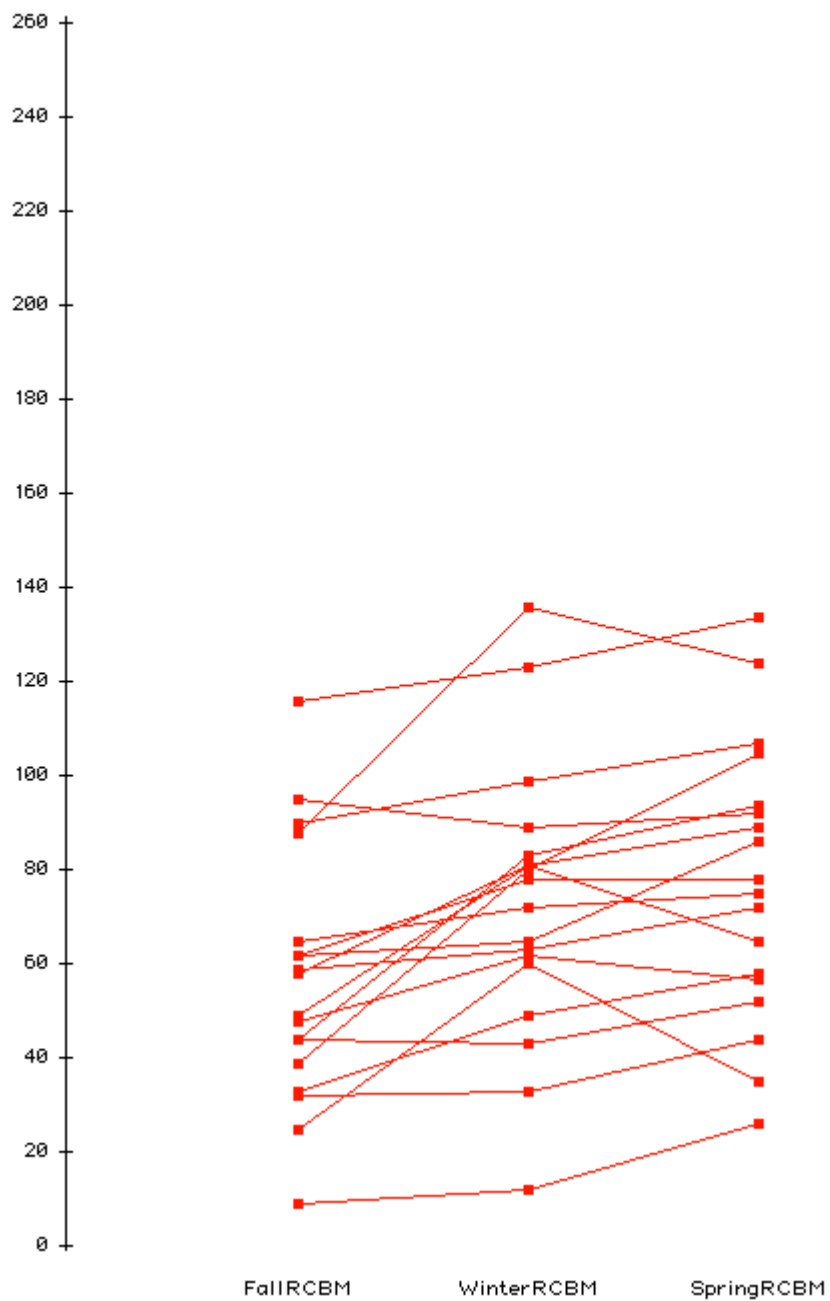


Figure 8. Parallel coordinate of ELL students' growth on R-CBM from Fall to Spring at AIMS level 1 (Falls far below).

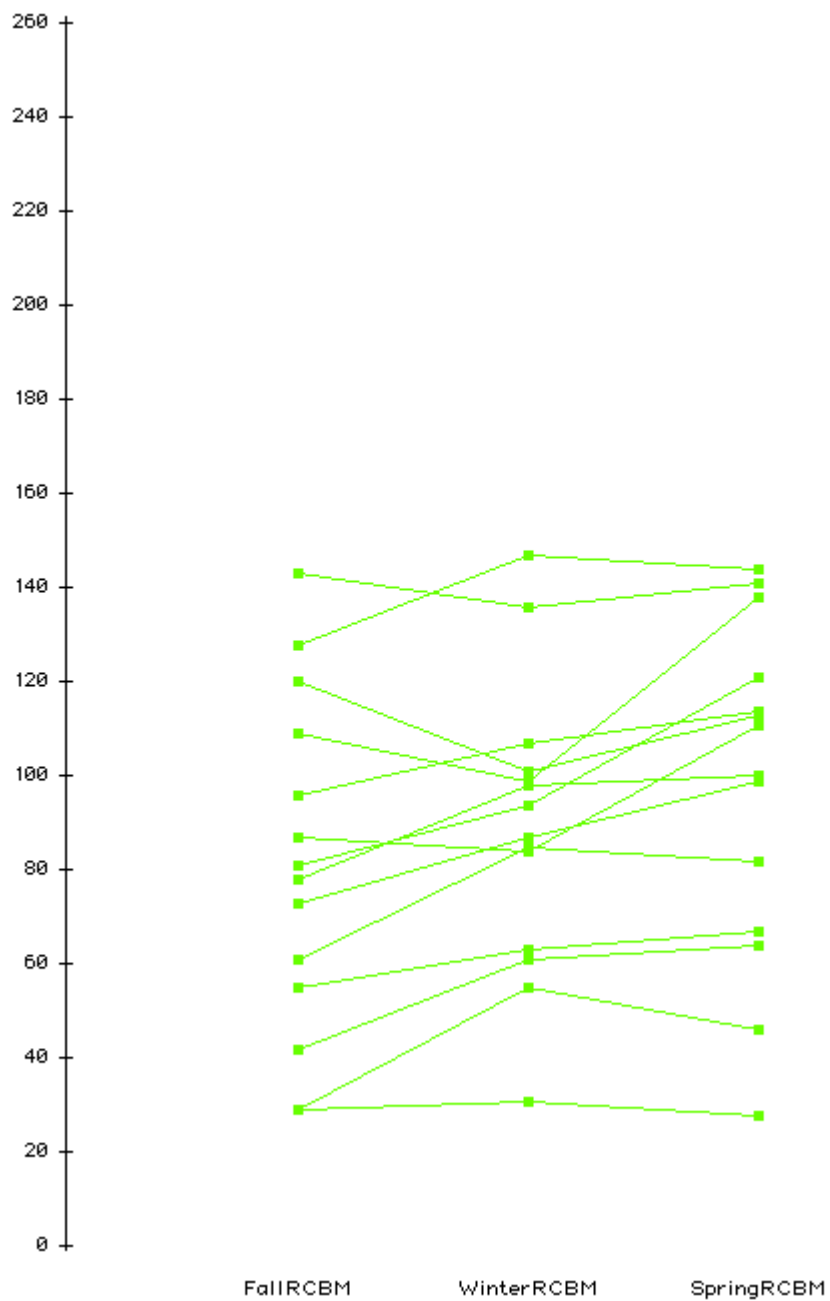


Figure 9. Parallel coordinate of non-ELL students' growth on R-CBM from Fall to Spring at AIMS level 1 (Falls far below).

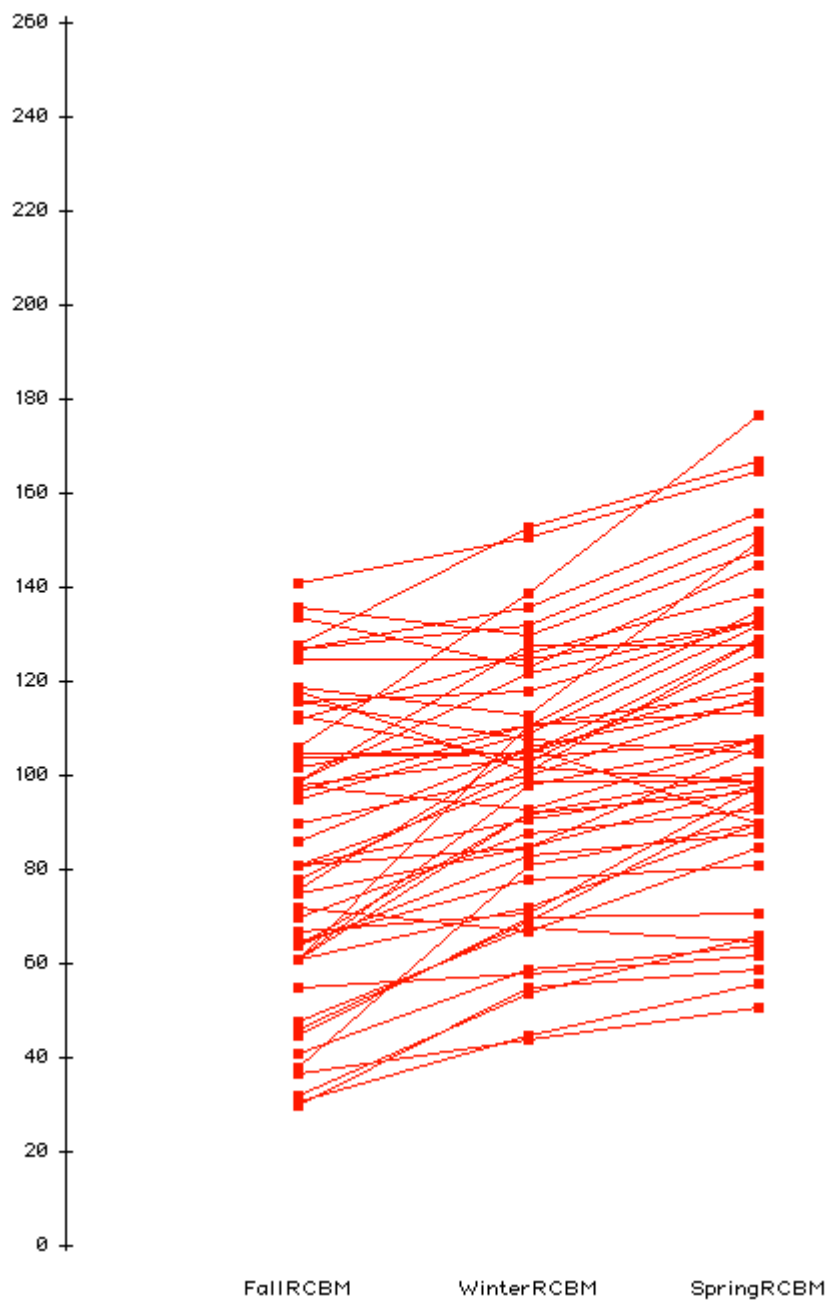


Figure 10. Parallel coordinate of ELL students' growth on R-CBM from Fall to Spring at AIMS level 2 (Approaches the standard).

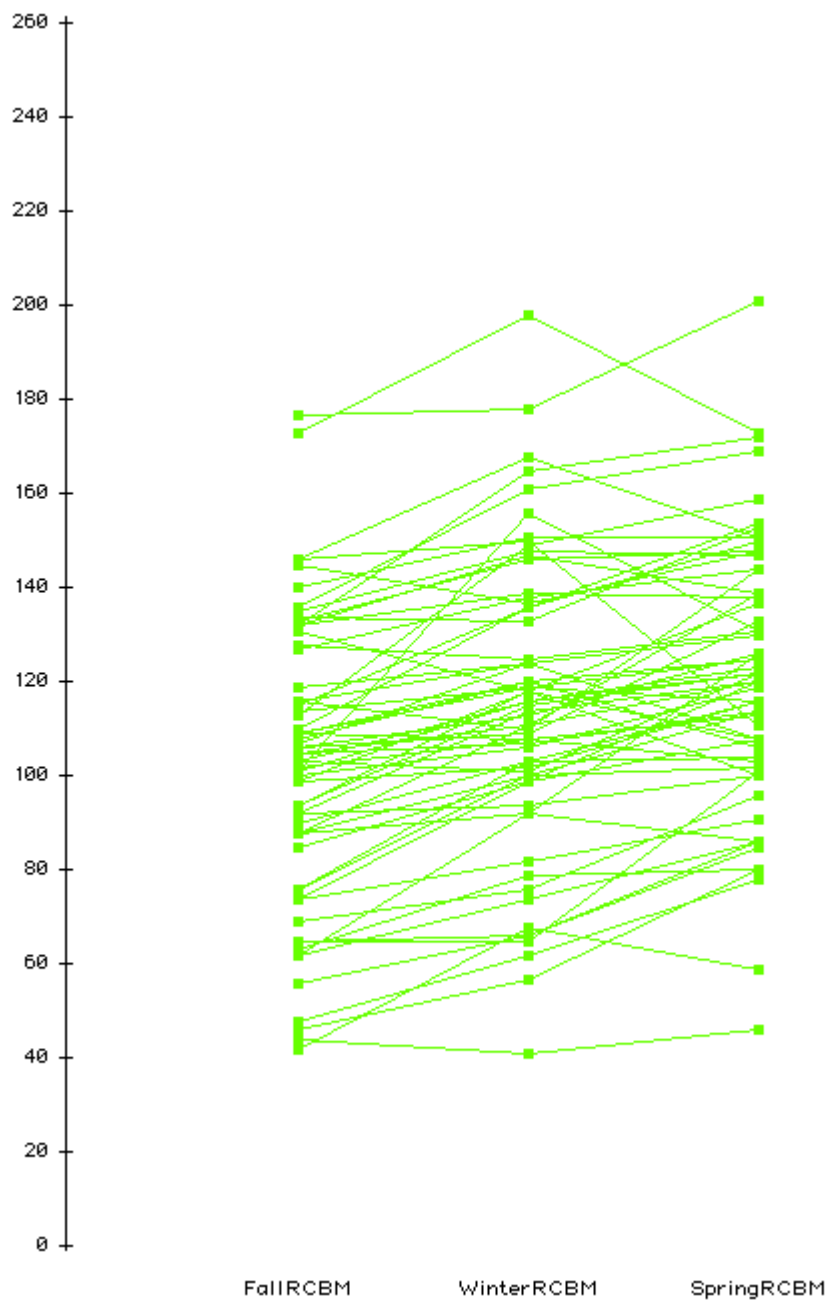


Figure 11. Parallel coordinate of non-ELL students' growth on R-CBM from Fall to Spring at AIMS Level 2 (Approaches the standard).

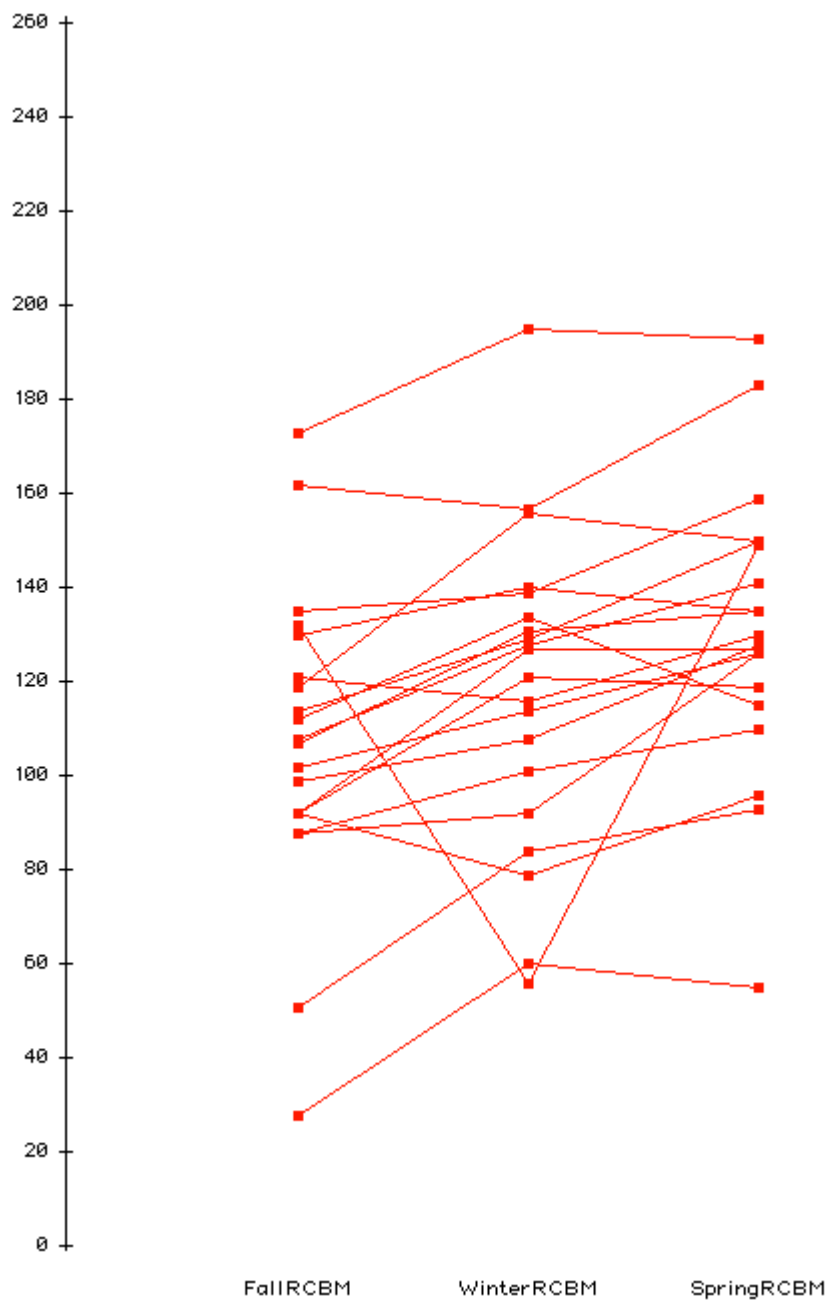


Figure 12. Parallel coordinate of ELL students' growth on R-CBM from Fall to Spring at AIMS level 3 (Meets the standard).

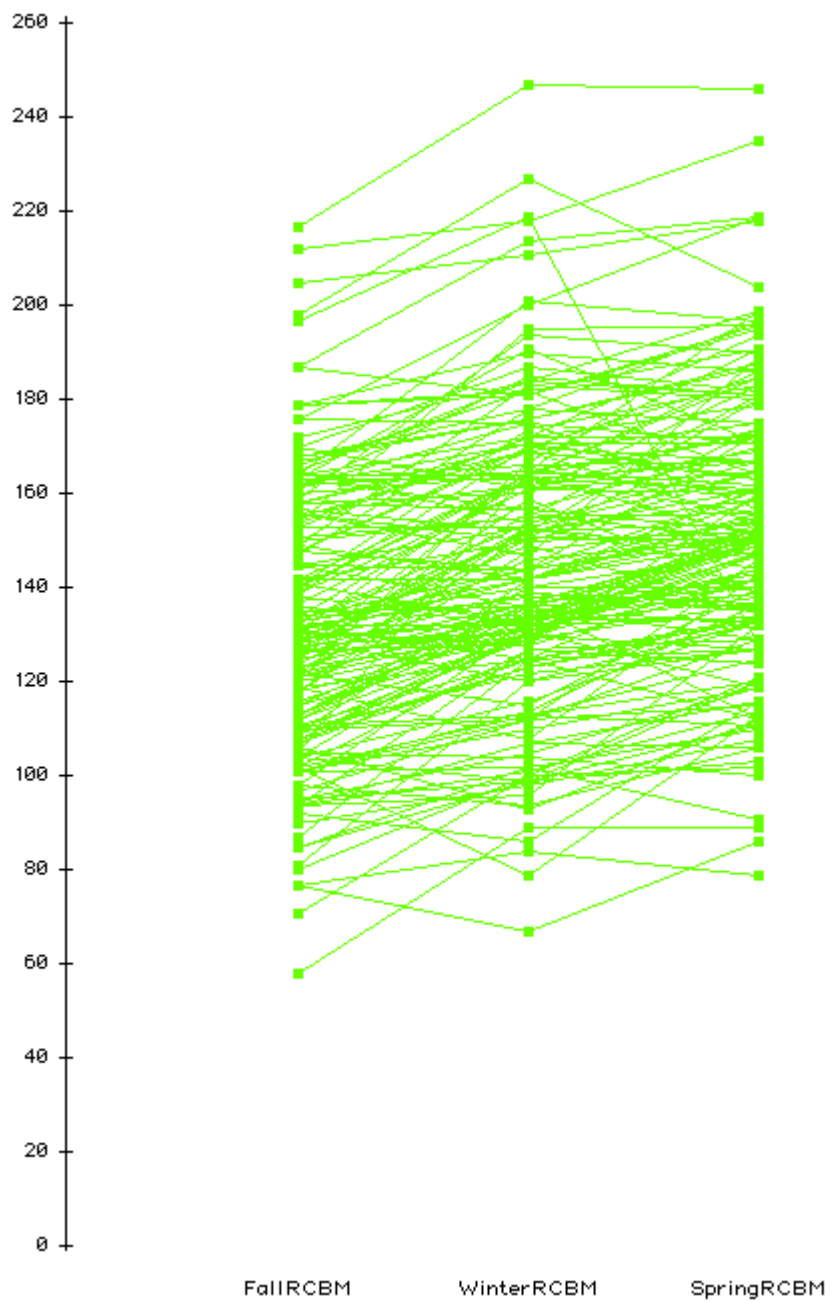


Figure 13. Parallel coordinate of non-ELL students' growth on R-CBM from Fall to Spring at AIMS level 3 (Meets the standard).

Prediction Analyses

As a first step in examining the predictor and outcome variables in this study, correlations were calculated between Fall R-CBM scores (initial status), student's growth on R-CBM from Fall to Spring, MAZE scores, and AIMS reading scores (see Table 20). Each pair of correlations was calculated using Pearson correlation. Results showed that the Fall R-CBM score was significantly correlated with student's growth from Fall to Spring on R-CBM, and MAZE, and the AIMS reading scores. Student growth on R-CBM from Fall to Spring was only significantly correlated to the Fall R-CBM score. The MAZE assessment was significantly correlated with the Fall R-CBM score and the AIMS reading score. The AIMS reading score was significantly correlated with the Fall R-CBM score and the MAZE score.

Table 20 *Correlation Matrix of Predictor and Outcome Variables*

Variable	FallRCBM	Growth	MAZE	AIMS
FallRCBM	--	-.271*	.599*	.577*
Growth	-.271*	--	.004	.049
MAZE	.599*	.004	--	.409*
AIMS	.577*	.049	.409*	--

Note: * $p < .01$.

In order to address the final research question, multiple regression analyses were conducted using SAS (2009) in order to establish the efficacy of predicting from fluency measures, R-CBM, to performance on a standardized achievement test, AIMS.

Specifically whether the R-CBM initial status (Fall score) and R-CBM rate of growth predict performance on the outcome measure of achievement (AIMS). In addition, ELL status and Maze assessment were also examined in terms of their predictive ability with AIMS performance. Additional regression analyses were conducted to examine the predictive ability of R-CBM initial status, R-CBM growth, and Maze relative to the student's ELL status. In using achievement as measured by the AIMS as the outcome (students performance on AIMS), multiple regression analyses were used to examine that outcome as a function of the student's Fall CBM score (initial status), the student's rate of growth on R-CBM from Fall to Spring, the student's ELL status (ELL vs. non-ELL) and the student's Maze score. The second multiple regression analysis was run comparing ELL versus non-ELL students AIMS outcome as a function of initial status, rate of growth, and Maze. All of the data assumptions for regression (residuals, VIF) were met and are displayed in Figures 14, 15, and 16. The residuals are random and normally distributed. Figure 14 shows the homoscedasticity and independence of residuals, which indicates that there are equal amount of residuals above and below zero and the scattering is random. Thus, there is no systematic bias. Figure 15 shows normality of residuals. The histogram in Figure 15 shows the actual residuals and the superimposed bell curve shows the ideal. They are very close to each other. Figure 16 also shows the normality of residuals through a scatterplot displaying the normal quantiles. If the residuals are normally distributed, they form a 45 degree diagonal line. Moreover, the Variance Inflation Factors (VIF) are low (below 10) as shown in Table 22 which indicates that multi-collinearity is not a serious problem in the regression model.

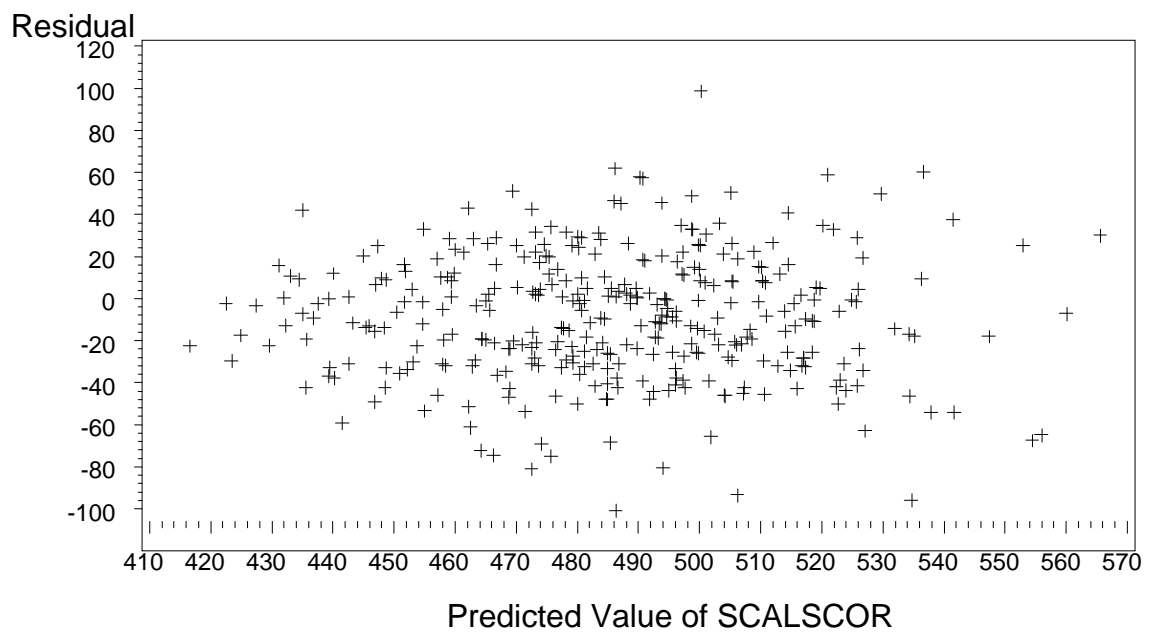


Figure 14. Homoscedasticity and independence of residuals.

Percent

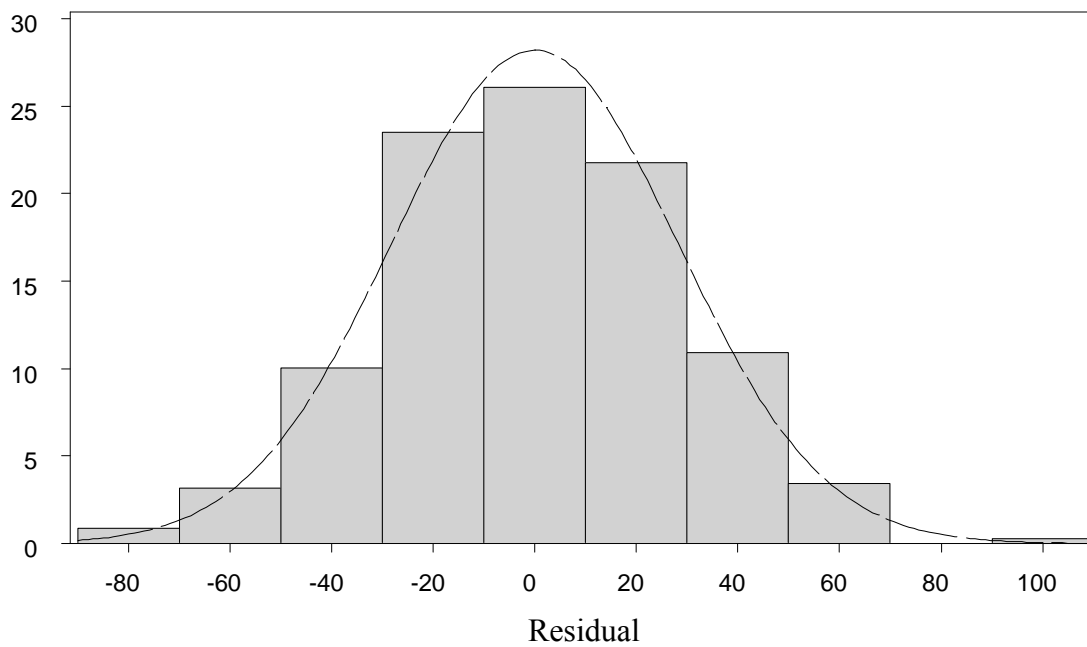


Figure 15. Normality of residuals.

Residuals

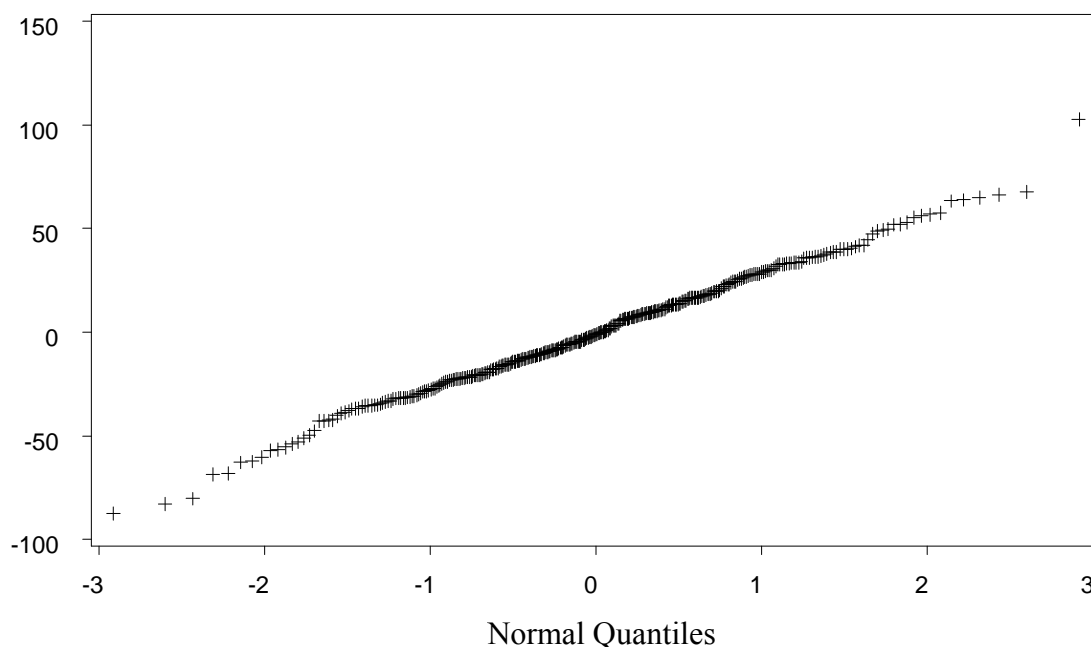


Figure 16. Scatterplot of the normality of residuals.

The results of the first multiple regression analysis are displayed in Tables 21 and 22. The overall model accounted for 48% of the variance on AIMS. The regression analysis showed that Fall RCBM (initial status), percentage of growth on RCB from Fall to Spring, and ELL status are all significant predictors of student performance on AIMS. The Maze assessment was not a significant predictor. In order to examine the term importance of each predictor on the AIMS outcome, TIBCO Insightful Miner (2009) was used. Results showed that Fall RCBM is the most important predictor of AIMS performance, with ELL status ranking second, percentage of growth third, and Maze as the least important predictor. Figures 17 through 20 show the scatter plots associated with the regression models. Figure 17 shows Fall RCBM with AIMS, and there appears

to be a consistent pattern. In other words, as students' Fall RCBM scores increase, so does their score on the AIMS assessment. Figure 18 shows ELL status in relation to AIMS performance. As a group, non-ELL students AIMS scores are higher than ELL students AIMS scores. While percentage of growth on R-CBM from Fall to Spring emerged as a significant predictor, the pattern is not clear from the scatterplot (see Figure 19). There are some students who grew by a significant percent from Fall to Spring, yet did not score higher on AIMS. Figure 20 shows the relationship of Maze and AIMS performance. Maze did not significantly predict performance on AIMS, and this is evident from the scatterplot. A consistent pattern does not emerge between the two variables.

Table 21 ANOVA: *Model and AIMS Reading Scores*

R Square		Sum of Squares	Mean Square	F	Sig.
.4828	Regression	260173	65043	80.27	.000
	Residual	2787738	810.28550		
	Total	538912			

Table 22 *Parameter Estimates of Predictor Variables on AIMS Outcome*

Variable	Label	DF	Parameter Estimate	Standard Error	<i>t</i> Value	Pr> <i>t</i>	VIF
Intercept	Intercept	1	400.10419	7.80022	51.29	.0001	0
FallRCBM	FallRCBM	1	0.67253	0.06473	10.39	.0001	2.6969
ELL		1	-15.36157	3.89130	-3.95	.0001	1.2482
Growth		1	0.27643	0.08224	3.36	.0009	1.7225
MAZE	MAZE	1	0.36818	0.33915	1.09	.2770	1.6872

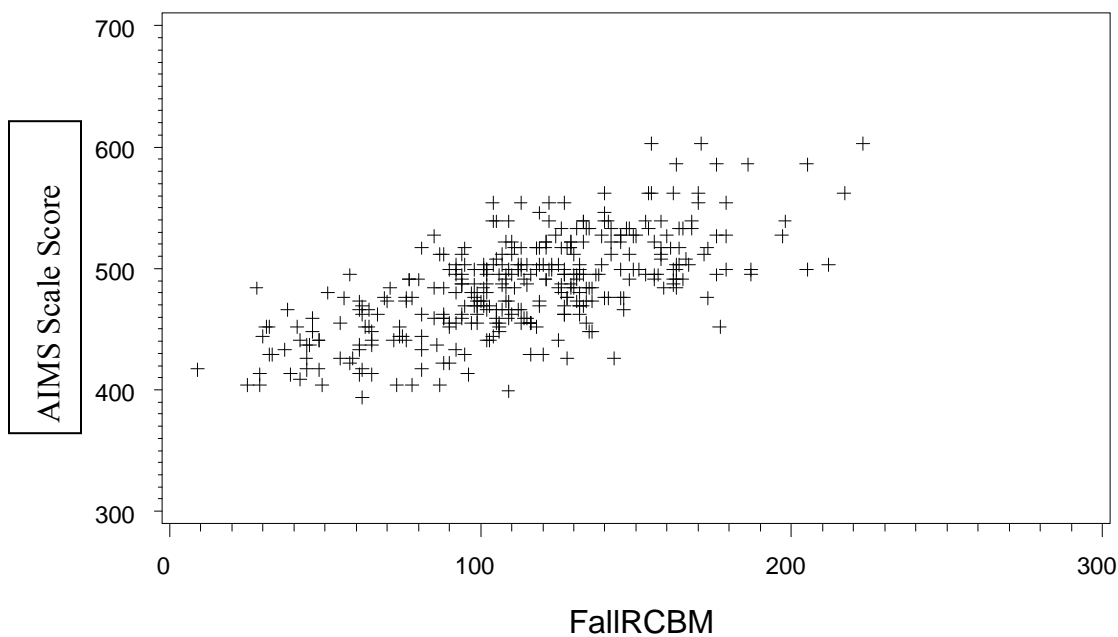


Figure 17. Scatterplot of Fall R-CBM by AIMS Scaled Score.

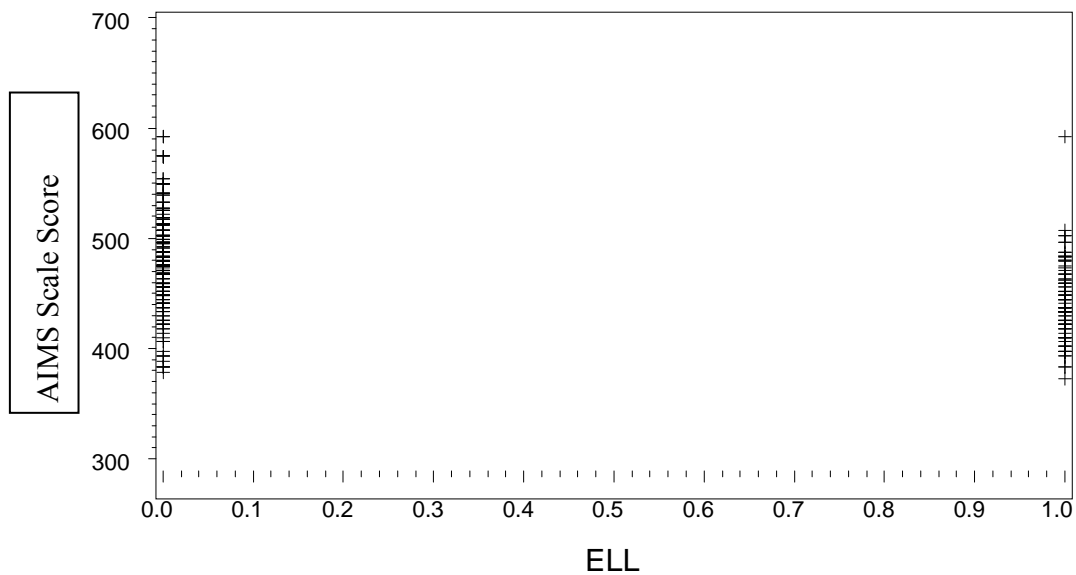


Figure 18. Scatterplot of ELL status by AIMS Scaled Score (ELL = 1, non-ELL = 0).

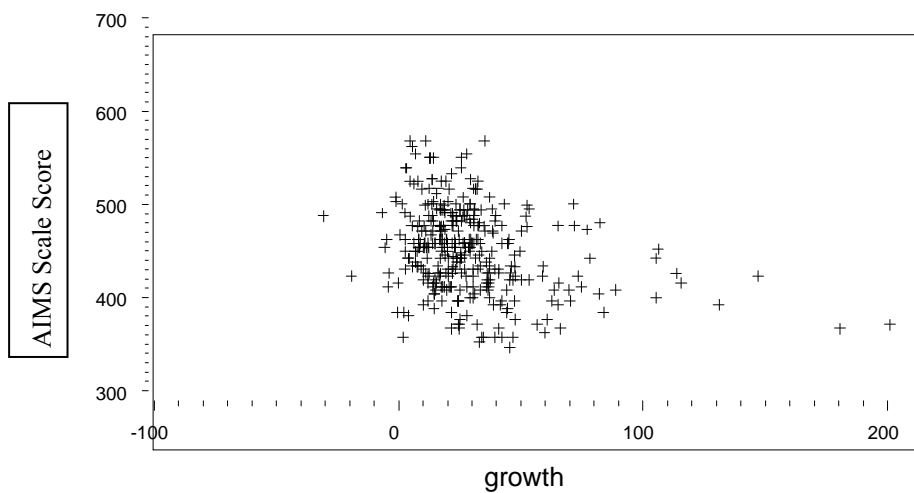


Figure 19. Scatterplot of Percentage of Growth on R-CBM from Fall to Spring and AIMS Scaled Score.

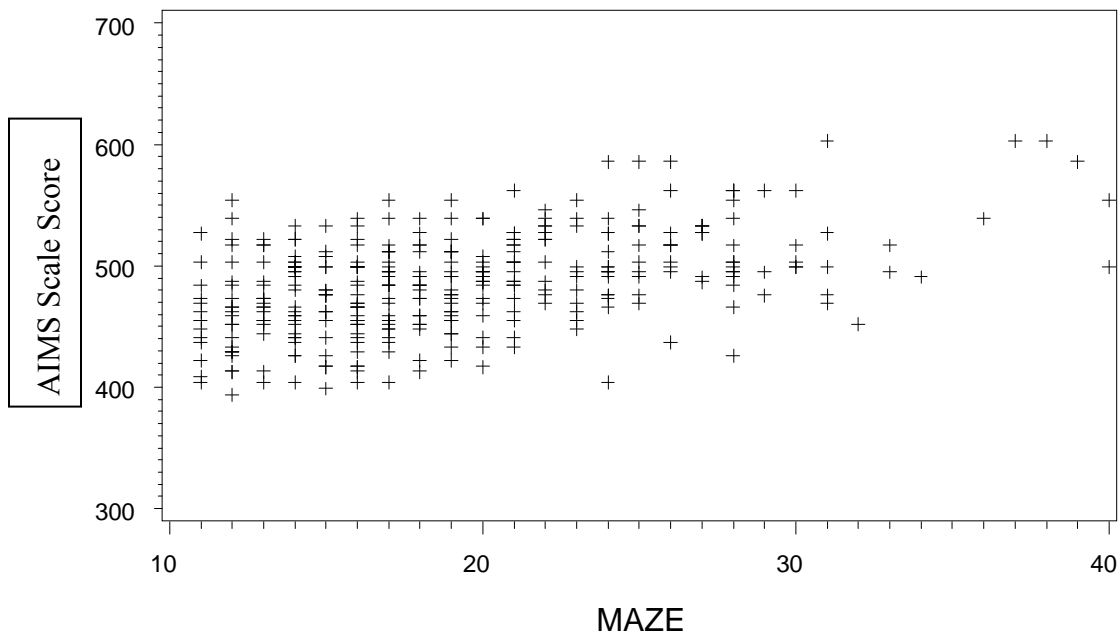


Figure 20. Scatterplot of Maze scores and AIMS Scaled Score.

The results of the second multiple regression analysis are displayed in Tables 23 and 24 (non-ELL) and 25 and 26 (ELL). The overall model for non-ELL students accounted for 38% of the variance on AIMS. The regression analysis showed that Fall RCBM (initial status) and percentage of growth on RCB from Fall to Spring are both significant predictors of non-ELL student performance on AIMS. The Maze assessment was not a significant predictor. The overall model for ELL students accounted for 37% of the variance on AIMS. The regression analysis showed that only Fall R-CBM (initial status) was a significant predictor of student performance on AIMS. Percentage of growth from Fall to Spring and Maze were not significant predictors of ELL students' performance on AIMS reading.

Table 23 ANOVA: *Non-ELL Model and AIMS Reading Scores*

R Square		Sum of Squares	Mean Square	F	Sig.
.3839	Regression	136181	45394	52.96	.0001
	Residual	218581	857.17873		
	Total	354762			

Table 24 *Parameter Estimates of Predictor Variables on AIMS Outcome for Non-ELL Students*

Variable	Label	DF	Parameter Estimate	Standard Error	tValue	Pr> t
Intercept	Intercept	1	394.09167	9.562	41.21	.0001
FallRCBM	FallRCBM	1	.72845	.08081	9.01	.0001
Growth		1	.44260	.14086	3.14	.0019
MAZE	MAZE	1	.17233	.39406	.44	.6622

Table 25 ANOVA: *ELL Model and AIMS Reading Scores*

R Square		Sum of Squares	Mean Square	F	Sig.
.3689	Regression	33881	11294	16.75	.0001
	Residual	57973	674.10567		
	Total	91854			

Table 26 *Parameter Estimates of Predictor Variables on AIMS Outcome for ELL Students*

Variable	Label	DF	Parameter Estimate	Standard Error	<i>t</i> Value	Pr> <i>t</i>
Intercept	Intercept	1	385.81217	14.37047	26.85	.0001
FallRCBM	FallRCBM	1	.60070	.11162	5.38	.0001
Growth		1	.17877	.10167	1.76	.0822
MAZE	MAZE	1	.89499	.72011	1.24	.2173

Figures 21 through 26 show the scatter plots associated with the ELL and non-ELL regression models. Figure 21 shows Fall RCBM with AIMS for non-ELL students, and there appears to be a consistent pattern. In other words, as non-ELL students' Fall RCBM scores increase, so does their score on the AIMS assessment. While percentage of growth on R-CBM from Fall to Spring emerged as a significant predictor for non-ELL students, the pattern is not clear from the scatterplot (see Figure 22). There are some students who grew by a significant percent from Fall to Spring, yet did not score higher on AIMS. There are also a few students who did not grow from Fall to Spring, but still scored higher on AIMS. Figure 23 shows the relationship of Maze and AIMS performance for non-ELL students. Maze did not significantly predict performance on AIMS, and this is evident from the scatterplot. A consistent pattern does not emerge between the two variables. Figure 24 shows Fall RCBM with AIMS for ELL students, and there appears to be a consistent pattern. In other words, as ELL students' Fall RCBM scores increase, so does their score on the AIMS assessment. Percentage of growth on R-

CBM from Fall to Spring did not emerge as a significant predictor for non-ELL students, and this is evident from the scatterplot (see Figure 25). There does not appear to be any relationship between percentage of growth on R-CBM and the AIMS scores. Figure 26 shows the relationship of Maze and AIMS performance for ELL students. Maze did not significantly predict performance on AIMS for ELL students, and this is evident from the scatterplot. A consistent pattern does not emerge between the two variables.

Scatterplot

ELL=0

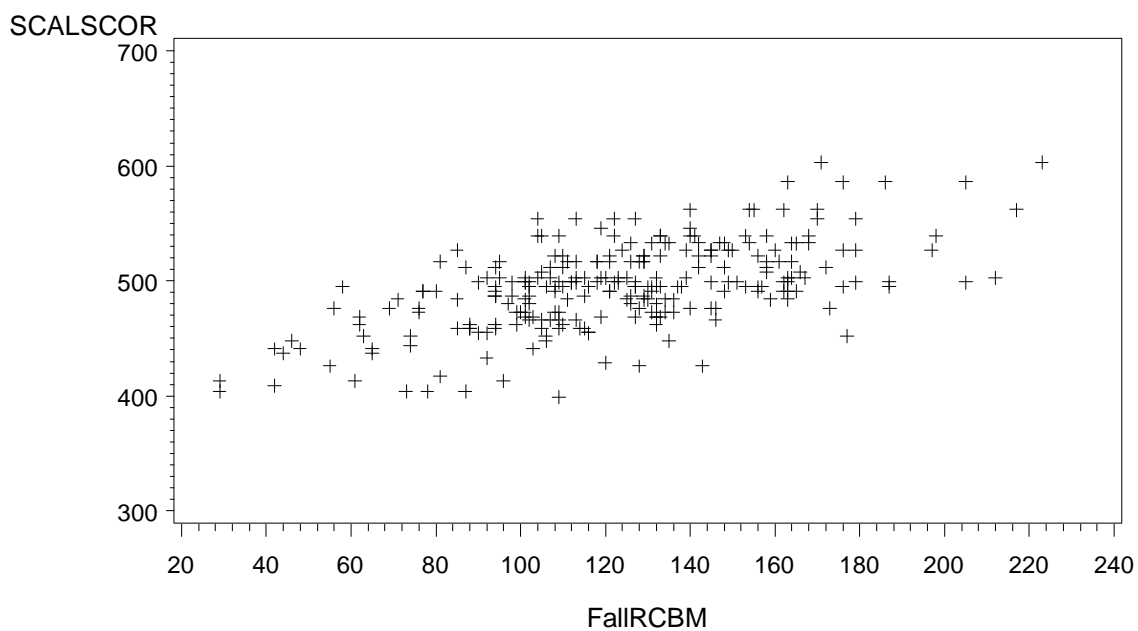


Figure 21. Scatterplot of R-CBM by AIMS Scaled Score for non-ELL students.

Scatterplot

ELL=0

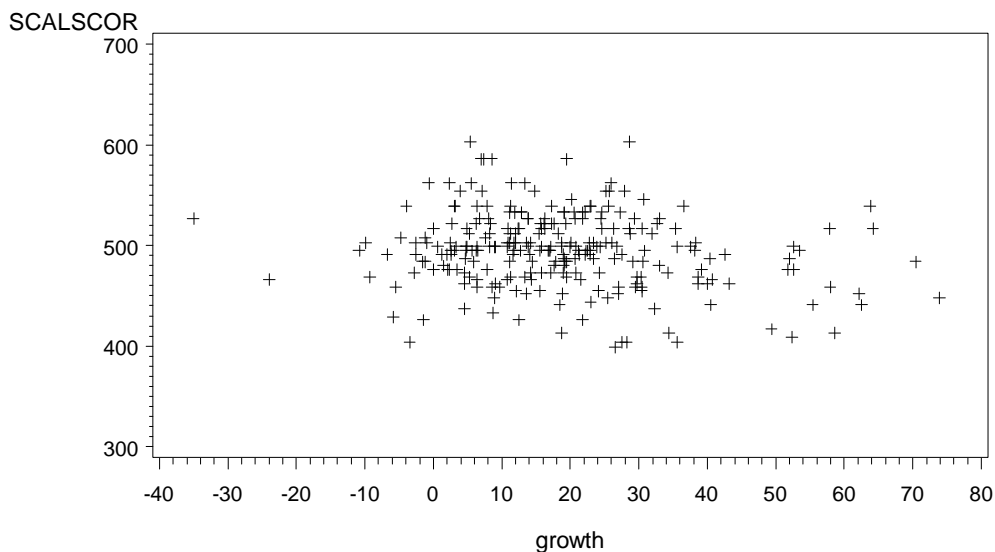


Figure 22. Scatterplot of R-CBM percentage of growth by AIMS Scaled Score for non-ELL students.

Scatterplot

ELL=0

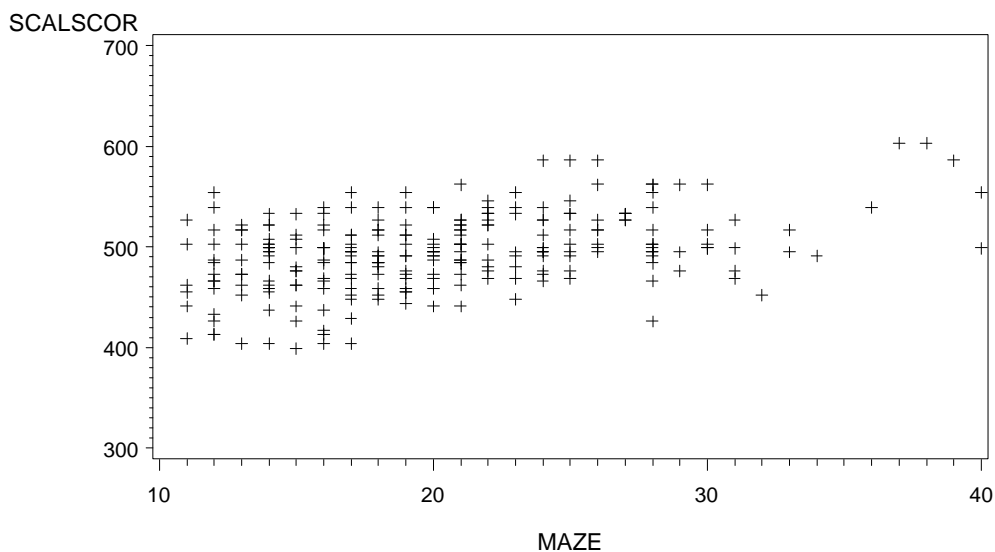


Figure 23. Scatterplot of Maze by AIMS Scaled Score for non-ELL students.

Scatterplot

ELL=1

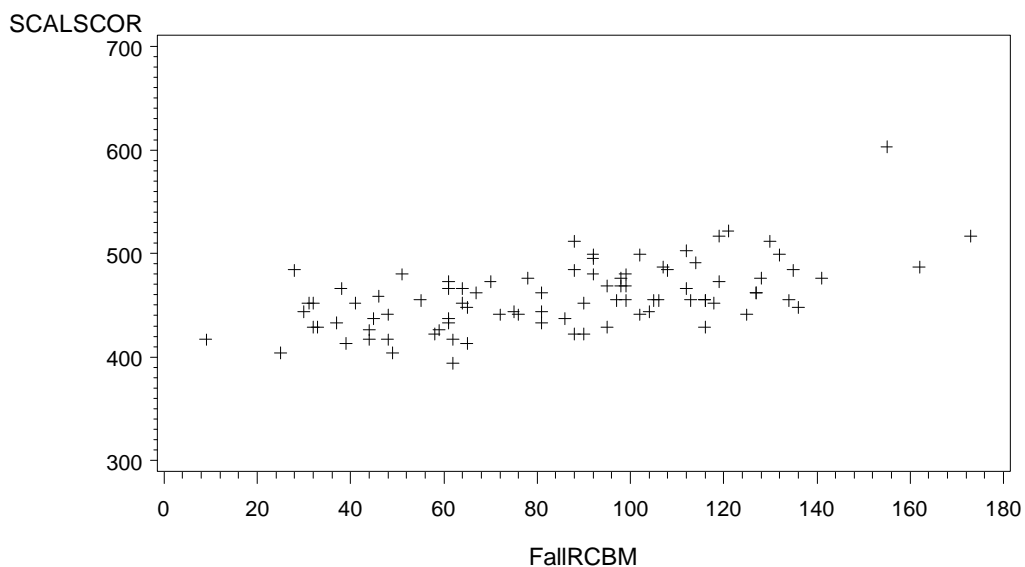


Figure 24. Scatterplot of Fall R-CBM by AIMS Scaled Score for ELL students.

Scatterplot

ELL=1

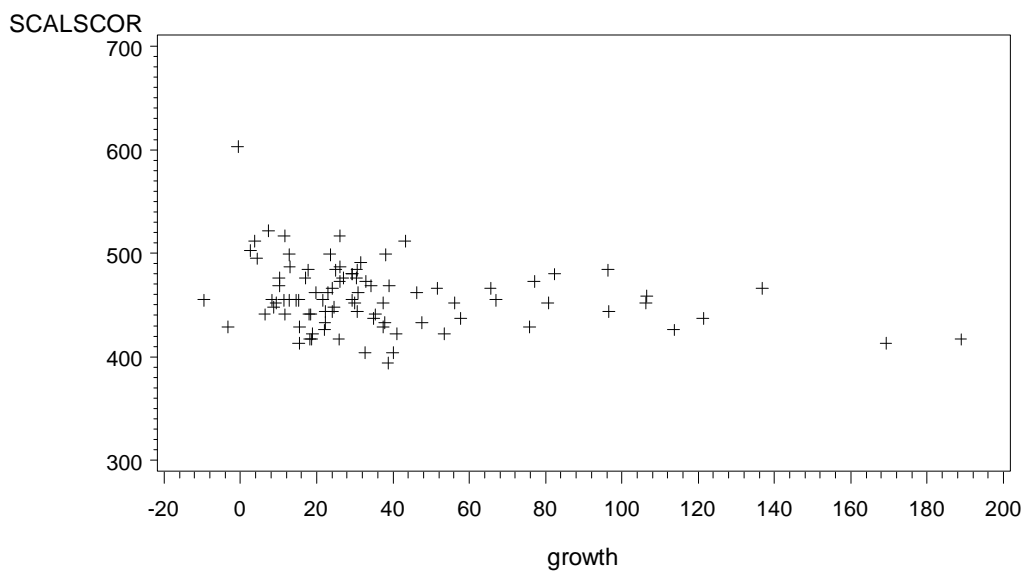


Figure 25. Scatterplot of R-CBM growth by AIMS Scaled Score for ELL students.

Scatterplot

ELL=1

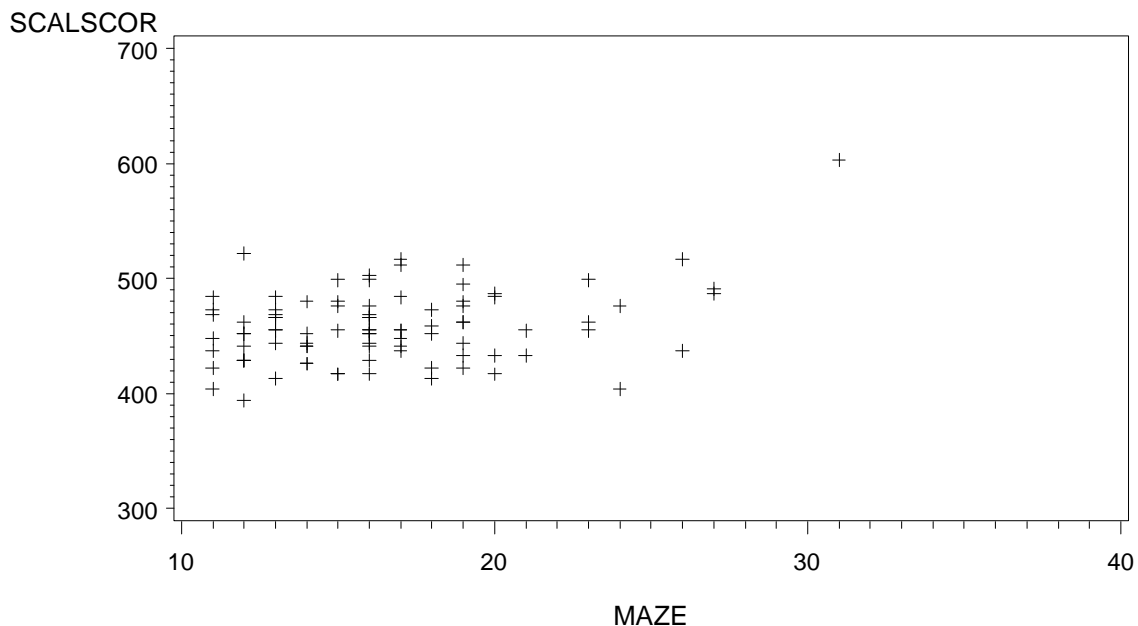


Figure 26. Scatterplot of Maze by AIMS Scaled Score for ELL students.

To further examine the relationships among the predictors and AIMS outcome, several trellis plots were generated (see Figures 27 through 30) using SAS JMP (2009). In examining the trellis plots, it appears that regardless of the conditioning variables (Maze, ELL), the scattergram of AIMS and growth of RCBM does not form a clear pattern. A different story emerges when the relationship between AIMS and the initial status of RCBM is examined. The slope is positive across all levels of different variables (ELL, Maze, etc). This implies that the growth pattern on RCBM has a wide variation, as indicated in the parallel coordinates shown previously. In other words, students who start at a low initial status on Fall RCBM might improve substantively, but students who have a high initial status might have minimal improvement, while some remain flat from Fall

to Spring. This patternless condition of growth is consistent across almost all levels of the conditioning variables (Maze, ELL, etc).

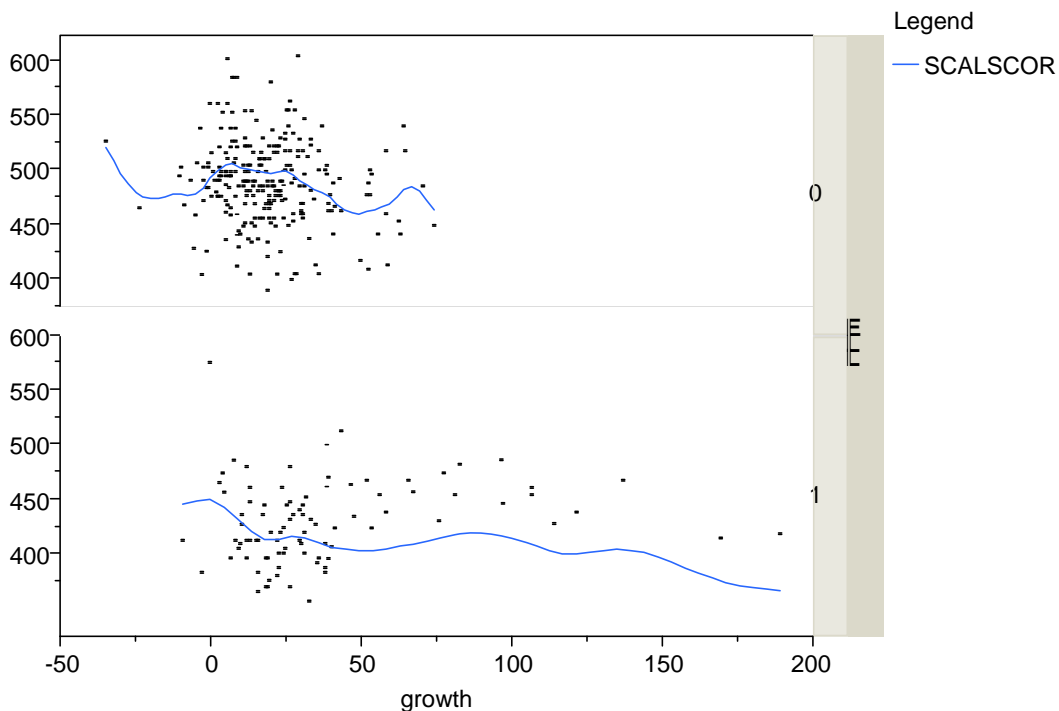


Figure 27. AIMS Scaled Score versus growth by ELL status.

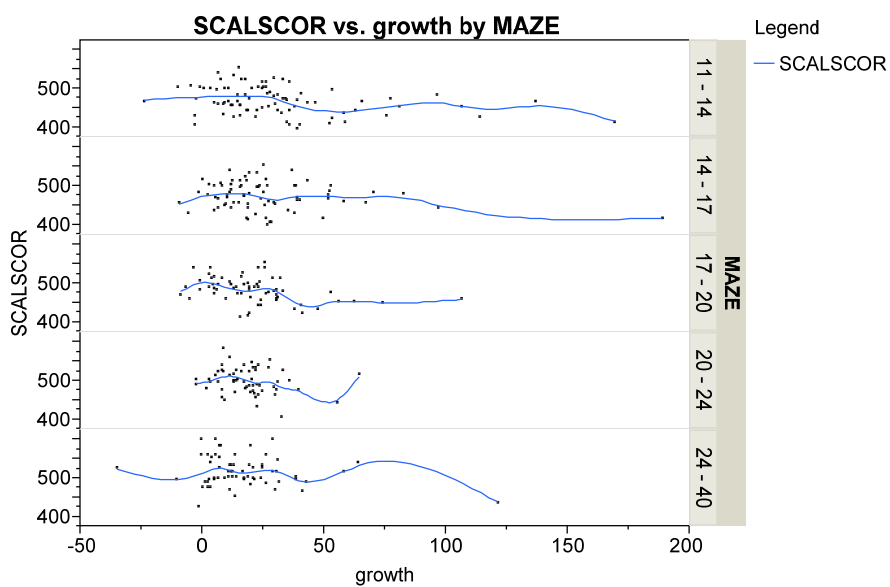


Figure 28. AIMS Scaled Score versus growth by Maze.

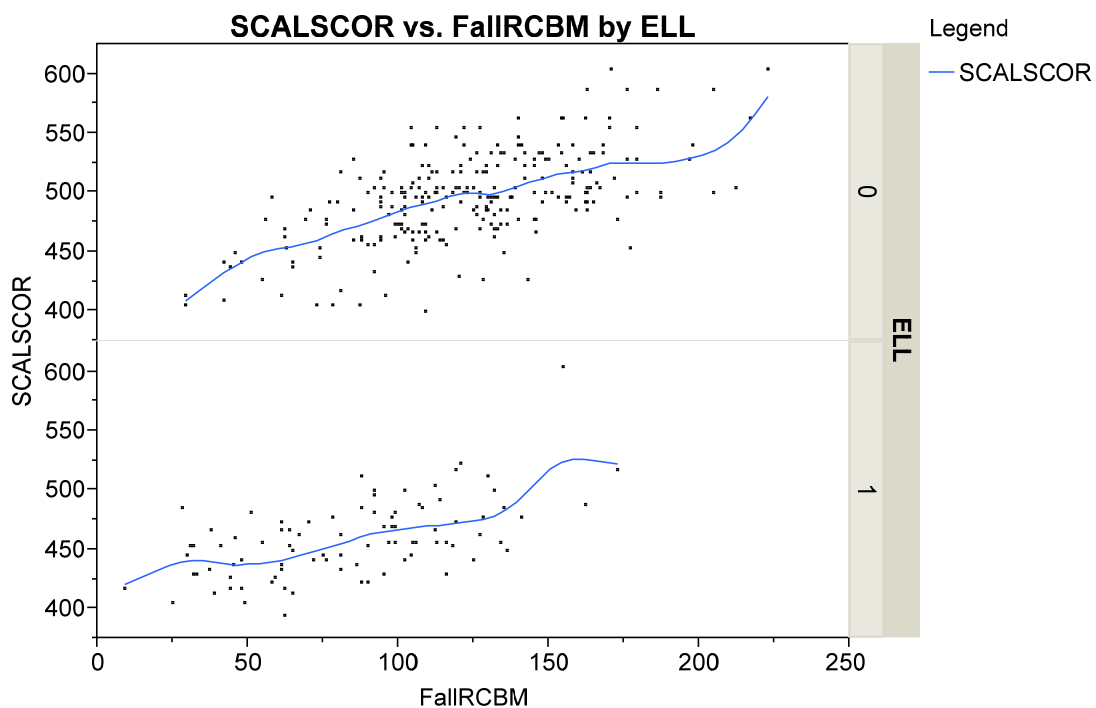


Figure 29. AIMS Scaled Score versus initial status (Fall R-CBM) by ELL status.

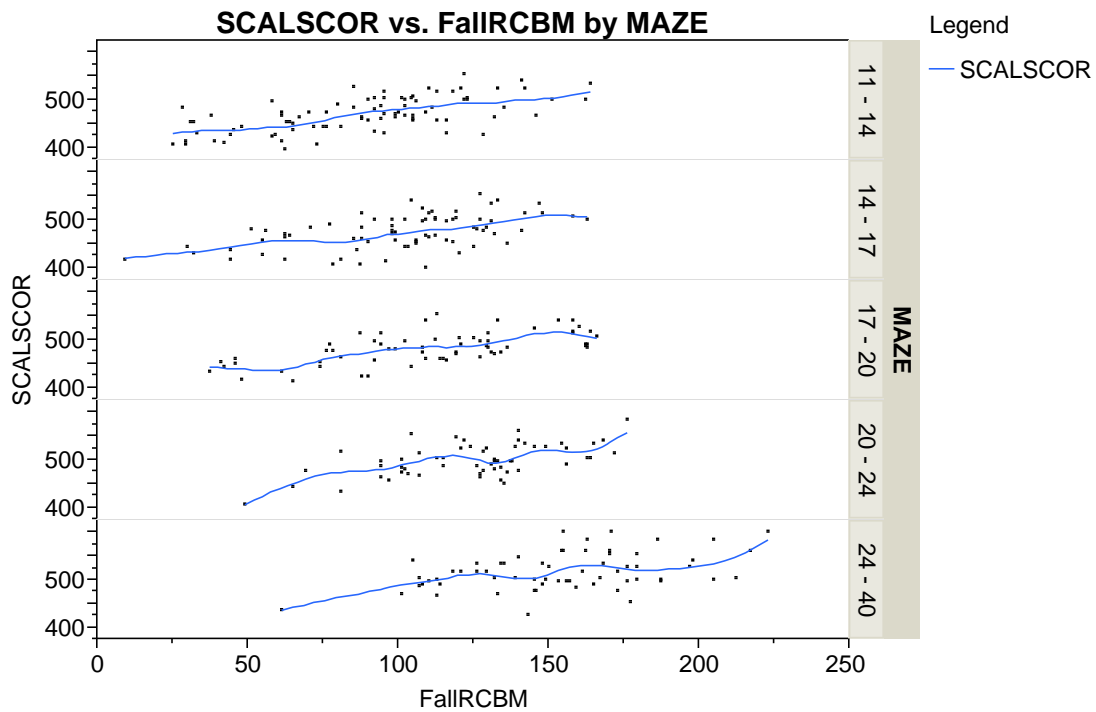


Figure 30. AIMS Scaled Score versus initial status (Fall R-CBM) by Maze.

CHAPTER V

DISCUSSION

This chapter discusses the findings and implications for practice of the current study in the context of past research. It also examines the limitations of the present study and potential areas of future research. This study addressed the initial status and rate of growth on reading curriculum-based measures with a population of sixth grade ELL and non-ELL students. The initial status (Fall R-CBM) and average rate of growth over one school year on R-CBM was examined relative to ELL status. Students' rate of growth was also examined in relation to their level of performance in reading on a high-stakes measure, Arizona Instrument to Measure Standards (AIMS). In addition, the efficacy and predictive validity of R-CBM fluency measures (initial status, rate of growth), as well as a Maze assessment and ELL status, was examined in relation to AIMS. In order to look for differences between ELL and non-ELL students in terms of the CBM predictors, further multiple regression analyses were conducted. These analyses looked at the predictive ability of the three measures (R-CBM initial status, R-CBM growth, and Maze) on student performance on AIMS relative to their ELL status. The current nonexperimental study utilized archival data from a group of 349 students in sixth grade from a school district on the west side of Phoenix, Arizona. Data was analyzed to determine if there was significant variation in students' initial status as measured by the Fall R-CBM score as well as growth rate on R-CBM from fall to spring relative to their

AIMS performance level and ELL status. Further analysis examined the Fall R-CBM measure, rate of growth from fall to spring on R-CBM, Maze, and ELL status to determine if they were predictive of students' performance on the AIMS reading assessment administered in the spring.

R-CBM Initial Status and Growth: Results

Initial status on R-CBM in the fall and growth curves of the overall population of students was examined in terms of an initial status (Fall R-CBM score) and an average growth rate from Fall to Spring on R-CBM. The results answer the first research question: What was the average rate of growth of the current population over one school year on the reading fluency R-CBM measures? The HLM analysis indicated that the average initial status as measured by the Fall R-CBM was 103.51 words read correctly per minute with an average growth rate (slope) of 10.36 words read correct from Fall to Winter, and from Winter to Spring. Thus, the overall growth was about 20.72 words over one school year. Further analysis established that the students' scores on the Fall R-CBM as well as the growth rates of all students from fall to spring were significantly different. The students, therefore, started at significantly different points in the fall on the R-CBM measure, and grew at significantly different rates from fall to spring.

R-CBM Initial Status and Growth Relative to ELL Status: Results

The results of the second HLM model were examined to determine if in fact students start at different points on the fall R-CBM measure and grow at different rates on R-CBM from over one school year relative to their ELL status. This section responds to the second research question: what is the relationship among the rate of growth on R-

CBM and level of English Language Proficiency? The results indicated that there is significant variation between ELL and non-ELL students in their initial status as measured by the Fall R-CBM score. Specifically, the mean R-CBM score for non-ELL students in the fall (initial status) was 114.01 words read correctly per minute, which was higher than that for ELL students by 40.68 words read correct on R-CBM. Thus, ELL students are starting the school year at a much lower status when compared to their non-ELL peers in terms of reading. In order to examine whether there were differences in the rate of growth for ELL versus non-ELL students, further analysis examined the average growth rate (slope) of non-ELL and ELL students from fall to spring on R-CBM. Results indicated the average growth rate for non-ELL students was 19.74 words read correct from Fall to Spring. ELL students' growth rate from Fall to Spring on R-CBM, on average, was 1.98 correct words more than non-ELL students. The mean difference in growth rate between the two groups (ELL versus non-ELL) was not significant. Thus, while ELL students score significantly lower in the Fall on R-CBM, they do not vary in their growth trajectories from non-ELL students. This is concerning for the ELL students, as the current study indicates that they continue to fall behind their non-ELL peers in terms of reading performance.

AIMS Performance and R-CBM Initial Status and Growth Rates: Results

The AIMS models examined the relationship among the four levels of AIMS reading and initial status and growth rates on R-CBM, as well as ELL status. The HLM analyses were conducted to answer the third research question: If students vary on their initial status (Fall R-CBM), growth rates (slopes) on R-CBM and/or performance on state

AIMS reading assessment, is this variation related to ELL status? Students fall in to one of four performance categories for the AIMS assessment. Students at AIMS level 1 fall far below the standard, students at AIMS level 2 approach the standard, students at AIMS level 3 meet the standard, and students at AIMS level 4 exceed the standard. Results showed that overall, students varied significantly based on their AIMS level in their initial status as measured by the fall R-CBM score. Further analysis revealed that students at AIMS Level 1 varied significantly from students at Level 4 in their initial status, while students at Level 2 and Level 3 did not vary significantly. Students do not vary significantly on their growth rates relative to their performance level on AIMS.

In order to further examine the complex relationships among the variables, data visualization was utilized. While there were no statistically significant differences found among rate of growth on R-CBM relative to the level of performance on AIMS, by examining the parallel coordinates, interesting patterns emerged. It appeared that students at the lowest level of performance on AIMS (Level 1 Falls Far Below) and students at the highest level of performance on AIMS (Level 4 Exceeds) exhibited slower rates of linear growth from Fall to Spring on R-CBM. Thus, for students who are exceeding standards on AIMS, their level of reading fluency may have hit a ceiling. For students who are falling far below standards on AIMS, their R-CBM scores in Fall are significantly lower and they do not appear to make much progress through the spring. The two middle levels on AIMS (Level 2 Approaches and Level 3 Meets) showed the most consistent linear growth rates from Fall to Spring on R-CBM. However, the growth

rates were not significantly different, and this needs to be kept in mind when examining and interpreting the data visualizations.

The interaction effects of AIMS, ELL Status, and R-CBM were examined. Results showed that both ELL status and AIMS level related significantly to initial status as measured by R-CBM. However, ELL status and AIMS level are not significantly related to rate of growth on R-CBM from Fall to Spring. The interaction effect between AIMS level, ELL status and rate of growth was not significant. To further examine the interaction effects, parallel coordinates were generated which displayed individual students' rates of progress on R-CBM from Fall to Spring by AIMS level relative to their ELL status. It was apparent from the data visualization that most ELL students at AIMS Level 1 showed slow growth from Fall to Spring, but there were a few students who made substantial growth from Fall to Winter but then dropped down from Winter to Spring. For non-ELL students at AIMS Level 1, their growth trajectories were flat, in that most students did not make much growth from Fall to Spring on R-CBM. Unlike the ELL students at Level 1, the non-ELL students did not appear to drop substantially from Winter to Spring. For both ELL and non-ELL students at AIMS Level 2 (approaches the standard), there seems to be a general upward trend from Fall to Spring on R-CBM. In other words, regardless of their ELL status, the students approaching the standard on AIMS showed growth from Fall to Winter, and from Winter to Spring on R-CBM. For ELL students at AIMS level 3 (meets the standard), a consistent pattern in growth from Fall to Spring does not emerge, which would not be expected given these students met the standard on the AIMS reading test. There were several students who made growth

from Fall to Winter, but then their performance leveled off from Winter to Spring. However, for non-ELL students at AIMS Level 3, a consistent pattern of growth emerged, where most students continued to grow consistently from Fall to Spring on R-CBM. This pattern would be expected given that these students met the standard on the AIMS state assessment in the spring. As with the AIMS only models, there were no statistically significant differences among the growth rates, so this limits the usefulness of the data visualizations.

Predictors for AIMS Performance: Results

Multiple regression analysis was conducted in order to address the final research question(s): Is R-CBM initial status, R-CBM rate of growth, and Maze predictive of reading achievement on the spring AIMS reading assessment relative to ELL status? Which of these factors is a more potent predictor of reading achievement on the AIMS reading assessment? The analyses were conducted in order to establish the efficacy of predicting from fluency measures, R-CBM, to performance on a standardized achievement test, AIMS. Specifically whether the R-CBM initial status (Fall score) and R-CBM percentage of growth predict performance on the outcome measure of achievement (AIMS). In addition, ELL status and Maze assessment were also examined in terms of their predictive ability with AIMS performance. Additional regression analyses were conducted to examine the predictive ability of R-CBM initial status, R-CBM growth, and Maze relative to the student's ELL status. In using achievement as measured by the AIMS as the outcome (students performance on AIMS), multiple regression analyses were used to examine that outcome as a function of the student's Fall

CBM score (initial status), the student's percentage of growth on R-CBM from Fall to Spring, the student's ELL status (ELL vs. non-ELL) and the student's Maze score. The second multiple regression analysis was run comparing ELL versus non-ELL students AIMS outcome as a function of initial status, percentage of growth, and Maze. The regression analysis for the overall student population showed that Fall RCBM (initial status), percentage of growth on RCB from Fall to Spring, and ELL status were all significant predictors of student performance on AIMS. The Maze assessment was not a significant predictor. Further analysis showed that Fall RCBM or initial status is the most important predictor of AIMS performance, with ELL status ranking second, percentage of growth third, and Maze as the least important predictor. Thus, it appears that a student's initial status, as measured by the Fall R-CBM measure, is the most important factor in predicting whether that student will meet the state standards on the AIMS assessment in the spring.

The regression analyses conducted that examined whether there were any differences in the predictors' significance for ELL and non-ELL students showed that for both groups, initial status (Fall R-CBM) was the most significant predictor of AIMS performance. However, while percentage of growth was a significant predictor for non-ELL students, it was not a significant predictor for ELL students. Thus, as non-ELL students' percentage for growth from Fall to Spring on R-CBM increased, so did their AIMS scaled scores in reading. This was not the case for ELL students. There was no consistent pattern between R-CBM percentage of growth and AIMS performance for this group. Maze was not a significant predictor for either group's performance on AIMS.

Thus, similar to the overall student population, a student's initial status as measured by R-CBM in the Fall is the most important predictor of how that student will perform on the high-stakes AIMS assessment in reading in the spring.

To further examine the relationships among the predictors and AIMS outcome, several trellis plots were generated. In examining the trellis plots, it appeared that regardless of the conditioning variables (Maze, ELL), the scattergram of AIMS and growth of RCBM did not form a clear pattern. In other words, regardless of a student's score on the Maze assessment or ELL status, percentage of growth on R-CBM is not a significant predictor of a student's AIMS performance. This implies that the growth pattern on RCBM has a wide variation, as indicated in the parallel coordinates shown previously. In other words, students who start at a low initial status on Fall RCBM might improve substantively, but students who have a high initial status might have minimal improvement, while some remain flat from Fall to Spring. This patternless condition of growth is consistent across almost all levels of the conditioning variables (Maze, ELL, etc). However, a different story emerges when the relationship between AIMS and the initial status of RCBM is examined. The slope was clearly positive across the two levels of different variables (i.e., ELL and Maze). Thus, initial status as measured by the Fall R-CBM, as supported by the regression analysis, seems to be the most potent predictor of a student's performance on AIMS reading, regardless of their ELL status or Maze performance.

It is important to note that while MAZE did not emerge as a significant predictor of AIMS performance for ELL and non-ELL students, it was significantly correlated with

the AIMS reading scores. Thus, MAZE should not be discounted as a potentially effective and valid measure of students' reading achievement. It is possible that since the MAZE was administered in the spring just before the AIMS in the current study, its predictive validity was decreased. Perhaps if the MAZE assessment was administered in the Fall at the same or similar time as the Fall R-CBM, its predictive power would have increased. Furthermore, the MAZE was only administered once during the school year. It would have been useful to examine student growth from Fall to Spring as measured by the MAZE in addition to the R-CBM.

Results in Relation to Previous Research

R-CBM: Initial Status and Rate of Growth

The use of CBM with elementary students has been extensively documented in the literature. This research has suggested that for the majority of elementary-aged students, the total number of words read correctly on a one-minute reading probe is a valid indicator of general reading proficiency (Shinn, 2009). Research has recently begun to also examine the use of CBM in measuring reading growth in middle and secondary school environments, in assessing and monitoring English Language Learners, and in predicting student performance on state accountability measures (Deno, 2003). The present study examined the initial status (Fall R-CBM score) and the typical rate of growth on R-CBM for sixth grade middle-school students, while also exploring the differences among ELL and non-ELL students in relation to R-CBM.

While screening for early reading difficulties has been well-established in the research with elementary-aged students, very little is known about valid and reliable

screening tools for use with older students (i.e., middle and secondary schools).

Currently, middle schools have limited technically adequate measurement systems to assist in both the identification and ongoing progress monitoring of students with reading difficulties. Curriculum-based measurement has the potential to become that technically adequate tool for evaluating the reading skills of older students, although the research is still emerging in this area. The current results showed that this population of students' initial status, or words read correct on the Fall R-CBM was about 103 words read per minute, with an average growth rate of about 10 words per minute from Fall to Winter and again from Winter to Spring, averaging about 20 total words growth over the school year. Both initial status and rate of growth were significant.

According to the Florida Center for Reading Research, a sixth grade Fall R-CBM score of 122 words per minute would be considered "low risk" (FCRR, 2006).

According to the Hasbrouck-Tindal (2005) table of oral reading fluency norms, students at the 50th percentile in sixth grade in the fall are reading 127 words correct within one minute. Based on this research, the current student population is scoring within the at-risk range and significant below the 50th percentile compared to the national average based on the Fall R-CBM scores. However, when the R-CBM Fall score was examined relative to the student's level of performance on the AIMS assessment, students who met the standard (Level 3) on AIMS had a mean Fall R-CBM score of 129.31, which is consistent with previous research. Thus, for those students in this population who were actually meeting the state standards in reading, their initial status as indicated by the Fall R-CBM score was higher than the overall population.

Based on the developmental reading literature, greater changes in oral reading fluency development occur more in the earlier grades (i.e., 2-3) as compared to the later grades (i.e., 4-6) (Adams, 1990 and Chall, 1983 as cited in Silberglitt & Hintze, 2007). As such, there are minimal studies which have examined the average rate of growth on R-CBM for middle school students, but one study by Silberglitt and Hintze showed an average rate of growth for a population of sixth grade students to be about 28 words read correct over a school year. This is slightly higher than the rate of growth for the students in the present study. However, in the Silberglitt and Hintze study, the population of students was different than the current population, with only .9% of students being considered ELL, and 93.8% of students classified as White, not of Hispanic origin. Regardless of the population of students, it appears that R-CBM is still sensitive enough to detect growth in reading over a school year at the sixth grade level.

R-CBM: ELL versus non-ELL Students

Based on the overrepresentation of ELLs in special education, as well as the rising achievement gap between the ELL and non ELL population, there is an obvious need for reliable and valid assessment tools for the purpose of early identification of ELL students who may be struggling in reading. However, such assessment tools are lacking for this population (Wagner, Francis, & Morris, 2005). Several studies have begun to explore the use of CBM in assessing ELL students and in monitoring their reading growth (Baker & Good, 1995; Betts et al., 2009; De Ramírez & Shapiro 2006; Gunn et al., 2005). The current study attempted to add to this research in terms of examining the usefulness of R-CBM with the ELL population.

Results of the present study indicated that ELL students' started at a significantly lower level than non-ELL students in terms of their Fall R-CBM scores. The average non-ELL student started the year reading approximately 114 words per minute, which was about 41 words higher, on average, than the ELL students. However, there were no significant differences detected among ELL and non-ELL students in terms of rate of growth over the school year on R-CBM. On average, ELL students grew 1.98 correct words more than non-ELL students. Thus, while ELL students score significantly lower in the Fall on R-CBM, they do not vary in their growth trajectories from non-ELL students. These results are consistent with a recent study by Betts et al. (2009) which examined oral fluency growth of ELL and non-ELL third graders. The study found that the oral reading fluency growth of ELLs in their study resembled that of nonELL students. However, it was noted that the ELLs actual oral reading fluency rate was substantially lower than what would be typically expected for third grade students, which is consistent with the results of this study. In other words, the current sixth grade ELL students R-CBM scores were also substantially lower than the "typical" or expected scores. Another study by De Ramirez and Shapiro (2006) examined the oral reading fluency growth rates among ELL and non-ELL students in grades 1-5. Results indicated that R-CBM can be used as a sensitive measure in assessing the rate of growth in reading for Spanish-speaking ELL students, although the rates of reading growth on English R-CBM measures were significantly slower for ELL students than for non-ELLs which is not consistent with the current results. However, the authors point out that rate of improvement, or slope, across grade levels substantially improved by fifth grade. Thus, it

is possible that as ELL students get older their rates of reading growth may increase and “catch up” to their non-ELL peers. However, because they are starting at such a lower reading rate, regardless if they are growing at the same pace as their non-ELL peers, they will still be ending the year significantly below grade level expectations.

CBM as a Predictor on High-Stakes Assessment

Studies have begun to examine the usefulness of curriculum-based measures in predicting future reading success in terms of state accountability outcomes. By screening students in the fall using R-CBM, research has shown that student performance on the early screening measures can predict whether or not the students will meet state standards on the high-stakes assessments in the spring (Crawford, Tindal, & Steiber, 2001; Hintze & Silberglitt, 2005; McGlinchey & Hixson, 2004; Silberglitt, Burns, Madyun, & Lail, 2006; Stage & Jacobsen, 2001; Wood, 2006). The current study utilized multiple regression analysis in order to examine the predictive ability of several variables (i.e., initial status as measured by the Fall R-CBM score, percentage of growth from Fall to Spring on R-CBM, ELL status, and Maze) on the students’ reading achievement in the spring as measured by a high-stakes assessment, AIMS. Results showed, for the overall student population, that initial status (Fall R-CBM score) was the most important predictor of a students’ AIMS performance in the spring followed by ELL status and percentage of growth over the school year on R-CBM. The Maze assessment was not a significant predictor of student achievement on AIMS. The current results are consistent with previous research by Stage and Jacobsen (2001) in which R-CBM was examined in relation to their predictive ability on a high-stakes reading assessment (WASL). Their

results indicated that the Fall R-CBM cut-scores increased the predictive power of whether students would pass or fail the WASL by 30% over base rate levels. Consistent with the present study, their findings also showed that R-CBM scores in the Fall were a better predictor of the state test performance than growth in R-CBM throughout the school year. While the Stage and Jacobsen study used a younger population of students (fourth graders) than the current study, it appears that Fall R-CBM continues to be the most significant predictor of student achievement as measured by high-stakes assessments as students get older. One explanation for the decreased importance of student growth, or slope, on predicting student achievement in later grades can be based on previous research on R-CBM growth in reading (Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993). Research has shown that although oral reading fluency increases in later grades, it does not increase as much as in earlier grades, which helps to explain why the increase in slope did not account for as much variance in the outcome measure in both the current study (AIMS) and the Stage and Jacobsen study (2001).

The present study adds to the existing research which has shown R-CBM to be a significant predictor of student performance on state accountability tests. However, very few studies have looked at the use of CBM with English Language learners, particularly whether they are useful in predicting future reading success on high stakes assessments. The current study examined the predictive ability of R-CBM initial status, R-CBM growth from Fall to Spring, and the Maze assessment relative to a student's ELL status. Regression analyses revealed that for both ELL and non-ELL students, the Fall R-CBM score was the most significant predictor of the student's performance on the AIMS

reading assessment in the spring. This is consistent with a study by Wiley and Deno (2005) which examined both oral reading fluency and Maze measures with ELL and non-ELL students in terms of their ability to predict student performance on the *Minnesota Comprehensive Assessment* (MCA). The sample consisted of 36 third-grade students (21 non-ELL and 15 ELL), and 33 fifth-grade students (19 non-ELL and 14 ELL). The students were administered measures of oral reading fluency, maze and the state assessment (MCA). Results showed moderately strong correlations between ORF and the MCA for non ELLs (.71) and moderate for ELLs (.61). For fifth grade students, the correlations were moderate for both ELL (.57) and non ELL students (.69). While the students in the Wiley and Deno study were younger and from a different language background (Hmong) than the current population, it is still evident that R-CBM (oral reading fluency) is an important tool to utilize with ELL students in order to predict future performance in reading.

Research has begun to examine the reliability and validity of the Maze task with the ELL population. Teachers often question the validity of oral reading fluency for ELL students, as many believe that ORF only measures the ability of ELL students to decode the words as opposed to measuring their reading comprehension skills (Wiley & Deno, 2005). As such, the Maze task may provide a more accurate measure of ELL students' overall reading proficiency versus oral reading fluency alone. The present study examined the predictive ability of the Maze task for both ELL and non-ELL students. The regression analyses did not find the Maze to be a significant predictor for the overall student population. Moreover, Maze was not a significant predictor for ELL and non-

ELL students. These results are somewhat inconsistent with previous research by Wiley and Deno. Their research found that while the Maze task was a better predictor than oral reading fluency for fifth grade non ELL students, it was less predictive than oral reading for the ELL population. Furthermore, the Maze scores did not significantly contribute to the performance on a high-stakes reading assessment for ELL students, while it did significantly contribute for non-ELLs. Unlike Wiley and Deno, the present study did not find the Maze task to be a significant predictor for non-ELL students. However, results are consistent in that the Maze task does not seem to be a significant predictor of student performance on high-stakes assessments for ELL students. This data refutes the assumption that the Maze task may provide a more accurate measure of ELL students' overall reading proficiency versus oral reading fluency alone. However, in another study by Vanderwood, Linklater, and Healy (2008), the Maze was used as an outcome measure along with R-CBM and a norm-referenced, standardized achievement test with ELL students. Results showed Maze to have strong positive correlations with Nonsense Word Fluency scores in fall, winter, and spring (.50, .51, and .54, respectively), with R-CBM (.66), and with ELL Level (.56). Maze was moderately correlated with the standardized achievement test in reading (.27). For the present study, the Maze was also significantly correlated with Fall R-CBM (.59) and AIMS reading scores (.41). It is evident from the varying results of the current research and the afore-mentioned studies that further research needs to investigate whether the Maze assessment could be utilized as a valid and reliable measure of reading in English Language Learners.

Implications for Instruction

Many secondary students are reading significantly below grade level and are making inadequate progress on the National Assessment of Educational Progress (U.S. Department of Education, 2005). According to Kamil (2003), more than 25% of middle-school students cannot read well enough to identify the main idea of reading passages. It is obvious from the research that the current educational system is not effectively meeting the needs of secondary students. Many secondary students who are reading below grade level will most likely continue to experience failure and may end of dropping out of school. The current educational system expects that secondary students are able to read fluently and accurately at grade level; however, some secondary students who are struggling with reading are lacking sufficient decoding, fluency, vocabulary, and comprehension skills (Kamil, 2003). As such, there is a need for a system in which student's at risk of reading failure are identified early in the school year and are then provided with appropriate intervention in order to remediate those weaknesses and improve reading achievement. Starting intervention immediately upon entering middle school or high school may make the difference in the success or failure for students who are struggling readers (Gibbs, 2009). A school-wide system of screening, targeted interventions, and progress monitoring is crucial for improving secondary students' achievement outcomes.

Response to Intervention (RTI) has been extensively researched with elementary-aged students and has shown to be a successful model for improving student achievement. Due to RTI's success in elementary schools, many school districts have

begun expanding their RTI models into middle school, and even high school. However, there is not nearly as much research on using an RTI framework with secondary schools. Nonetheless, RTI shows promise in improving student achievement even at the secondary level. In a meta-analysis by Edmonds et al. (2009) which examined the effectiveness of reading interventions with older struggling readers, it was found that students with reading difficulties in grades 6-12 could improve their comprehension when provided with targeted intervention in comprehension and/or multiple reading components. Identifying screening assessments that can be administered in the Fall that are predictive of later achievement in reading allows students at-risk to be targeted early in the school year for intervention. The results of the current study demonstrate that Fall R-CBM is a significant predictor of success or failure on a high-stakes reading assessment, AIMS for sixth grade ELL and non-ELL students. When students are receiving intervention, their response to those interventions needs to be directly and frequently monitored. The current study showed that R-CBM is still a useful tool in monitoring student growth over one school year, even at the sixth grade level. However, other research has shown that increased accuracy and fluency rates do not always lead to an improvement in reading comprehension (Allinder et al., 2001; Kuhn & Stahl, 2003). Thus, while the present study showed that reading fluency as measured by R-CBM is still an effective screening tool to identify sixth grade students at-risk of reading failure and in monitoring their growth over one school year, the interventions should not focus solely on building fluency. For students who are struggling with word-level skills, it is crucial to build

those word-reading skills (i.e., word recognition, fluency, word attack) while also explicitly teaching comprehension strategies (Edmonds et al., 2009).

ELL students tend to be at a higher risk for reading failure compared to their non-ELL peers. Hispanic students are almost twice as likely as non-Hispanic Whites to be reading below the expected level for their age (Snow, Burns & Griffin, 1998 in Gunn et al., 2000). The present study supports this notion, in that the ELL students' initial status as measured by the Fall R-CBM assessment was significantly lower than their non-ELL peers. What's more, there was not a significant difference in growth rates from Fall to Spring between these two populations. Thus, the ELL students are starting lower and continue to stay lower than their non-ELL peers in terms of their reading performance. This again points to the importance of using screening tools, such as R-CBM, early in the school year to identify these students at risk of failure and provide intensive levels of intervention in order to change their growth trajectories. Based on this researcher's knowledge of the academic setting of the present study, limited efforts were made to provide consistent, intensive levels of intervention for at-risk students. A voluntary after-school "tutoring" program was offered three days per week for students who were classified as "at-risk" on the R-CBM benchmark assessments. However, student attendance was inconsistent, and specific intervention programs were not utilized. It is hoped that if the ELL students in the present study who had been identified as at-risk based on their Fall R-CBM scores were provided with consistent, appropriate intervention, their growth trajectories would have increased significantly during the school year, eventually closing the gap between the ELL and non-ELL students.

What constitutes an “appropriate intervention” for older ELL students is an important question to consider. There has been very little research conducted specifically with intervention effectiveness with older Spanish-speaking ELL students. One study by Denton et al. (2008) examined the effectiveness of a multicomponent reading intervention with middle-school students (grades 6-8) with severe reading difficulties. The students in the intervention group received daily explicit and systematic small-group intervention for 40 minutes over 13 weeks. The intervention consisted of a modified version of a phonics-based remedial program, as well as instruction in vocabulary, fluency, and comprehension strategies. Results revealed that students’ in the intervention group did not demonstrate significantly higher outcomes in word recognition, comprehension, or fluency than the students’ receiving the typical instruction. The authors cite several reasons for the minimal response to intervention, including language development, self-efficacy beliefs, and low oral vocabulary. The authors conclude that very little is known about effective interventions for this specific population, and discussed several instructional factors to be considered when designing and implementing interventions for older ELL students. These instructional factors included the intensity of the intervention (e.g., 2 hours per day versus 40 minutes per day; smaller student: staff ratio, etc.), the length of the intervention (e.g., 2 years versus 13 weeks), and specific instructional components (e.g., having a stronger focus on oral language development or vocabulary). It may be very challenging for secondary schools to provide this type of intensive intervention. However, schools need to strongly consider the needs of their students and decide upon what is the most important goal for these students. If the goal is

it to improve the students' reading achievement, then that needs to become the focus of the school day. In order to meet this goal, schedules may need to be rearranged, and staff may need to be more flexible. This will most likely be a challenging task, but by investing time and energy in designing and implementing effective interventions for students, the improvement in student outcome will provide lasting benefits for both the individual students and the school.

Limitations for the Present Study

Certain limitations should be taken into account when considering the results of the present study. First, the power of the study would be increased if there were a larger number of students included in the study. For the purposes of the HLM analyses, only those students who took all of the R-CBM (Fall, Winter, and Spring), Maze, and AIMS assessments could be included in the study. With this requirement, only 349 total students were included. Moreover, all students were from one middle-school. Future studies should attempt to include a greater number of schools and students in the sample to provide further support for the R-CBM and Maze predictions on student outcome, as well as initial status and average rate of growth on R-CBM for this age group.

Another limitation of the current study was its inability to take in to account student's actual level of language proficiency. While students were grouped into ELL versus non-ELL based on whether they took the AZELLA (only ELL students are required to take the test), they could not be further classified into levels of language proficiency due to the lack of students at certain levels (e.g., no students in Level 2, Level 1 had only 2 students, and Level 3 had only one student). It would be beneficial for

future studies to examine students' level of language proficiency and how that may impact their initial status and growth on R-CBM, as well as the predictive ability of R-CBM and Maze on high-stakes assessments.

For the current study, a Maze assessment was only given in the spring prior to the administration of the AIMS assessment. As such, it could not be examined as another tool for monitoring student growth over the school year. Moreover, the Maze was not found to be a significant predictor of student performance on the AIMS reading assessment. It is quite possible that the outcome may have been different if a Maze assessment had been given in the Fall and could have been examined as another measure of initial status along with the Fall R-CBM score. Further research should continue to examine the usefulness of the Maze in predicting older ELL and non-ELL students' outcomes on high-stakes measures, as well as examine the effectiveness of the Maze in detecting changes in student growth over time.

Another limitation of the present study relating to language proficiency is the potential psychometric issues with ELL students and their performance on standardized achievement tests, in this case, the AIMS. The present study found that ELL students performed significantly lower as a whole on the AIMS reading assessment in the spring. Previous research has shown that ELL students' assessment results on standardized achievement tests may be confounded by their language background, which leads to problems with test reliability and construct validity (Abedi, 2002). This needs to be taken into consideration when examining the current results. In order to help remediate these psychometric issues, Abedi recommends that the psychometric characteristics of

standardized assessments should be carefully reviewed with ELL students and that when assessing ELL students, efforts should be made to reduce the confounding effects of language background on assessment outcomes. These issues may be difficult to remediate, as long as states continue to require all students, regardless of language background, to take the same high-stakes assessment to measure student achievement. Nonetheless, future studies should take this into account by examining level of language proficiency and its effect on student performance on high-stakes standardized achievement tests.

A final limitation of the current study is that it utilized archival data from a school district. As such, the researcher was unable to examine interscorer reliability on the R-CBM and Maze measures. Although the staff that administered the screening assessments had been trained in administration and scoring, the extent of this training is not known. As such, a representative estimate of interobserver agreement was unable to be generated.

Suggestions for Future Research

The implications of the present study point to the need to further examine R-CBM growth for ELL and non-ELL students at the middle school level. Further research also needs to investigate the use of curriculum-based measures in preventing reading failure on high-stakes outcomes, specifically for older students. Although this study addressed the need for research on R-CBM with the older ELL population, further research needs to address this with larger sample sizes, over a longer period of time, and across schools and districts. Specifically, future research could examine average rate of growth on R-CBM

throughout middle-school, from sixth grade through eight grade, to determine whether it continues to be a useful tool for monitoring student growth for both ELL and non-ELL students. Another important area for future research is to determine if the current results generalize to other school districts. Further research needs to include not only a larger number of subjects and over a longer period of time, but also include other high-stakes reading measures used by other school districts to ensure generalizability of the results of this study.

Although the results from the current study are promising in terms of the usefulness of R-CBM with middle school ELL and non-ELL students, more research is needed on the potential of R-CBM and Maze for this population. The results of the current study conflict with a previous study that examined the usefulness of Maze in predicting student outcome on high-stakes testing (Wiley & Deno, 2005). While Wiley and Deno (2005) found Maze to be a significant predictor for non-ELL students, the present study did not support this finding. The Maze was not a significant predictor for either ELL or non-ELL students on the AIMS reading assessment. Due to this conflicting data, future research needs to examine the Maze assessment along with R-CBM and its potential in measuring student growth and predicting success on high-stakes outcomes.

The use of data visualization was critical to the current study in understanding the complex relationships among all of the variables (i.e., ELL status, AIMS performance level, R-CBM initial status and growth, etc.). While certain interaction effects were not statistically significant, by examining the parallel coordinates, trellis plots, and

scatterplots, interesting patterns emerged. For instance, the parallel coordinates of student growth patterns by their AIMS level revealed that students at AIMS Level 2 (approached the standard) and AIMS level 3 (meets the standard) showed more consistent rates of growth from Fall to Spring on R-CBM than students at AIMS Level 1 (falls far below the standard) and Level 4 (exceeds the Standard). These findings have important instructional implications in that students at Level 1 that are failing far below the standard on AIMS will require more intensive intervention in order to change their growth trajectories on R-CBM. By changing their growth trajectories, they may have a better chance of approaching or meeting the standard on the AIMS reading assessment in the spring. It would be valuable for future research, particularly multi-level modeling, to utilize data visualization to further examine the complex relationships among the variables (e.g., R-CBM rate of growth, Maze, ELL status, level of language proficiency, etc.).

While the current study did not investigate intervention effectiveness, the results did point to the importance of providing effective intervention to students at-risk for reading failure as measured by the Fall R-CBM score, particularly ELL students. The current results found that ELL students were significantly lower than non-ELL students in their initial status on R-CBM in the Fall, but their growth rates from Fall to Spring did not vary significant from their non-ELL peers. In other words, ELL students started the year and ended the year significantly below their non-ELL peers in reading. These results support the need for future research on intervention effectiveness in reading for older students, particularly older ELL students. It will be crucial for future research to

examine what constitutes an effective intervention program for these students in terms of intensity, length of the intervention, frequency, and the actual instruction/programs that are most effective in improving student achievement in reading.

Conclusion

Currently, there is a political and educational emphasis on accountability, high-stakes assessment, and student outcomes. Districts are being held accountable for the outcomes of all students based on the results of standards-based, high stake assessments. Schools not only have to show that 95% of their students are proficient on state tests, but individual subgroups, including English Language Learners, must also meet the 95% goal (Wiley & Deno, 2005). This puts schools in a difficult position, because by the time this assessment is given (usually in the spring), it is too late to identify students at risk as well as to determine whether reading instruction was effective in improving student outcomes. For this reason, many districts have been using CBM as a way of identifying early in the school year students who are at risk of reading failure, as well as measuring student progress over time. In addition, recent research has been promising in the use of oral reading fluency measures as a predictor of whether students will meet standards on state tests. While CBM has been well established as an effective assessment tool for English speaking students, there has been limited research on the use of CBM with ELL students. According to McCardle, Mele-McCarthy and Leos (2005), there is a need for accurate and user-friendly assessment tools that schools can utilize for screening and progress monitoring with the ELL population. Although limited, the body of research on ELL students and reading in the elementary grades is growing. However, there has been

minimal research done with middle school ELL students (i.e., grades 6-8). According to Denton et al. (2008), the research on effective intervention for older readers with reading difficulties is lacking, particularly with students who are English Language Learners (ELLs). Often, students are starting middle school (sixth grade) with reading skills that are falling significantly below grade level expectations. More research is needed on how to effectively identify, intervene, and monitor these students in order to improve reading outcomes for this population.

The results of the current study support the use of R-CBM with sixth grade ELL and non-ELL students to identify students at risk of reading failure, monitor student growth over one school year, and predict student performance on a high-stakes assessment in the spring. It appears from the current research that a student's initial status as measured by the Fall R-CBM is the most potent predictor of student performance on a high-stakes reading assessment. A student's percentage of growth from Fall to Spring was also a significant predictor of performance on the high-stakes assessment, but only for the non-ELL population. The current research also supports the notion that ELL students start the school year significantly below their peers in terms of reading skill, and that they remain behind their peers as indicated by their lower performance on the high-stakes reading assessment in the spring. This is also evidenced by the fact that ELL students do not vary significantly in their growth rates from Fall to Spring compared to their non-ELL peers. As such, they are starting behind and ending behind, and without some type of intervention to change that growth trajectory, they will continue to remain behind. Developing a school-wide system of assessment and

intervention at the secondary level to monitor progress over time can be a powerful tool to alter the course of reading trajectories for both ELL and non-ELL students and may help in ensuring later success on high-stakes assessments.

Several studies have revealed that R-CBM is both an accurate and efficient way of predicting students who are likely to pass the reading portion on state accountability tests as far back as kindergarten and first grade (Hintze & Silbergitt, 2005; Sibley, Biwer, & Hesch, 2001; Vanderwood, Linklater, & Healy, 2008). It is obvious from this research that it is imperative for school districts to be utilizing some form of CBM in reading to screen and identify students at risk for reading failure. It is well documented in the literature that early intervention is key in remediating academic deficits, yet many students are still entering middle school with significantly below grade-level academic skills, particularly in reading. The students in the present study, on average, were falling significantly below sixth grade level expectations in reading based on the average Fall R-CBM score. What's more, the ELL students were significantly below their non-ELL peers in their initial status in the Fall on R-CBM, and did not make sufficient growth throughout the school year to close the gap. While it would be easy to look back and blame these students' poor reading performance on the elementary school and what they did not do, that will not solve the problem of what is happening for these students now. The present study emphasizes the absolute urgency for all schools, including middle and secondary schools, to put in place a system utilizing CBM to identify students at-risk of reading failure and provide consistent, intensive intervention to identified students. By the time low performing students have reached middle school, they may have

experienced years of frustration and academic failure. It is the school systems' responsibility to stop this cycle of poor performance and intervene. It is never too late to help a student become more successful and prevent them from dropping out of school.

REFERENCES

- Abedi, J. (2002). Standardized achievement tests and English language learners: Psychometric issues. *Educational Assessment, 8*(3), 231-257.
- Adams, M.J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Allinder, R.M., Dunse, L., Brunken, C.D., & Obermiller-Krolikowski, H.J. (2001). Improving fluency in at-risk readers and students with learning disabilities. *Remedial and Special Education, 22*, 48-45.
- Ardoin, S.P., Witt, J.C., Suldo, S.M., Connell, J.E., Koenig, J.L., Resetar, J.L., Slider, N.J., & Williams, K.L. (2004). Examining the incremental benefits of administering a maze and three versus one curriculum-based measurement reading probes when conducting universal screening. *School Psychology Review, 33*(2), 218-233.
- Arizona Department of Education. (2009). *Arizona's Instrument to Measure Standards: 2007 Technical Report*. Retrieved August 6, 2009, from www.ade.state.az.us/standards/AIMS/Technical/Legal/AIMStechreport2007.pdf.
- Arizona Department of Education. (2007). *Arizona English Language Learner Assessment (AELLA): Technical Manual*. Retrieved August 7, 2009, from www.a.ed.gov/oelas/AELLA/AELLAA-1TechnicalManual.pdf.
- August, D., & Hakuta, K. (Eds.). (1997). *Improving schooling for language-minority children: A research agenda*. Washington, DC: National Academy Press.
- Baker, S.K., & Good, R. (1995). Curriculum-based measurement of English reading with bilingual Hispanic students: A validation study with second-grade students. *School Psychology Review, 24*, 561-578.
- 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.
- Barwick, M., & Siegel, L.S. (1996). Learning difficulties in adolescent clients of a shelter for runaway and homeless street youth. *Journal of Research on Adolescence, 6*, 649-670.

- Belsley, D.A., Kuh, E., & Welsch, R.E. (1980). *Regression Diagnostics: Identifying influential data and sources of collinearity*. New York: John Wiley.
- Berry, W.D., & Feldman, S. (1985). *Multiple regression in practice*. London: Sage Publications.
- Brown-Chisdey, R., Johnson, P., & Fernstrom, R. (2005). Comparison of grade-level controlled and literature-based maze CBM reading passages. *School Psychology Review, 34*(3), 387-394.
- Brown-Chisdey, R., Davis, L., & Maya, C. (2003). Sources of variance in curriculum-based measures of silent reading. *Psychology in the Schools, 40*, 363-377.
- Bryk, A.S., & Raudenbush, S.W. (1992). *Hierarchical linear models*. Newsbury Park, CA: Sage.
- Chall, J.S. (1983). *Stages of reading development*. New York: McGraw Hill.
- Cohen, J., Cohen, P., West, S.G., & Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Crawford, L., Tindal, G., & Stieber, S. (2001). Using oral reading rate to predict student performance on statewide achievement tests. *Educational Assessment, 7*(4), 303-323.
- Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in Bilingual children* (pp.70-89). Cambridge, MA: Cambridge University Press.
- Cummins, J. (1984). *Bilingual and special education: Issues in assessment and pedagogy*. Austin, TX: PRO-ED.
- Data Description Inc. (2008). DataDesk 6.2. Ithaca, NY: Author.
- Deno, S.L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*, 219-232.
- Deno, S.L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education, 37*, 184-192.
- Deno, S.L., Mirkin, P.K., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children, 49*, 36-45.

- Denton, C.A., Wexler, J., Vaughn, S., & Bryan, D. (2008). Intervention provided to linguistically diverse middle school students with severe reading difficulties. *Learning Disabilities Research and Practice, 23*(2), 79-89.
- De Ramirez, R.D., & 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.
- Edmonds, M.S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K.K., Schnakenberg, J.W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research, 79*(1), 262-300.
- Ehren, B.J., Lenz, B.K., & Deshler, D.D. (2004). Enhancing literacy proficiency with adolescents and young adults. In K. Apel, B.J., Ehren, E.R. Silliman, & C.A. Stones (Series Eds.) & C.A. Stones, E.R. Silliman, B.J. Ehren, & K. Apel (Vol. Eds.), *Challenges in language and literacy; Handbook of language and literacy: Development and disorders* (pp. 680-701). New York: Guilford.
- Espin, C.A., & Tindal, G. (1998). Curriculum-based measurement for secondary students. In M.R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 214-253). New York: Guilford Press.
- Fletcher, J.M. (2007, February). Overview of the Texas Center for Learning Disabilities. Paper presented at the Pacific Coast Research Conference, San Diego, CA.
- Fuchs, L.S. (2004). The past, present, and future of curriculum-based research. *School Psychology Review, 33*(2).
- Fuchs, L.S., Fuchs, D., & Maxwell, L. (1988). A validity of informal reading comprehension measures. *Remedial and Special Education, 2*, 20-28.
- Fuchs, L.S., Fuchs, D., Hamlett, C.L., Waltz, L., & Germann, G. (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. (2000). Issues in the assessment of reading disabilities in L2 children. *Dyslexia, 6*, 13-28.

- Gersten, R., & Baker, S. (2000). What we know about effective instructional practices for English-language learners. *Exceptional Children*, 66(4), 454-470.
- Gersten, R., & Geva, E. (2003). Teaching reading to early language learners. *Educational Leadership*, 60(7), 44-50.
- Gibbs, D. P. (2009). *RTI in middle and high school: Strategies and structures for literacy success*. Horsham, PA: LRP Publications.
- 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.
- Gunn, B., Smolkowski, K., Biglan, A., Black, C., & Blair, J. (2005). Fostering the development of reading skill through supplemental instruction: Results for Hispanic and Non-Hispanic students. *The Journal of Special Education*, 39(2), 66-85.
- Gunn, B., Biglan, A., Smolkowski, K., & Ary, D. (2000). The efficacy of supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school. *The Journal of Special Education*, 34(2), 90-104.
- Haager, D., & Windmueller, M. P. (2001). Early reading intervention for English language learners at-risk for learning disabilities: Student and teacher outcomes in an urban school. *Learning Disability Quarterly*, 24, 235-250.
- Hasbrouck, J.E., & Tindal, G. (2005). *Oral reading fluency: 90 years of measurement. Technical Report*. Retrieved May 27, 2010 from www.brtpjects.org/publications/tech_reports/techRpt33_FluenctNorms.pdf.
- Hasbrouck, J.E., & Tindal, G. (1992). Curriculum-based oral reading fluency norms for students in grades 2 through 5. *Teaching Exceptional Children*, 24(3), 41-44.
- Hintze, J.M., Callahan, J.E., Matthews, W.J., Williams, S.A.S., & Tobin, K.G. (2002). Oral reading fluency and prediction of reading comprehension in African-American and Caucasian elementary school children. *School Psychology Review*, 31, 540-553.
- Hintze, J.M., & Silberglitt, B. (2005). A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high-stakes testing. *School Psychology Review*, 34, 372-386.

- Howe, K.B., & Shinn, M.M. (2002). *Standard reading assessment passages (RAPs) for use in General Outcome Measurement: A manual describing development and technical features* [Technical Manual]. Eden Prairie, MN: Edformation.
- Kamil, M.L. (2003). *Adolescents and literacy: Reading for the 21st century*. New York: Alliance for Special Education, Carnegie Foundation.
- 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.
- Klein, J.R., & Jimerson, S.R. (2005). Examining ethnic, gender, language, and socioeconomic bias in oral reading fluency scores among Caucasian and Hispanic students. *School Psychology Quarterly, 20*(1), 23-50.
- Klinger, J.K., Artiles, A.J., & Barletta, L.M. (2006). English Language Learners who struggle with reading: Language acquisition or LD? *Journal of Learning Disabilities, 39*(2), 108-128.
- Kranzler, J.H., Miller, M.D., & Jordan, L. (1999). An examination of racial ethnic and gender bias on curriculum-based measurement in reading. *School Psychology Quarterly, 14*, 327-342.
- Kranzler, J.H., & Taylor, G.S. (2000, March). Another look at racial ethnic and gender bias on CBM. Poster presented at the Annual Convention of the National Association of School Psychologists, New Orleans, LA.
- Kuhn, M.R., & Stahl, S.A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95*(1), 3-21.
- Leach, J.M., Scarborough, H.S., & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of Educational Psychology, 95*(2), 211-224.
- Linan-Thompson, S., Vaughn, S., Hickman-Davis, P., & Kouzekanani, K. (2003). Effectiveness of supplemental reading instruction for second-grade English language learners with reading difficulties. *The Elementary School Journal, 103*(3), 221-238.
- Lipka, O., Lesaux, N.K., & Siegel, L.S. (2006). Retrospective analyses of the reading development of grade 4 students with reading disabilities: Risk status and profiles over 5 years. *Journal of Learning Disabilities, 39*(4), 364-378.

- Macias, R.F. (2000). *Summary report of the survey of the states' limited English proficient students and available educational programs and services, 1997-98*. Washington, DC: National Clearinghouse for Education.
- Marston, D.B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and how to do it. In M.R. Shinn (Eds.), *Curriculum-based measurement: Assessing special children* (pp.18-78). New York: The Guilford Press.
- McBride, H., & Siegel, L.S. (1997). Learning disabilities and adolescent suicide. *Journal of Learning Disabilities, 30*, 652-659.
- McCardle, P., Mele-McCarthy J., & Leos, K. (2005). English language learners and learning disabilities: Research agenda and implications for practice. *Learning Disabilities Research and Practice, 20*(1), 68-78.
- McLaughlin, B. (1992). *Myths and misconceptions about second language learning: What every teacher needs to unlearn*. Washington, DC: National Center for Research on Cultural Diversity and Second Language Learning.
- 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.
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.
- Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models, applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Rueda, R., & Windmueller, M.P. (2006). English language learners, LD, and overrepresentation: A multiple-level analysis. *Journal of Learning Disabilities, 39*(2), 99-107.
- SAS Institute. (2009). JMP 8 [Computer software and manual]. Cary, NC: Author. SAS Institute. (2008). SAS 9.2 [Computer software and manual]. Cary, NC: Author.
- Scammacca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J., Reutebuch, C. K., et al. (2007). *Reading interventions for adolescent struggling readers: A meta-analysis with implications for practice*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.

- Scribner, A.P. (2002). Best assessment and intervention practices with second language learners. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology IV* (pp.1485-1498). Bethesda, MD: National Association of School Psychologists.
- Shin, J., Espin, C.A., Deno, S.L., & McConnell, S. (2004). Use of hierarchical linear modeling and curriculum-based measurement for assessing academic growth and instructional factors for students with learning difficulties. *Asia Pacific Education Review*, 5(2), 136-148.
- Shinn, M.R. (2009). Best practices in using curriculum-based measurement in a problem-solving model. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 243-261). Bethesda, MD: National Association of School Psychologists.
- Shinn, J., Deno, S.L., & Espin, C. (2000). Technical adequacy of Maze task for curriculum-based measurement of reading growth. *Journal of Special Education*, 34, 264-173.
- 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.
- Silberglitt, B., & Hintze, J.M. (2007). How much growth can we expect? A conditional analysis of R-CBM growth rates by level of performance. *Exceptional Children*, 74(1), 71-84.
- Silberglitt, B., Burns, M.K., Madyun, N.H., & Lail, K.E. (2006). Relationship of reading fluency assessment data with state accountability test scores: A longitudinal comparison of grade levels. *Psychology in the Schools*, 43(5), 527-535.
- 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.
- Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360-406.

- Stecker, P.M., Lembke, E.S., & Foegen, A. (2008). Using progress monitoring data to improve instructional decision making. *Preventing School Failure, 52*(2), 48-58.
- TIBCO (2009). Spotfire Miner [Computer software and manual]. Palo Alto, CA: Author.
- Twyman, T., & Tindal, G. (2007). Extending curriculum-based measurement into middle/secondary schools: The technical adequacy of the concept maze. *Journal of Applied School Psychology, 24*(1), 49-67.
- U.S. Census Bureau, (2000). *Census 200 Summary File 3*, Washington, DC: Bureau.
- U.S. Department of Education, Institute of Education Sciences, National Assessment of Educational Progress, National Center for Education Statistics. (2005). *The nation's report card, reading*. Washington, DC.
<http://nces.ed.gov/nationsreportcard/>.
- Vanderwood, M.L., & Nam, J. (2009). Best practices in assessing and improving English language learners' literacy performance. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 243-261). Bethesda, MD: National Association of School Psychologists.
- 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.
- Vaughn, S.R., Fletcher, J.M., Francis, D.J., Denton, C.A., Wanzek, J. Wexler, J., Cirino, P.T., Barth, A.E., Romain, M. A. (2008). Response to intervention with older students with reading difficulties. *Learning and Individual Differences, 18*, 338-345.
- Wagner, R.K., Francis, D.J., & Morris, R.D. (2005). Identifying English language learners with learning disabilities: Key challenges and possible approaches. *Learning Disabilities Research and Practice, 20*(1), 6-15.
- 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.
- Wood, D.E. (2006). Modeling the relationship between oral reading fluency and performance on a statewide reading test. *Educational Assessment, 11*(2), 85-104.

VITA

Nicole received her undergraduate degree in psychology from the University of Illinois at Urbana-Champaign and her specialist's degree in School Psychology from Eastern Illinois University. Before returning to school to obtain a doctoral degree in School Psychology at Loyola University Chicago, Nicole worked for four years as a school psychologist in the Chicago suburbs. Nicole continued to practice school psychology full-time while she pursued her doctoral degree. She is currently a certified School Psychologist employed with the Litchfield Elementary School District in Litchfield Park, Arizona. Upon completing her dissertation, Nicole will graduate in December 2010 with her Doctor of Philosophy Degree in School Psychology from Loyola University Chicago.

DISSERTATION APPROVAL SHEET

The Dissertation submitted by Nicole Osterman Stokes has been read and approved by the following committee:

David Prasse, Ph.D., Director
Dean, School of Education
Loyola University Chicago

Theresa Pigott, Ph.D.
Associate Professor, School of Education
Loyola University Chicago

Diane Morrison, Ed.D.
Clinical Assistant Professor, School of Education
Loyola University Chicago

The final copies have been examined by the director of the Dissertation Committee and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the Dissertation is now given final approval by the committee with reference to content and form.

The Dissertation is therefore accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Date

Director's Signature