Utah State University
DigitalCommons@USU

All Graduate Theses and Dissertations | Graduate Studies, School of

1-1-2012

# An Exploration of Aggregated Patterns of Student Curriculum-Based-Measurement Outcome Data Within a Response to Intervention Program

Elizabeth Findlay

Recommended Citation
Findlay, Elizabeth, "An Exploration of Aggregated Patterns of Student Curriculum-Based-Measurement Outcome Data Within a Response to Intervention Program" (2012). *All Graduate Theses and Dissertations.* Paper 1433.
http://digitalcommons.usu.edu/etd/1433

This Thesis is brought to you for free and open access by the Graduate Studies, School of at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact becky.thoms@usu.edu.

UtahStateUniversity
MERRILL-CAZIER LIBRARY

# AN EXPLORATION OF AGGREGATED PATTERNS OF STUDENT CURRICULUM-BASED-MEASUREMENT OUTCOME DATA WITHIN A RESPONSE TO INTERVENTION PROGRAM

by

Elizabeth Findlay

A thesis submitted in partial fulfillment of
the requirements for the degree

of

Educational Specialist

in

Psychology

Approved:

| | |
|---|---|
| Donna Gilbertson, Ph.D.<br>Major Professor | Tim Slocum, Ph.D.<br>Committee Member |
| Gretchen Gimpel Peacock, Ph.D.<br>Committee Member | Mark R. McLellan, Ph.D.<br>Vice President for Research and<br>Dean of the School of Graduate Studies |

UTAH STATE UNIVERSITY
Logan, Utah

2012

Copyright © Elizabeth Findlay 2012

All Rights Reserved

# ABSTRACT

An Exploration of Aggregated Patterns of Student Curriculum-Based-Measurement Outcome Data Within a Response to Intervention Program

by

Elizabeth Findlay, Educational Specialist

Utah State University, 2012

Major Professor: Donna Gilbertson, Ph.D.
Department: Psychology

One major concern when developing a response to intervention (RTI) program is to select effective practices that will be successfully implemented and sustained with adequate organizational guidance and support. The purpose of this study was to explore patterns of student tier placement data as a school-based case example of the nature and utility of RTI in an applied setting. Specifically, this study aimed to explore the extent that the percentages of students placed in a three-tier program based on student oral reading fluency (ORF) level and growth trajectories reflect the standard RTI tier placement (80%, 15%, and 5%) at fall, winter, and spring in a school setting. Percentages of the total student population tier placement were explored with ORF data from third- and fourth-grade students ($N = 429$) at two schools in fall, winter, and spring. Results showed that school and ORF data reflected the standard percentages of student populations within each tier in fall, winter, and spring. However, slope data showed

greater percentages of students in the more intensive tiers. Moreover, flexible grouping, or movement between tiers occurred for few students when movement occurred based on school or ORF level data. No significant differences were found between the school and ORF student tier placements in fall, winter, and spring. A significant difference was found in spring between placement methods with a larger proportion of students in Tier 1 based on the school assignments and a larger proportion of students in Tier 2 and Tier 3 based on ORF slope assignments.

(88 pages)

# PUBLIC ABSTRACT

An Exploration of Aggregated Patterns of Student Curriculum-Based-Measurement Outcome Data Within a Response to Intervention Program

by

Elizabeth Findlay, Educational Specialist

Utah State University, 2012

Identifying and determining what a student needs in order to progress and succeed in school is an important aspect of education. One proposed model for doing so is called response to intervention (RTI). This model states that the degree to which a student does or does not respond to high-quality interventions can help predict future performance and provide needed insight into what skills a student does and does not have. A student receives more or less services based on his or her level of responding interventions provided. The standard RTI model indicates that 80% of a student population should respond to typical classroom instruction (Tier 1), 15% of students will not respond to this instruction and will require more intense interventions to progress (Tier 2), and 5% of a student population will not respond to Tier 1 or Tier 2 instruction and intervention and will, therefore, need even more intense and individualized help to progress (Tier 3). This model aims to help the most students in the most effective way. However, there are many unanswered questions about RTI's utility within the school systems. For RTI to be an effective means of helping students, it is critical that research be conducted to determine how it should be utilized within a school system. Therefore, the following thesis investigated a functioning RTI model within an applied setting.

Data were collected and analyzed from a school system that utilized RTI during the 2009-2010 school year. This thesis looked specifically at if and how RTI worked in identifying and supporting students in the area of reading. Results showed that overall RTI can be a viable option for identifying and helping students who need support. Many interesting patterns were found in this study including noting that student placement within the tiers of RTI depends on method chosen to evaluate along with particular criterion chosen. Further, results indicated that flexible grouping, or movement between tiers, has the potential to help many students, however how frequently movement should be evaluated still needs to be determined. Finally, this project further supported previous research noting that allocation of resources and a systematic and structured delivery method of RTI are important for success.

## ACKNOWEDGMENTS

First and foremost I would like to thank Dr. Donna Gilbertson for her continual effort and support through not only the thesis process, but also throughout my entire graduate career. Dr. Gilbertson has been there to support me and provide greatly needed insight. She has helped me through each and every step of this monumental process from deciding a topic, to conducting a literature review, to writing results and discussion, and finally preparing for a defense. She has believed in me and this project when even I did not. Dr. Gilbertson is a truly exemplar professor, mentor, chair, and friend. I would also like to acknowledge my committee members, Dr. Gretchen Gimpel-Peacock and Dr. Timothy Slocum, for their input, guidance, and support as well.

I also need to thank my dedicated family and group of supportive friends who have helped me through this journey. Like Dr. Gilbertson, my friends and family have supported and guided me even when I was ready to give up. They have provided endless encouragement and late-night companions as I have worked to reach this accomplishment. I will always be indebted to them for their confidence and support. I am forever grateful to each and every person who has touched my life as I have undertaken and accomplished my educational goals.

Elizabeth Findlay

# CONTENTS

Page

ABSTRACT ..... iii

PUBLIC ABSTRACT ..... v

ACKNOWLEDGMENTS ..... vi

LIST OF TABLES ..... ix

CHAPTER

I. INTRODUCTION ..... 1

II. LITERATURE REVIEW ..... 6

Prevalence of Reading Problems and Problems with Traditional Identification Procedures ..... 7
Proposed RTI Models in the Literature ..... 11
RTI Empirical Support ..... 13
RTI and Progress Monitoring ..... 24
Empirical Support for CBM ..... 25
Purpose of Research ..... 29

III. METHODS ..... 32

Participants and Setting ..... 32
General RTI Program Overview ..... 32
Materials ..... 39
Dependent Variable ..... 40
Overview of Progress Monitoring Data Collection Procedures ..... 41
District RTI Progress Monitoring Database System ..... 41
Procedure ..... 42

IV. RESULTS ..... 43

Descriptive Summary of Data ..... 43
Research Questions ..... 45

Page

V. DISCUSSION ..... 52

Percentage of Students Served in Each Tier ..... 53
Growth Rates Across Time and Grades ..... 55
Flexible Movement of Student Tier Placement ..... 58
Limitations ..... 60
Practical Implications ..... 61

REFERENCES ..... 64

APPENDIX ..... 72

# LIST OF TABLES

Table Page

1. Demographic Data ..... 33
2. Percentage of Students Passing State English Exam ..... 33
3. AIMS Criterion for Oral Reading Fluency Level ..... 37
4. Mean Oral Reading Fluency Across Time at Each Grade Level Per School .... 44
5. Mean Slope for Oral Reading Fluency Across Time at Each Grade Level Per School ..... 44
6. Percentages of Students in School and ORF-Level-Determined Tiers..... 46
7. Percent of Student Showing Adequate Growth in School-Determined Tiers.... 48
8. Movement Between Tiers ..... 49
9. Spring Outcome of Students Moved Between Tiers in Winter ..... 51

# CHAPTER I

# INTRODUCTION

Reading is a fundamental skill all children need to learn. If a child obtains a good grasp of basic reading principles, it will affect other areas of his or her academic life. Gettinger and Stoiber (2007) have noted that children have roughly a 75% chance of not reading at grade level by the end of elementary school if they enter the second grade without being able to read. Thus, it is imperative that children be given every opportunity to learn to read early in their school years. In order for this to occur, there needs to be an effective way to identify early onset of reading problems to provide instruction to help remediate reading difficulties to prevent the development of a severe reading deficit.

Traditionally, special education has been the most common instructional support program option to supplement the general education curriculum in a school setting. As a result those students who experienced difficulties in the general education curriculum were frequently referred for special education evaluations to determine if poor performance was due to a learning disability. The discrepancy model has been most commonly used to identify any type of specific learning disability (SLD), including reading (Vellutino, Scanlon, Small, & Fanuele, 2006). According to this model, an SLD can be identified when a student's score on a test of general intelligence suggest that the student has the cognitive ability to learn but substantially lower scores on the achievement test suggests that the student is not learning as expected. There are several concerns regarding schools' use of the IQ-achievement discrepancy model as the sole method for identifying students with learning disabilities. Primary concerns of this model

include difficulty of early identification before reading problems become severe which leads to problems being more difficult to remediate and the lack of information provided about specific student instructional needs (Barnes & Harlacher, 2008; Burns & Ysseldyke, 2005).

A promising approach that has potential to help resolve the problems of the discrepancy model is response to intervention (RTI), authorized in the Individuals with Disabilities Education Act. RTI is a multi-tiered approach to providing behavioral and academic interventions to struggling learners at increasing levels of intensity. This model is based on the assertion that the degree to which a child does or does not respond to a series of high-quality interventions implemented with integrity is a good predictor of future performance and provides information about the type of support that a child may need to successfully catch up with his or her peers (Barnes & Harlacher, 2008). Further, continued low performance combined with effective teaching provides data that suggest that the poor performance is not due to poor instruction but may be due to a learning disability (Healy, Vanderwood, & Edelston, 2005).

RTI involves a shift from traditional psychometric standardized approaches of assessment to a more pragmatic, educationally relevant model focused on measuring changes in individual academic performance over time. This includes moving away from a “within child” deficit paradigm to a contextual perspective with greater emphasis on instructional intervention and progress monitoring prior to special education referral. Potential advantages to this model include a focus on obtaining positive academic outcomes for all students and better allocation of support that matches individual student

needs when problems first emerge through systematic decision making and progress monitoring (Burns & Ysseldyke, 2005; Vaughn & Fuchs, 2003).

Although there is emerging research validating the potential advantages of this model for identifying reading problems early and providing intervention support for more students, specifics of RTI procedures and processes remain inconsistent across school settings (Marston, 2005; Tilly, 2003; Vaughn, 2003). Interventions used in the RTI model will likely vary to meet the needs of different student populations across school settings. Given this uncertainty of the effectiveness of any RTI model that is developed to meet the needs of a certain school population, L. Fuchs and D. Fuchs (1998) have noted that psychometrically validated progress monitoring measures for decision making is critical for RTI to function at an optimally effective level. Frequent examination of student reading progress on progress monitoring measures several times a year provides data to make decisions about the effectiveness of the selected interventions at each tier and about appropriate student placement in each tier. To make instructional decisions, one progress monitoring tool, curriculum-based measures (CBM), are frequently incorporated into RTI (Shinn & Bamonto, 1998). CBM provides a psychometric validated system of measurement that is sensitive to individual responsiveness within a short period of time (Jenkins & Terjeson, 2011; Reschly, Busch, Betts, Deno, & Long, 2009; Wayman, Wallace, Wiley, Ticha, & Epsin, 2007). Ideally, educators frequently review student CBM data based outcomes to determine students need for more or less intensive instructional support (L. Fuchs & D. Fuchs, 2007; Healy et al., 2005; O'Connor, Harty, & Fulmer, 2005; VanDerHeyden, Snyder, Broussard, & Ramsdell, 2007a).

Despite the promising results of current research on RTI and CBM, there are still areas that need to be investigated more thoroughly to fully understand the optimal benefits of this model on student outcomes. For example, research has not been conclusive about the exact relationship between results on CBM data and student tier placement. Further, an RTI model is based on allocation of school based resources at various intervention tiers to ensure that students make academic progress (L. Fuchs & D. Fuchs, 2007; Healy et al., 2005; O'Connor et al., 2005; VanDerHeyden et al., 2007a). To meet this goal, the standard RTI model proposes that 80% of children within a school are adequately progressing with the universal general education curriculum (Tier 1) given that teachers are the primary resource in a school setting. This provides support from more specialized but fewer school-based personnel (e.g., reading specialists, school psychologist) to assist the remaining 15% of children who may be at-risk for problems and are placed and progressing in a more intensive small group instructional program (Tier 2) and 5% of students are high risk students placed and adequately progressing with individualized instruction (Tier 3). Ideally, an examination of aggregates of student CBM data for each tier will show correspondence to the suggested tier composition of the basic RTI model (80%-15%-5%). One important aspect of this process is the flexibility to move students between tiers as student needs change. However, it is unknown how this flexibility influences the balance between resources allocated to each tier and program success within a given school year. Despite the complexity of this process, only some of the current research has examined how the outcome data and the percentages of students in each tier in an RTI model corresponds with the standard RTI model and with expected

outcomes. To date, typical patterns of CBM performance, percentages of students within each tier, group flexibility, and how these patterns correspond to end of the year testing in an applied school setting remains unknown. Findings from such an exploration of CBM patterns would further facilitate research on relations among CBM-related patterns and RTI programs targeting academic outcomes in schools. Therefore, the purpose of this study is to explore the patterns based on real time CBM data and student Tier placement as a school based case example of the nature and utility of the CBM data within an RTI model in an applied setting.

# CHAPTER II

# LITERATURE REVIEW

One of the most basic skills a child needs to learn is how to read. The RTI model is a multi-tiered approach that seeks to prevent reading failure by providing several layers or tiers of increasingly intensive high quality instruction that can match different student needs. Student outcome performance is frequently monitored and used to inform decisions about each student's instructional needs. Student outcome data is also collected at each tier to determine which instruction should be maintained and modified. The use of CBM as progress monitoring (PM) tools is commonly incorporated into RTI to monitor student outcomes (Jenkins & Terjeson, 2011; Reschly et al., 2009; Wayman et al., 2007). Theoretically, through the tiered approach of RTI and its use of progress monitoring measures, struggling students would receive the appropriate frequency and intensity of intervention and support (L. Fuchs & D. Fuchs, 2007; Healy et al., 2005; O'Connor et al., 2005; VanDerHeyden et al., 2007a).

The purpose of this literature review is to summarize the rationale for RTI programs and previous research on the effectiveness of an RTI program on student reading performance. Thus, the first objective for this review is to provide a summary of the need for interventions for struggling readers and problems with traditional methods for identifying students who are struggling in reading. Second, a description of the RTI model as an alternative identification method will be presented followed by a review of research on the RTI program. The final purpose of this review is to summarize studies on the CBM progress monitoring tools that are most frequently used to make decisions

regarding the effectiveness of a RTI program.

For this literature review, a systematic review of the peer reviewed literature was conducted using electronic databases, ERIC, PsycINFO, Google Scholar and Academic Search Premier. Review and journal articles were reviewed containing empirically based studies on RTI programs published in the years 2003-2012. This search was conducted using the following descriptors were used to locate studies: RTI, Response to Intervention, Implementation of RTI, Progress Monitoring, RTI Validity, RTI Reliability, Application of RTI, RTI Program Implementation, and RTI Implementation Problems. Review and empirical studies on the psychometric properties and the utility of CBM for decision making were also identified and reviewed using the following descriptors to locate studies: curriculum based measurement, reading, progress monitoring, and oral reading fluency.

## Prevalence of Reading Problems and Problems with Traditional Identification Procedures

Learning to read is one of the most basic skills a child needs as he or she begins gaining an education. Unfortunately, the prevalence of reading problems among school age children is quite high. According to the Kennedy-Krieger Institute, reading disabilities account for 80% of all learning disabilities (Kennedy Krieger Institute, 2005). Further, specific learning disabilities account for a significant proportion of students who receive services under IDEA. In the 30th Annual Report to Congress from the Office of Special Education Programs (2007), 6,081,890 children ages 6-21 were receiving

services under IDEA (9.1% of the general population) from fall of 1997 through fall of 2006. Of those children, 44.6% (4% of the general population) were receiving services for specific learning disabilities. More current Annual Reports to Congress have not yet been released, however according to the recent 2011 Condition of Education report published by the National Center for Educational Statistics (2010), in the 2008-2009 school year 38.2% (2,476,000) of the 6,483,000 students served under IDEA fell within the Specific Learning Disability category (Aud, Hussar, & Kenna, 2011). These students also have a significantly lower chance of obtaining sufficient, or passing, national standards. Based on results from the 2011 Nation's Report Card on Reading, 67% of fourth-grade students performed at or above the Basic level, 34% performed at or above the Proficient level of reading, and 8% performed at or above the advanced level (Aud et al., 2011). Research has also shown that a child only has a 25% chance of reading at grade level by the end of elementary school if he/she enters the second grade without being able to read (Gettinger & Stoiber, 2007). Thus, it is imperative that to ensure that the literacy needs of all students are addressed and that the specific needs of struggling readers are identified and met in a timely manner.

To prevent reading difficulties, early identification of students experiencing reading problems is becoming a more prominent goal in educational settings. Research has shown that early identification as well as early intervention consistently helps poor readers catch up with their peers and helps improve overall outcomes (Vaughn, Wanzek, Murray, Linan-Thompson, & Woodruff, 2009). This increased effort to identify children with reading difficulties early in their educational careers has also been spurred due to

many problems with traditional identification procedures. The traditional means of identifying a child with a learning disability is to use the IQ-achievement discrepancy model. This model identifies specific learning disabilities (SLD) by demonstrating a severe discrepancy between performance on intelligence tests indicating adequate cognitive ability to learn and achievement tests indicating learning has not been achieved as cognitively expected (Vellutino et al., 2006).

Arguments in the literature about the measurement and conceptual flaws of the discrepancy model have resulted in an increased dissatisfaction with this approach. A primary argument is that the discrepancy model is not conducive to early identification and intervention. This is based on the assumption that few struggling children meet requirements for the IQ-Achievement discrepancy until about third grade; hence, young children tend to fail for two or three years without receiving services (Restori, Gresham, & Cook, 2008). Thus, as schools wait for a wide gap between actual and expected performance to develop, the student's academic needs are not met during an interval when intervention may effectively decrease the achievement gap between a child's and his or her peers' reading performance (Vaughn & Fuchs, 2003).

A second argument is that there are various formulas and definitions across the states that are used to obtain an IQ-achievement discrepancy (D. Fuchs, Mock, Morgan, & Young, 2003). As a result, these varying definitions have led to unreliable and inconsistent identification rates of students with SLD between states. Further, many researchers argue that the use of one data point taken at one point in time also leads to poor reliability of decision making (Fletcher et al., 2002).

A third argument is that the use of this model makes it difficult to adhere to the exclusionary clause in IDEA that does not allow learning disabilities to be diagnosed when it is primarily the result of poor instruction or due to social, cultural, or language experience. Because this type of model assesses skills and characteristics associated with the child, other environmental factors influencing poor academic performance (e.g., whether or not children have received appropriate instruction, curriculum issues, etc.) may not be evaluated. As result of these complications, studies indicate that over identification and under identification frequently occur (Gresham, 2007; Speece & Shekitka, 2002; Vaughn, 2003).

A fourth argument against the use of the discrepancy model is that assumptions that would support the utility of the IQ-achievement discrepancy model have not been supported. For example, academic performance of poor readers with a discrepancy does not differ from that of students without a discrepancy (Gresham, 2002). Several studies suggest that young, poor readers with and without an IQ-achievement discrepancy show similar phonological-processing deficits and perform similarly on many reading-related tasks (e.g., Foorman, Francis, & Fletcher, 1995; Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Stanovich, 1999; Torgesen, Morgan, & Davis, 1992; Vellutino et al., 2006). Further, there is a lack of emprical support on the utility for selecting instructional techniques that differentially benefit students with LD as opposed to students without a LD that also may struggle with reading (Gresham & Witt, 1997).

Given the continuing dissatisfaction with the discrepancy outcomes, the 2004 reauthorization of the IDEA no longer required states to use the discrepancy approach to

identify of students with learning disabilities (L. Fuchs & D. Fuchs, 2007). Alternatively, IDEA 2004 specifically states that a child's response to a series of increasingly more intensive interventions can be used if empirically supported interventions are administered to a student (Jimerson, Burns, & VanDerHeyden, 2007). This option is an approach called RTI. The following section discusses the RTI model, its potential to improve the process of identification and intervention in the school, and surrounding empirical support.

## Proposed RTI Models in the Literature

The RTI approach to identifying learning disabilities essentially proposes that the extent to which a child does or does not respond to a high-quality intervention that is implemented with integrity is a good predictor of future performance support. Further, observed inadequate gains with an intervention that incorporates effective teaching provides current data to help rule out prior poor instruction as a primary factor for low achievement. Continued low performance with several levels of more intensive intervention supports provides additional data that suggests that a student is struggling due to a learning disability (Healy et al., 2005).

In general, RTI approaches incorporate research based programs and practices for intervention planning and validated curriculum based evaluation measures for decision making (L. Fuchs & D. Fuchs, 1998). The majority of RTI models incorporate a multiple tiered intervention structure. Tier 1 is the universal intervention which consists of the general education classroom. When this level of intervention is effective, most students

(approximately 80%) perform as expected. Screening procedures or progress monitoring of student improvement occurs several times a year to evaluate if the program meets this criterion indicating that the general education program is effective for most students. This screening data is further used to identify students who are not responding to the general education universal program. Any identified struggling students in Tier 1 (i.e., approximately 20%) are provided with Tier 2 intervention to supplement the general education program with more intensive instruction. Tier 2 can be more intensive by giving more instructional time, more practice with corrective feedback and conducted with small groups of students to decrease the teacher and student ratio. In Tier 2, progress monitoring is more frequent (e.g., once a month) to make decisions about the effectiveness of the intervention for group and student within a short time frame. Progress monitoring data of Tier 2 intervention is used to identify students who continue to struggle with the intervention methods in Tier 2 and may require more intensive Tier 3 level of intervention. In Tier 3, students receive even more intense intervention, usually individualized, with more frequent progress monitoring (e.g., every two weeks) or students are referred to special education services (L. Fuchs & D. Fuchs, 2007; Healy et al., 2005; O'Connor et al., 2005; VanDerHeyden, Witt, & Gilberston, 2007b).

An historical difficulty with multiple tiers may be the creation of multiple tiers of preventative intervention that can be reliably distinguished from both general and special education (L. Fuchs & D. Fuchs, 2007). Despite the neatness of the traditional RTI model in the literature, there is still debate on whether or not special education services should occur after Tier II has failed, if Tier 3services should include special education, and if

special education should be a fourth tier for students who do not respond to Tier 1, 2 or 3 (Barnes & Harlacher, 2008). Typically, the different opinions on this matter follow views on what needs to be done after Tier 2 has proven ineffective (Reschly, 2005). The question of how many tiers are needed in an effective RTI model depends on what one's primary purpose for this model entails (D. Fuchs, Stecker, & L.Fuchs, 2008, p. 73). That is, proponents of "fewer tiers" (i.e. those supporting Tier 3 as special education) view RTI's primary purpose as disability identification whereas those who support "more tiers" see the main goal of RTI as early intervention (D. Fuchs et al., 2008, p. 74). Due to this on-going debate, RTI models can vary significantly from school to school. Thus far, there is not enough research to determine which, if any, of these approaches consistently yields the greatest improvement for the highest percentage of students. The following section summarizes the current research supporting the advantages of the implementing a RTI model with at least a Tier 2 level in place in a school setting.

## RTI Empirical Support

In the past, systematic allocation of resources to provide various levels of supplemental instructional support has not been consistently implemented. The RTI model has a decision making process that is maintained continually through progress-monitoring of all students to allocate the right amount of support that is aligned with specified instructional outcomes at various levels of support (Compton, Fuchs, Fuchs, & Bryant, 2006). By systematically using progress monitoring to identify struggling students to aid in systematic academic program planning, the tiered approach of RTI

models aims to help all students achieve better academic outcomes. Thus, this type of program encompasses services for many more students than those who identified with a disability (Barnes & Harlacher, 2008).

Although the research is still emerging on RTI programs, there have been a few studies whose results suggest that following the RTI program increases reading performance over time and helps a large amount of students exit an "at-risk" status. For example, Simmons and colleagues (2008) conducted a longitudinal four year study on the effects of an RTI program on changeability of risk status for reading difficulties of students attending seven Title I elementary schools from kindergarten through third grade. In the kindergarten year, prior to any intervention, students were given early reading curriculum-based measures (letter-naming frequency [LNF] task and initial sounds fluency [ISF]), the word identification subtest of the Woodcock-Johnson Psychoeducational Battery and the Peabody Picture Vocabulary Test. Students were classified within an at-risk status for poor reading performance if a student score fell below the 30$^{th}$ percentile on the LNF task, in the bottom quartile on the ISF tests, and if poor performance was confirmed by their teachers. All identified at-risk students received additional support that consisted of small group interventions given 5 days a week for between 30 and 45 minutes. Although 117 (25%) of the original 464 kindergarten students screened were classified as being at risk, there were only 41 of these students who had completed the study at the end of third grade. Of the remaining students, at the end of the study, performance scores increased to above the 46$^{th}$ percentile on a phonemic segmentation fluency assessment and at or above the 63$^{rd}$ percentile for a nonsense word

fluency assessment at the end of the study. Further, 38 of the 41 (93%) children were classified as out of risk (above 30th percentile) on the word identification task, 39 out of 41 (95%) were out of risk on the word attack task, and 20 out of 41 (49%) students were out of risk on the task of oral reading fluency.

Results from the study suggested that the use of an RTI model and early intervention is an effective way to help a relatively high percentage of students who are initially seen as "at-risk" attain and maintain a status of "out-of-risk." The traditional RTI model suggests that an ideal percentage of students considered out-of-risk is 85%. Interestingly, the percentages for the word identification and word attack tasks (93% and 95%, respectively) correspond to higher than expected percentages of out-of-risk students than the RTI model suggests. In contrast, the percentage of students out-of-risk based on the oral reading fluency task (49%) corresponds to a lower than expected percentage than the RTI model suggests.

Even though results are very promising, there are some limitations of the study that warrant continuing research in this area. First, the high attrition may have resulted in a biased sample of students who performed differently than those students who did not remain in the study. Also, lack of a control group made it difficult to ascertain how many students would have fallen within the at-risk status without exposure to the RTI program. Finally, instead of actively problem solving and allowing for flexible grouping throughout the year, the study only assessed individual student placements at the beginning and end of the school year.

An important emphasis of RTI is a strong focus on the identification and provision of effective interventions when problems first emerge in addition to the identification of students with SLD (Restori et al., 2008). O'Connor and colleagues (2005), for example, conducted a study examining the effectiveness of a three tiered RTI program on remediating reading difficulties for kindergarten through third-grade students. In this study, the researchers collected reading outcome data from kindergarten through first grade the first year of the study and added a grade level to data collection each subsequent year of the 3-year study. Outcome performance was compared between at-risk students (those who were not achieving set benchmark performance standards) participating in an RTI program and students in a control group who did not participate in the RTI. Tier 2 intervention in this study consisted of approximately 15 minutes of supplemental instruction administered to about three students three times a week. Tier 3 intervention consisted of 30 minutes daily, individual or small group instruction. Results indicated that the students who received tiered intervention made significantly more progress than students who did not receive intervention with an effect size of 0.6 between Tier 2 and the control group performance. Further, with the use of Tier 2 and Tier 3 interventions, fewer students were identified as needing special education (control = 15%, Tiers 2 and 3 = 8%).

Data presented by Vellutino and colleagues (2006) also suggested that the use of RTI effectively distinguishes responders from nonresponders and helps struggling readers to reach and maintain near normal levels of reading. In this study, 3% of students participating in a RTI program whereas 9% of students who did not participate in the RTI

program were performing below the 30th percentile on the Woodcock-Johnson test of Reading Mastery after one semester. Similarly, a study carried out by Torgesen and colleagues (2001) found that 31.4% of elementary school students already known to have a learning disability effectively responded to an RTI program. In sum, several studies have concluded that a systematic RTI Tier 2 and Tier 3 interventions in first through third grades improves the ability of school personnel to identify and help struggling students. Further, studies indicate that a RTI program can have a substantial impact on student performance in a relatively short amount of time and decrease the number of at-risk students for reading problems.

Several studies have also suggested that use of a RTI approach reduces over representation of minorities in special education. Marston, Muyskens, Lau, and Canter (2003), for example, reviewed the effects of a problem solving RTI model used by Minneapolis Public Schools on overrepresentation of minority students in special education. In this problem solving approach, the first stage (or tier) focuses on problem solving to help teacher support students in the general education curriculum. The second stage uses a multidisciplinary team in order to further problem solve and create interventions for struggling students. Students who are not showing sufficient response in either of the first two stages are moved into the third stage with more intensive problem solving and evaluation. In the year prior to implementation of the model 44% of the student population was African American, however 64% of referrals, 69% of students evaluated, and 69% of those found eligible for special education were African American students. Four years later, after the problem solving model had been in place 45% of the

students were African American and 59% of referrals, 58% of students evaluated, and 55% of students eligible for special education were African American students. Further, although data from the Harvard Civil Rights Project reported that the odds-ratio that an African American in Minnesota is labeled as LD was 2.7, the Minneapolis Public Schools odds-ratio for African American students ranged from 1.9 to 2.1 after 3 years of implementation of the RTI model.

As a result of emerging support for an RTI approach, a survey of department of education representatives from all 50 states conducted by Berkeley, Bender, Peaster, and Saunders (2009) indicated that many schools across the nation have developed and are implementing an RTI model. These models vary between schools but in general are based on two approaches: a standard protocol approach and/or a problem solving model. A standard protocol employs one empirically support program to address reading problems for all children within a tier whereas as problem solving approach uses a four step problem solving steps to identify the type of the reading problem and develop an individualized intervention to address the student needs (D. Fuchs et al., 2003). Most schools reported using a blended approach using a standardized protocol for Tier 2 and problem solving at Tier 3.

Although different RTI models are used, there is a general consensus on key components that are needed for any type of RTI to be effective. As part of their definition of RTI, Burns and Gibbons (2008) noted four components that are essential: methodical use of assessment data, effective allocation of resources, enhanced learning, and applicability to all children. L. Fuchs and D. Fuchs (2006) further identified integrity,

feasibility, and efficacy as other key components of any RTI model. Integrity is generally thought of as monitoring the implementation of interventions to check for compliance to standard protocols and consistency across providers (Glover & DiPerna, 2007). Questions surrounding feasibility issues facilitate proper use of resources and help measure costs of training, progress monitoring, and implementing interventions (L. Fuchs & D. Fuchs, 2006). The efficacy of RTI involves the level of responsiveness of students and how well interventions seem to be functioning at each tier.

One important feature of all RTI models is flexibility, or the ability of students to enter to more intense tiers immediately when reading difficulties emerge and to exit to a less intense tier when problems are remediated (Burns & Gibbons, 2008). Students experiencing severe problems may be expedited directly to the most insensitive tier level or special education. Flexible movement is based on the progress monitoring data.

While group flexing has not been thoroughly researched, what research has been done has resulted in positive findings (Glover & DiPerna, 2007; Healy et al., 2005). Vaughn, Linan-Thompson, and Hickman (2003), for example, conducted a study that incorporated flexible grouping into a RTI model. Forty-five second-grade students who were identified as at risk for reading problems were provided with supplemental instruction in groups of three for 35 minutes daily. Supplemental instruction focused on five components of reading: phonemic awareness, phonics, fluency, reading comprehension, and spelling. Progress monitoring measures were administered every 10 weeks, and students who met exit criteria no longer participated in supplemental instruction. Exit criteria involved a student receiving a passing score on a screening

measure (Texas Primary Reading Inventory), a reading comprehension task, and a score of 50 correct words per minute on a second-grade level fluency task for three consecutive weeks. At the end of the study, the students were grouped into four groups: early exit (10 students who exited after 10 weeks), mid exit (14 students who left after 20 weeks), late exit (10 students who left after 30 weeks), and no exit. Twenty-two of the 24 students who exited early and at mid point continued to make adequate progress without the supplemental instruction and 23 of these students maintained fluency reading scores receiving only core instruction. However, approximately 34% of students, consisting mainly of the students in the late exit group, were not able to achieve even minimal progress in the remaining weeks after supplemental instruction was removed. These results indicate that the appropriate use of progress monitoring and the decision making process increases the likelihood that students will receive the correct amount of support for their needs, and will be given the opportunity to return to solely the general education classroom when sufficient progress has not only been reached but maintained. Further, the results demonstrate the necessity for progress monitoring and evaluation to continue even after a student has left supplemental instruction. However, few studies investigated the long term effects of tiered, flexible grouping or progress of students who exit a Tier 2 or Tier 3 program. Moreover, additional research on when and how frequently students stop making adequate progress and need to return to more intensive intervention and characteristics of these students is needed.

Given that RTI models have been successfully conducted with well-funded and researcher-supported projects, recent studies have further investigated whether RTI

methods can be implemented by school districts without this type of support and get similar results. A meta-analysis of the current studies ranging between 1996 and 2004 that was conducted by Burns, Apppleton, and Stenhouer (2005) compared outcomes of studies of RTI programs developed and implemented in an applied setting with and without support from a research team. An important finding in this meta-analysis was that RTI models implemented without support from a university research team had larger median effect sizes than those implemented by university research teams (1.42 and .92, respectively). This research also found supporting evidence for the assumption that an RTI actively in use leads to improved student and system outcomes. Compared to research developed RTI, field based RTI resulted in positive systemic improvements three times greater with effect sizes of 1.80 and .47. Further, this meta-analysis also found that as a result of RTI, only 1.26% of the student population was referred for special education services and on average only 1.68% of student populations were placed into special education. Additionally, of the 6% of the student population within RTI, less than 2% received special education placement indicating that approximately just over 4% of the student population benefitted from the implementation of RTI.

In sum, the promising results from current RTI outcome research provide a promising framework for timely service delivery for all students (Fletcher & Vaughn, 2009). Research has suggested various models are effective such as the problem solving (PS) and the standard protocol (SP) model (Marston, 2005), but what interventions are utilized and how they are systematically implemented varies across schools (Berkeley et al., 2009). Thus, there are many unanswered questions regarding RTI programs in the

applied setting that have yet to be resolved (Reynolds & Shaywitz, 2009).

For example, there is still some debate on such issues as how many tiers are needed, the most beneficial types of interventions as well as the ideal intensity and most effective duration of chosen interventions, what intervention in each tier should look like, how long children need to be in each tier and how and when do children move between tiers, and on whether or not Tier 3 needs to be distinct and separate from Special Education placement (Vaughn, 2003). Reynolds and Shaywitz (2009) posed that current concepts of RTI may place too much emphasis on the quality and level of instruction in student achievement levels without operationally defining how "appropriate" instructional levels would or should be determined (Reynolds & Shaywitz, 2009). Swanson (2008) also stated that the lack of consideration for environmental and individual differences that can affect outcomes is a major concern regarding RTI.

Additionally, more investigation needs to be done to determine the utility of different assessment tools and different decision making criteria with the multiple tiered models as well as on both individual and aggregated outcomes related to tier specific interventions (Glover & DiPerna, 2007). Burns and Ysseldyke (2005) also reported some questions that research needs to more fully answer regarding the RTI process based on a meta-analysis of the RTI literature. These questions include whether or not adequate personnel training is occurring, whether or not all interventions being used are empirically supported, if there are certain leadership components that lead to success and the utility of RTI data for accurate identification of students with learning disabilities, and how integrity should be assessed. Reynolds and Shaywitz (2009) purported that the

current data on student outcomes that guide the critical components of the RTI model in practice is too scarce and without comparisons with other potential models to make clear conclusions on whether or not RTI approach is the best service delivery model or classification process for students with disabilities.

Another major criticism of RTI according to Reynolds and Shaywitz (2009) related to RTI's apparent alteration of the definition of learning disability. As a result of this, there may be a large portion of students who are not reaching their full potentials and who would benefit from intervention and accommodations that are not able to access this help because they do not necessarily fall into a higher Tier of intervention services indicating a clear need for additional support. The authors agree with previous researchers that the current RTI model may actually be defining the term learning disabled simply as a group of low-achieving students who did not respond to good instruction and who may be performing at grade level on some measures, but would actually show severe discrepancies if given more comprehensive assessments (Boada, Riddle, & Pennington, 2008; Swanson, 2008).

Despite growing empirical support for use of an RTI approach, clearly additional investigations are warranted to further our understanding of the RTI process in the applied school setting. Reynolds and Shaywitz (2009) reported an insufficient research base to give reliable guidance on how to implement an effective RTI model because many aspects need elaboration and clear definitions. Given the number of uncertain aspects of RTI models and wide variance in specific RTI practices, it is important that data are collected and used to make decisions on whether or not a RTI program is

resulting in positive student outcomes. Monitoring student outcomes using a valid, high utility assessment tool for decision making at all levels is especially warranted given the complexity of the process and many unknown factors that support an effective RTI program (Glover & DiPerna, 2007). The following section reviews the current literature on progress monitoring tools that may be used in the RTI system.

## RTI and Progress Monitoring

Given the limitations of RTI research, frequent progress monitoring (PM) of student outcomes is the critical feature to ensure that each school's RTI model results in positive student outcomes. When examining student PM data, educators need to determine if (a) instruction is supporting most students in each tier, (b) each tier is serving approximately the right percentage of students to adequately allocate school resources to match student needs, (c) students are not given higher levels of services unless their performance indicates the need for these services, and (d) students showing consistently inadequate growth are given higher levels of services. Each of these four outcomes indicates that an RTI program is producing the expected results in an efficient manner. Once it is determined that a program is not working, then educators can work on determining why and what aspects of the model need to be implemented, supported or modified. To increase the efficacy of the RTI process, it is critical that PM data are used to determine when a program is not working as expected in a reasonable period of time.

The technical features of a useful PM tool require that a measurement system can be administered on frequent regular intervals without the influence of practice effects, be

sensitive to individual responsiveness within a short period of time, easy to administer, and reliably predict important outcomes (e.g., proficient reading in later grades, scores on end of year tests (Francis et al., 2008). One well-researched measure that meets these criteria is curriculum-based measurement (CBM; Shinn, 1989; Shin & Bamonto, 1998). To date, CBM has been extensively used to evaluate RTI programs in outcome studies. The following section reviews the current literature on the psychometrics and utility of CBM as a useful progress monitoring tool in the RTI system.

## Empirical Support for CBM

CBM was developed to be an inexpensive, quick and simple method to frequently index student performance in key academic areas (Burns & Gibbons, 2008; Deno, 2003). CBM administration and scoring procedures have been standardized to provide a reliable measurement system to routinely monitor and evaluate individual student progress and instructional effectiveness (Shapiro, 1996; Shinn, 1989). When administering a reading CBM (R-CBM), for example, a child is asked to read aloud for one minute while an observer records oral-reading fluency (ORF). A child's performance, or decision making metric, is gauged on the number of correctly read words per minute. Individual student level of ORF performance and growth over time (slope) is formatively compared to benchmarks or levels of peer performance and peer growth rates to assess a student's response to intervention (L. Fuchs & D. Fuchs, 1998).

R-CBM has been well–researched and found to have adequate psychometric properties (Burns & Gibbons, 2008; Daniel, 2010; Marston, 1989; McMaster & Espin,

2007; Wayman et al., 2007). Several studies report test-retest reliability of R-CBM measures in reading to range between .82-.97 and alternate-form reliability to range between .84-.96 (Wayman et al., 2007). According to L. Fuchs, D. Fuchs, and Speece (2002), CBM measures in reading have demonstrated good criterion validity when compared with popular commercial reading tests with correlations between ORF and test cores ranging between .63 and .90. Hintze and Silberglitt (2005) reported moderate predictive validity (range, $r$ = .49 to $r$ = .61) and concurrent validity ($r$ =.69) between R-CBM and the Minnesota Comprehensive Assessment. Hosp and Fuchs (2005) also conducted a construct validity study to determine if the relationship between CBM and specific reading skills remained strong at different grade levels. First through fourth graders were administered grade-appropriate CBM reading passages as well as subtests from the Woodcock Reading Mastery Tests. Two to 3 weeks after completing the subtests, the CBM reading measures were readministered. Results showed criterion related validity correlations between CBM and the word identification, word attack, basic skills and passage comprehension subtests to range between .71 and .91. Further, a discriminative analysis indicated that R-CBM score correctly identified adequate performing students on each of the specific reading skills and identified students who scored below a standard score of 90 at Grade 1 and 85 at Grades 2 through 4 on each reading sub skills on the standardized test. Results from this study provide further evidence that R-CBM is appropriate for monitoring performance on specific reading sub skills, such as decoding, word reading, and comprehension. Finally, several studies have also verified that CBM performance is sensitive to growth in reading over brief periods of

time (Marston, Fuchs, & Deno, 1986). This sensitivity allows teachers to be able to continually collect and review data to identify when instruction is or is not working for each student (Burns & Gibbons, 2008; Shinn & Bamonto, 1998). R-CBM data can be utilized in multiple ways including helping to establish standards for screening and identifying students who potentially need Special Education services, monitoring, planning, and determining the effectiveness of educational programs given to all students (Stecker, Fuchs, & Fuchs, 2005).

Despite the research demonstrating potential benefits of R-CBM, the limitations of this assessment tool should be noted. One major concern is the variability in student ORF data (level and slope data) that may be due to differences in text difficulty in grade CBM reading passages rather than student response to intervention. Difficulty in using readability formulas to identify grade level reading passages that are equivalent with respect to difficulty level reduces the utility of the CBM assessment as a measure of student progress (Ardoin, Suldo, Witt, Aldrich, & McDonald, 2005). Generalization of R-CBM across students is also limited given that R-CBM studies have been conducted with primarily with elementary students with few studies conducted with students of diverse backgrounds. Moreover, initial studies have been conducted on the utility of R-CBM measures for evaluating the effects of instruction for students with disabilities while studies investigating the utility of R-CBM for the purposes of establishing normative levels and growth rates to make high-stakes decisions within the RTI model, although promising, is just emerging in the literature.

In sum, despite its limitations, research has provided support for the use of CBM

as one of the few progress monitoring tools that serves as a decent screening tool or indicator of student performance and progress in reading (Deno, 1985; Stecker et al., 2005). Moreover, its sensitivity to student growth within a short period of time provides a valuable tool for monitoring interventions designed to maximize student learning. When schools effectively use CBM, progress monitoring can be on a frequent basis in order to problem solve when reading difficulties first emerge. Early effective problem solving increases a child's chance of obtaining the necessary help that he or she needs in order to succeed in school. Unfortunately, there is also concern about the type and duration of training teachers and school personnel need in order to be effective in interpreting CBM data (Wayman et al., 2007). Just how well teachers and school personnel understand CBM and how to use it is still unclear. If teachers are interpreting the same data in different ways, then its generalizable utility could be questioned. Research and practice of CBM has been done for many years, yet it appears to still be underutilized as a tool to help decision making (Wayman et al., 2007). If CBM is being underutilized and misunderstood, it may mean that RTI is also being underutilized and misunderstood (Shinn, 2007).

Since the primary use of CBM is to monitor student progress, it is a necessary component to help ensure the effective use of the RTI model. With teachers having a good understanding of CBM, it is more likely that the data obtained will be used to help in the decision making process. Thus, more research may need to be conducted on the utility of CBM for various decision making conducted by educators within a RTI program.

## Purpose of Research

An alternative method for systematically providing reading support that matches student needs that has gained popularity and empirical support is RTI. According to this model, the extent to which a child does or does not respond to a high quality intervention that is implemented with integrity is a good predictor of future performance and need for support (Healy et al., 2005). The RTI model typically follows a two- or three-tiered approach with children in higher tiers receiving more intervention support. Ideally, RTI replaces the IQ-academic performance discrepancy approach by improving reading before problems become extreme (L. Fuchs & D. Fuchs, 2007; Healy et al., 2005; O'Connor et al., 2005; VanDerHeyden et al., 2007b).

Even though RTI has an increasingly amount of empirical support, there are still areas that are lacking research. RTI involves a lot of components that are necessary for its proper implementation and there is not considerable support on what those components are or how they should be implemented (Vaughn, 2003). Thus, once a school commits to implementing various levels of instruction within an RTI approach, educators need to collect, analyze, and respond to information about student academic outcomes many times throughout a school year to ensure that a RTI program is functioning as expected. Given that the goal of the program is to result in many students performing as expected within the general education curriculum (~80%) or showing adequate growth with additional instructional support (~15%), tracking patterns of student outcomes is a major component of any RTI model to gauge achievement toward this RTI goal (Burns & Gibbons, 2008; VanDerHeyden et al., 2007b).

Progress monitoring is the valuable process within a RTI model to gather information about student outcomes patterns to assess the status of all intervention levels and to construct instructional programs based on the data. Progress monitoring involves the frequent evaluation of student performance to see if the RTI system is effective as expected and to help teachers and school personnel make appropriate decisions about students and curriculum (Burns & Gibbons, 2008). The most common form of student progress monitoring is CBM. CBM typically takes only a minute to administer and detects individual changes in skill level over a short period of time. The findings of prior research on CBM supports the utility of R-CBM data as a valid index to evaluate individual student progress, to predict student progress on standardized high stakes testing, and to guide decision making for instructional practices for group or individual students (Shinn, 1989; Shinn & Bamonto, 1998). Thus, student CBM data can be useful indicators of academic patterns to gauge the progress towards the RTI goal (i.e., 80% children are performing as expected in Tier 1 and 15% in Tier 2). Aggregates of student CBM outcome data that parallel intervention programs within an RTI model "targets" such as school wide, grade level, Tiers, and/or subgroups of students can be used as outcome measures of academic interventions at all intensity levels (Deno, 1985; Stecker et al., 2005; Wayman et al., 2007).

Most studies have drawn conclusions from CBM data collected and aggregated from small groups of students who are receiving some sort of instructional support within Tier level(s) but few have reported or analyzed patterns of collected and aggregated CBM data at several levels of an RTI program (Fewster & MacMillan, 2002; Wayman et al.,

2007). This makes it difficult to describe expected patterns of CBM performance, percentages of students within each tier, and group flexibility in the applied school setting. Findings from such an exploration of CBM patterns would further facilitate research on relations among CBM-related patterns and RTI programs targeting academic outcomes in schools. Therefore, this evaluation study aimed to explore the patterns based on real time CBM data and student Tier placement as a school based case example of the nature and utility of the CBM data within an RTI model in an applied setting. Three descriptive analyses of universal screening and CBM data were conducted in both third and fourth grade RTI programs within two schools during one school year (2009-2010). During the school year, CBM universal screening data was collected three times (fall, winter, spring) and the following analyses were conducted from that data:.

1. To what degree did school placement in the three tiers reflect placement using Aims benchmark percentiles?
2. Based on school placement, what percentages of students are being served within each tier across benchmark period?
3. Are students benefiting from instruction provided within each Tier based on the rates of improvement across benchmark periods?
4. Based on school placements, what percentage of students move to more or less intensive tiers between benchmark periods?
5. Are students benefitting from instruction following a tier movement placement based on the rates of improvement (in subsequent measurement periods)?

# CHAPTER III

# METHODS

## Participants and Setting

The participants for this study included all students in 3rd and 4th grade classrooms from two schools within a school district in an eastern state implementing a RTI program. Participating schools in this study implemented a RTI program during the 2009-2010 school year in the third and fourth grades and collected and sent progress monitoring data to the district curriculum administrators. Both schools utilized in this study are considered Title I schools with approximately 4% of the student population identified as English Language Learners and 8.5% identified as students with disabilities. Table 1 presents demographic information about each of the schools used for the study. Table 2 presents the percentages of students passing the end of year state English exam.

## General RTI Program Overview

Because specific details of programs between schools varied, specific aspects on the RTI model for each school were collected and reported through interviews. The interviews were conducted with one administrator from each of the two schools participating in this study to gather more specific information on how the RTI model looks and functions within their respective school. A summary of the interviews will be discussed next.

Both schools administered universal benchmarks three times per year in fall,

Table 1

*Demographic Data*

| | School 1 | | School 2 | |
|---|---|---|---|---|
| School demographics | *n* | % | *n* | % |
| Student | | | | |
| Total | 732 | | 662 | |
| Males | 374 | 51 | 358 | 54 |
| 3rd grade | 109 | | 99 | |
| 4th grade | 126 | | 99 | |
| Free/reduced lunch | 264 | 37 | 227 | 36 |
| Race | | | | |
| White | 365 | 50 | 315 | 47 |
| Black | 138 | 18 | 185 | 28 |
| Latino/Hispanic | 159 | 21 | 109 | 16 |
| Native American | 1 | <1 | 1 | <1 |
| Asian | 24 | 3 | 20 | 3 |
| Other (unspecified, Pacific Islander) | 45 | 6 | 32 | 5 |
| Student/teacher ratio | 19/1 | | 17/1 | |
| Classroom teachers | | | | |
| % fully licensed teachers | | 82 | | 91 |
| % provisional teachers | | 18 | | 9 |
| % w/bachelor's | | 47 | | 51 |
| % w/masters | | 51 | | 49 |
| % w/doctorate | | 2 | | 0 |
| % teaching outside endorsement area | | 9 | | 10 |

Table 2

*Percentage of Students Passing State English Exam*

| | School 1 | | | School 2 | | |
|---|---|---|---|---|---|---|
| Grade | Total | 3rd | 4th | Total | 3rd | 4th |
| All students | 84 | 88 | 79 | 86 | 79 | 91 |
| Female | | 92 | 82 | | 77 | 94 |
| Male | | 83 | 77 | | 80 | 89 |
| Black | 75 | 81 | 68 | 77 | 66 | 86 |
| Hispanic | 79 | 87 | 72 | 94 | 86 | 100 |
| White | 90 | 92 | 87 | 86 | 84 | 88 |
| Students w/disabilities | 81 | 73 | 80 | 62 | Not enough data | Not enough data |
| Economically disadvantaged | 70 | 79 | 63 | 80 | 68 | 93 |
| Limited English proficiency | 78 | 94 | 64 | 88 | 88 | 92 |

winter, spring. In both School 1 and 2, the teachers were primarily in charge of administering the universal benchmarks. Additionally, both schools utilized their building reading specialists and county personnel to conduct training. For the 2009-2010 year, School 1and 2 teachers were trained on AIMSweb. School 2 teachers also received practical training through observations and feedback. In addition, School 2 conducted a "data walk" to discuss administration and data collected (including interpretation) from benchmarks.

Within Tier 1, there was no specific reading program in School 1 or 2. Individual teachers determined what lessons to teach or follow. And in both School 1 and 2, while teachers were primarily in charge of administering the Tier 1 program, when possible they were supported through paraprofessionals in the classroom. At School 1, the daily reading lessons were conducted for 2 hours of "language arts" a day during which 45 minutes is set aside as "guided reading" time where the class is broken into small groups for more individualized instruction. School 2 had 90 minutes set aside every day for reading and language arts. For both schools, a general day of language arts/reading included both whole group and small group instruction. Whole group activities vary from day to day depending on student needs. Frequently, students have different centers to work on writing, fluency and comprehension. Small group instruction typically separates students by their level, or tier. Teachers gained information about student strengths and weaknesses through administrations of such measures as the Qualitative Reading Inventory (QRI) and the Reading Level Indicator (RLI).

In Tiers 2 and 3, progress monitoring occurred more frequently in both schools.

School 2 indicated that progress monitoring for both T2 and T3 occur on a weekly basis and in School 1 progress monitoring for T2 occurs weekly to every other week while T3 progress monitoring occurs one to two times a week. In both schools, classroom teachers and/or paraprofessionals administered progress monitoring measures in Tier 2 and the reading specialist or another member of the special education staff administers these measures to Tier 3 students.

In regards to instructors, both schools reported Tier 2 interventions were implemented by classroom teachers and Tier 3 interventions were primarily implemented by special education teachers or the reading specialist. Instruction in Tier 2 and 3 differed between the two schools. In School 1, Tier 2 and Tier 3 students received 45 minutes of direct instruction or intervention within the classroom through time imbedded in the schedule. Tier 2 students received 30 minutes of additional help 4 to 5 days a week during "intervention or extension" (IE) time. During IE time, Tier 3 students were pulled out by the reading specialist or special education teacher for even more targeted help. School 2 Tier 2 and Tier 3 students receive Tier 1 instruction plus additional services. Tier 2 students were pulled every other day during IE time (for a total of 45 minutes each time) to receive additional help. Tier 3 students were pulled 30 minutes every day for more individualized interventions.

In both schools, Tier 2 and Tier 3 interventions were as individualized as possible for the students in each group and materials and practice can look different on any given day. In School 1, Tier 2 groups typically had 6-9 students during IE time and between 10 and 15 students for intervention time within the classroom. Tier 3 groups in School 1

typically had six to eight students. In School 2, Tier 2 was comprised of four to five students and Tier 3 was comprised of two to three students. Similar to T1, T2 and T3 did not have any assigned or specific program used for intervention in either school. Both schools allowed teachers flexibility to use and tweak what works for them and their specific class. School 1 also distributed the book "Strategies that Work" by Stephanie Harvey and Anne Goudvia to give teachers ideas for lessons as well as strategies for implementation. School 2 reported the following programs as commonly being used in some form or fashion: PALS (Phonological Awareness Literacy Screening) lessons; Florida Reading Research lessons (Florida Center for Reading Research), SOAR (Set goals; Organize, Ask Questions, Record Progress) to Success (Houghton Mifflin) strategies; Houghton Mifflin "Intervention Kit" and Intervention Reading Series; and LLI (Leveled Literacy Intervention).

Specific skills were also targeted. School 1 indicated they first tried to target specific strategies students need: visualize, look backs, inferences, text connection, synthesize, and comprehension. Additionally, to target fluency, School 1 practiced sight words. Similar to School 1, School 2 also tried to target specific student needs. Specifically, School 2 used results from reading inventories to target specific student deficiencies. Both schools tried to incorporate and teach essential components of reading: phonological awareness, phonics, fluency, vocabulary and comprehension.

The two schools also reported the decision making and data analysis process. Both School 1 and School 2 used AIMSweb determined percentile scores from AIMSweb student aggregates to determine what Tier a student should be placed in. Specifically, the

10th and 25th ORF percentiles benchmark scores from AIMS web student aggregates per grade level were used for decision-making. These percentiles were based on information gathered from the AIMSWeb computer based program the district paid to access and was then provided to the primary researcher. As presented in Table 3, ORF scores that fell below 10th, between 10th and 25th, and above 25th ORF percentiles were placed in Tier 3, Tier 2, and Tier 1 level of supports, respectively.

Additionally, School 1 also took into consideration a student's placement on a Leveled Literacy program. In addition to AIMSweb cut-off scores, School 2 also used teacher input in determining tier placement; if a student falls on the low end but within an acceptable Tier 1 progress range and a teacher feels he or she needs additional help, that student can be placed in Tier 2 instead of remaining in Tier 1.

Both schools heavily relied on benchmark and PM data to determine when a student needs a change in placement. Common tools used for benchmarks and PM were the previously discussed measures: AIMSweb probes, RLI and RAI results, and PALS

Table 3

*AIMS Criterion for Oral Reading Fluency Level*

| Level/status | Percentile | Fall | Winter | Spring |
|---|---|---|---|---|
| Third grade three assessment periods per year | | | | |
| Tier 3 | 1 to 10 | 0 - 33 | 0 - 49 | 0 - 63 |
| Tier 2 | 11 to 25 | 34-53 | 50 - 78 | 64 - 91 |
| Tier 1 | >25 | > 54 | >78 | > 91 |
| Fourth grade Three Assessment Periods Per Year | | | | |
| Tier 3 | 1 to 10 | 0 - 54 | 0 - 69 | 0 - 81 |
| Tier 2 | 11 to 25 | 55 - 78 | 70 - 95 | 82- 105 |
| Tier 1 | >25 | > 78 | > 95 | > 105 |

data. Each school has a data or RTI team that made the decisions about student placement within each tier. At School 1, this team was comprised of ESL teacher, general education teacher, reading specialist, technology specialist, principal, and assistant principal. At School 2, the data team consisted of the general education teachers, reading specialist, school psychologist, educational diagnostician, assistant principal, and any other specialist the team feels is needed (e.g., ELL or special education teacher). At School 1, each grade level also had 1 meeting a month to review the data with the reading specialist to discuss data and progress of students. Additionally, grade level teachers met on a weekly basis to discuss data and the entire team met once every other month for half a day to discuss data and placement. School 2 had a similar schedule for meeting to discuss data. Once a month each grade level met on their own to discuss data and once a month the entire team met to discuss data and tier placement. School 2 also indicated that teachers and specialists met informally as frequently as needed to discuss interventions and placement.

Each teacher at School 1 kept a "running record" for each student and class that includes benchmark and PM scores as well as analysis of areas of strength and weakness. These data are discussed each time the team meets. School 1 stated that in general, students transition between tiers on a quarterly basis. As a general guideline at School 1, after a student has three or four data points on his/her current leveled reading that are 95% accurate or above, movement to a less intensive tier is considered. Similarly, if a student's data points consecutively (3-4 times) show low or decreasing scores, transition to a more intensive tier is considered.

At School 2, student data was discussed at each meeting to determine if a change in placement needs to occur. As a general guideline, if the team notices a student achieving above his/her goal line for six to seven consecutive data points (PM measures) or if a student is flat lining or falling under his/her goal line for six to seven consecutive data points movement is considered. Additionally, when the team was discussing a student who may need more intensive interventions, consideration of what may be the cause for the flat line or decrease is also taken into account. The team tried to determine if a student is not responding to a specific intervention or if the instruction or means of instruction needs to change.

## Materials

The benchmark reading passages used by each school came directly from the AIMSweb program (AIMSWEB). This assessment program is based on curriculum-based measurement standardized procedures (Shinn, 1998) to monitor student progress in early literacy skill development. As part of this program for third and fourth grade students, progress monitoring reading passages are to be administered three times per year (fall, winter, spring) to an entire student population to monitor school wide and district wide progress The AIMSweb program provides reading passages for each of the three benchmark administrations conducted throughout a school year and in addition a variety of additional passages that can be used throughout the year as progress monitoring measures. There are 23 passages for first grade (3 benchmark and 20 progress monitoring) and 33 passages (3 benchmark and 30 progress monitoring) for grades two

through eight. Third grade passages were 300 words each and fourth grade passages were 350 words each. Average reliability (alternate form) for third-grade passages was .86 and for fourth-grade passages was .85. Readability correlations were also calculated and compared using a variety of readability formulas including Fry (.94), Lexile (.97) and Spache (.97). Median readability was calculated to be .95 for the probes used specifically for benchmark testing and .90 for progress monitoring passages (Howe & Shinn, 2002).

## Dependent Variable

Student intervention progress was evaluated by measures of oral fluency rates (ORF), which were assessed with AIMSweb third and fourth grade reading passages. Oral reading rate (ORF) was determined by the number of correct words read per minute (WRC) from three administered reading passages. Standardized directions, as described by AIMSweb (Shinn & Shinn, 2002) were given by the examiners requesting students to read aloud from three different reading passages for one minute while the examiner followed along on a separate copy of the passage. A correctly read word was defined as an unprompted word that was read aloud by the student with correct pronunciation in 3 seconds. If a student did not attempt to read a word within 3 seconds, the examiner read the word for the student and marked the word as an error. Words were also counted as an error if the student omitted or mispronounced the word in the passage. ORF was calculated by subtracting incorrect words from the total words read. Total ORF was the median of the three administered passages (Shinn & Shinn, 2002). Average WRC for third-grade passages was found to be 107.6 and 121.5 for fourth-grade passages (Howe &

Shinn, 2002). Correlations between ORF and standardized comprehension measures of reading range from *r* = .63 to *r* = .90 (Shinn, 1989). In addition, ORF in general measures have been reported to have high test-retest reliability estimates (range, *r* = .92-.97) and alternate-form reliability estimates (range, *r* = .89-.94; Baker & Good, 1995; Good, Simmons, Kame'enui, Kaminski, & Wallin, 2002b; Shinn, 1989). As indicated in the previous section, alternate form reliability for AIMSweb reading passages ranged from *r* = .80 to *r* = .90. Standard error of measurement for AIMSweb passages were consistent with previously published studies (Howe & Shinn, 2002) Criterion-related validity studies of ORF probes ranged from .52 to .91 from eight separate validity studies (Good & Jefferson, 1998, Good & Kaminski, 1998; Good, Gruba, & Kaminski, 2002a).

## Overview of Progress Monitoring Data Collection Procedures

Three reading probes were administered to students three times during the study school year in fall, winter, and spring. All schools scheduled personnel to administer the CBM within the same week as the administration times are determined by the district and administration is done district-wide. School personnel were trained by the district administration to follow AIMSweb standardized procedures.

## District RTI Progress Monitoring Database System

Participating schools reported ORF data to the administration by completing and submitting an online data summary for each student via the AIMSweb database system. Through the computer-based data system that the school district paid to access, educators

enter individual data to track individual, class, school, and/or district level progress. Schools reported the ORF data that is collected at least three times a school year (fall, winter, spring) with each third- and fourth-grade student participating in the school-wide reading program. Student names were linked to the entered data at this site; however, the data were exported by district personnel and attached student codes rather than student names to the exported data. The data provided through this database was used in this study to answer the research questions.

## Procedure

This study utilized an archival data collection method. After obtaining written permission from the district research review board to use data from the district program evaluation database and approval for procedures from the Utah State University human subjects review board, researchers were provided with a de-identified dataset that has been obtained directly from each school's AIMSweb results for the 2009-2010 school year. All of this data was coded and entered into an Excel file.

Additionally, as stated earlier, key personnel at each school were interviewed to find out more specific information about the RTI model and process within their specific school (see the Appendix for list of questions).

# CHAPTER IV

# RESULTS

Data were collected to explore the degree an RTI system is working in an applied setting. Descriptive statistics were calculated and used as the primary method of analysis to answer the research questions. First a descriptive summary of the collected data will be presented. Next, results are presented to answer each of the research questions.

## Descriptive Summary of Data

The mean and standard deviation of ORF level and slope for the overall sample as well as for each individual school and each grade level within each school at fall, winter, and spring progress monitoring administrations are presented in Tables 4 and 5. To determine slope, student-level data were compared using oral reading growth rates between fall, winter, and spring screening assessment times. Specifically, for each individual, the winter benchmark was subtracted from the fall benchmark to determine the Time 2 growth trajectory benchmark. Likewise, for each individual, the spring benchmark was subtracted from the Winter benchmarks to determine the Time 3 growth trajectory benchmark.

In general, mean scores at both schools and in both grades increased across time although the mean score for School 1 was greater than that for School 2 on most assessments. Mean slope scores also show that School 1 had more improvement in ORF than School 2 and that children in grade 3 showed slightly greater growth than children in grade 4. Finally, the mean slope for both schools indicated that the mean growth slowed

Table 4

*Mean Oral Reading Fluency Across Time at Each Grade Level Per School*

| | Fall | | | Winter | | | Spring | | |
|---|---|---|---|---|---|---|---|---|---|
| Grade level | Total | School 1 | School 2 | Total | School 1 | School 2 | Total | School 1 | School 2 |
| Grade 3 | | | | | | | | | |
| *N* | 207 | 107 | 100 | 206 | 107 | 99 | 208 | 109 | 99 |
| Mean | 85.3 | 89.1 | 81.2 | 108.1 | 113.8 | 102.0 | 121.1 | 129.6 | 111.7 |
| *SD* | 39.2 | 42.3 | 35.3 | 39.1 | 42.2 | 34.6 | 43.7 | 46.7 | 38.2 |
| Minimum | 5 | 5 | 5 | 8 | 11 | 8 | 8 | 13 | 8 |
| Maximum | 183 | 183 | 164 | 202 | 199 | 202 | 254 | 254 | 203 |
| Grade 4 | | | | | | | | | |
| *N* | 222 | 125 | 97 | 224 | 125 | 99 | 223 | 124 | 99 |
| Mean | 105.8 | 104.1 | 107.8 | 125.1 | 127.0 | 122.7 | 137.7 | 139.5 | 135.5 |
| *SD* | 33.1 | 36.3 | 28.4 | 33.6 | 37.7 | 27.5 | 34.6 | 38.5 | 29.1 |
| Minimum | 14 | 14 | 46 | 0 | 0 | 62 | 26 | 26 | 70 |
| Maximum | 206 | 206 | 184 | 230 | 230 | 200 | 242 | 242 | 208 |

Table 5

*Mean Slope for Oral Reading Fluency Across Time at Each Grade Level Per School*

| | Winter | | | Spring | | |
|---|---|---|---|---|---|---|
| Grade level | Total | School 1 | School 2 | Total | School 1 | School 2 |
| Grade 3 | | | | | | |
| *N* | 199 | 103 | 96 | 202 | 106 | 96 |
| Mean | 23.0 | 25.5 | 20.4 | 12.9 | 15.8 | 9.7 |
| *SD* | 16.2 | 17.0 | 14.9 | 14.9 | 15.7 | 13.3 |
| Minimum | -24 | -15 | -24 | -27 | -16 | -27 |
| Maximum | 98 | 98 | 53 | 95 | 95 | 40 |
| Grade 4 | | | | | | |
| *N* | 217 | 122 | 95 | 216 | 120 | 96 |
| Mean | 19.8 | 23.0 | 15.7 | 13.8 | 13.8 | 13.7 |
| *SD* | 16.1 | 18.3 | 11.6 | 12.2 | 12.9 | 11.3 |
| Minimum | -57 | -57 | -16 | -16 | -16 | -9 |
| Maximum | 98 | 98 | 58 | 67 | 67 | 53 |

down between winter/spring as compared to fall/winter in both grades.

## Research Questions

Research Question 1: *To what degree did school placement in the three tiers reflect placement using AIMS benchmark percentiles?*

Given that the schools reported use of other data with ORF results to make tier placement decisions, the first research question examined the extent that school placement in the three tiers reflected placement using AIMSweb 10$^{th}$ and 25$^{th}$ ORF percentiles benchmark scores. The number and percentages of students in each Tier based on school and AIMS ORF placement categories are presented in Table 6. As presented in Table 6, placement of students in tier based on school criteria were almost identical to number of students placed in a tier based solely on AIMS benchmarks.

Research Question 2: *Based on school placement, what percentages of students are being served within each tier across benchmark period?*

The second research question examined the extent the standard RTI Tier student placement percentages (80%, 15%, and 5%) at Time 1, 2, and 3 (fall, winter, and spring) are reflected in the percentages of students in each tier as reported by the school As presented in Table 6, the percentages of students as placed by school personnel closely matched the RTI standard student population percentages (80-15-5), although slightly less so in Tier 1.

Research Question 3. *Are students benefiting from instruction provided within each tier based on the rates of improvement across benchmark periods?*

Table 6

*Percentages of Students in School and ORF-Level-Determined Tiers*

| | School | | ORF | |
|---|---|---|---|---|
| Level | Frequency | Percent | Frequency | Percent |
| Fall | | | | |
| Tier 1 | | | | |
| Total | 326 | 76.30 | 328 | 76.50 |
| School 1 | 169 | 73.20 | 170 | 73.30 |
| School 2 | 157 | 80.10 | 158 | 75.20 |
| Tier 2 | | | | |
| Total | 72 | 16.90 | 71 | 16.60 |
| School 1 | 44 | 19.00 | 44 | 19.00 |
| School 2 | 28 | 14.30 | 27 | 12.90 |
| Tier 3 | | | | |
| Total | 29 | 6.80 | 30 | 7.00 |
| School 1 | 18 | 7.80 | 18 | 7.40 |
| School 2 | 11 | 5.60 | 12 | 5.70 |
| Winter | | | | |
| Tier 1 | | | | |
| Total | 348 | 80.90 | 348 | 80.90 |
| School 1 | 190 | 81.90 | 190 | 81.90 |
| School 2 | 158 | 79.80 | 158 | 79.80 |
| Tier 2 | | | | |
| Total | 56 | 13.00 | 56 | 13.00 |
| School 1 | 26 | 11.20 | 26 | 11.20 |
| School 2 | 30 | 15.20 | 30 | 15.20 |
| Tier 3 | | | | |
| Total | 26 | 6.00 | 26 | 6.00 |
| School 1 | 16 | 6.90 | 16 | 6.90 |
| School 2 | 10 | 5.10 | 10 | 5.10 |
| Spring | | | | |
| Tier 1 | | | | |
| Total | 338 | 78.60 | 342 | 79.40 |
| School 1 | 183 | 78.90 | 185 | 79.40 |
| School 2 | 155 | 78.30 | 157 | 79.30 |
| Tier 2 | | | | |
| Total | 59 | 13.70 | 56 | 13.00 |
| School 1 | 30 | 12.90 | 29 | 12.40 |
| School 2 | 29 | 14.60 | 27 | 13.60 |
| Tier 3 | | | | |
| Total | 33 | 7.70 | 33 | 7.70 |
| School 1 | 19 | 8.20 | 19 | 8.20 |
| School 2 | 14 | 7.10 | 14 | 7.10 |

The percentages of students who showed adequate progress rate within each tier program over time was also explored to examine the degree that students were benefiting from instructions provided within each tier. Adequate growth was defined in this study as meeting a slope criterion that was calculated for the winter and spring assessments. Slope criterion was calculated from the AIMSweb percentile score that was the benchmark score for Tier 1 that was shown in Table 3. The Winter slope criterion was calculated as the difference in AIMS Tier one benchmark scores between winter and fall (i.e., 25 for third grade; 17 for fourth grade) and between spring and winter (i.e., 13 for third grade; 10 for fourth grade). Each student was coded as obtaining adequate growth if his or her slope met or was above the grade level calculated slope criterion score and inadequate if below the slope criterion.

The percentages of students who showed adequate growth are presented in Table 7. These results reveal that when examining growth rates, the percentages of students showing adequate growth were lower in the more intensive tier levels and all percentages of students responding were less than the expected 80% of students that should be responding within a tier.

Research Question 4: *Based on school placements, what percentage of students moved to more or less intensive tiers between benchmark periods?*

To explore the degree of movement of student placement between tiers, the number of students who remained in the same tier as well as those that entered or exited tiers between fall and winter and between winter and spring were examined. An analysis of student movement between Tiers 1, 2, and 3 of the RTI model showed that

Table 7

*Percent of Student Showing Adequate Growth in School-Determined Tiers*

| Level | Winter ORF % | Spring ORF % |
|---|---|---|
| Tier 1 | | |
| Total | 55 | 62 |
| School 1 | 63 | 63 |
| School 2 | 53 | 45 |
| Grade 3 | 52 | 57 |
| School 1 | 55 | 67 |
| School 2 | 49 | 44 |
| Grade 4 | 58 | 67 |
| School 1 | 70 | 67 |
| School 2 | 43 | 67 |
| Tier 2 | | |
| Total | 53 | 37 |
| School 1 | 42 | 58 |
| School 2 | 52 | 48 |
| Grade 3 | 46 | 23 |
| School 1 | 50 | 11 |
| School 2 | 43 | 39 |
| Grade 4 | 59 | 50 |
| School 1 | 63 | 42 |
| School 2 | 54 | 67 |
| Tier 3 | | |
| Total | 12 | 19 |
| School 1 | 40 | 19 |
| School 2 | 30 | 0 |
| Grade 3 | 11 | 18 |
| School 1 | 18 | 17 |
| School 2 | 0 | 20 |
| Grade 4 | 14 | 22 |
| School 1 | 20 | 33 |
| School 2 | 0 | 0 |

approximately 22% of the student population moved up or down a tier during the school year with School and ORF level based decisions. Of the 96 school decision movements, 44 were positive and 52 were negative. As presented in Table 8, there were approximately 50 student placement changes to less intensive tiers ($n$ = 52) and 50 to

Table 8

*Movement Between Tiers*

| Level | Placement | Frequency | *SD* |
|---|---|---|---|
| Tier 1 | Students moved to Tier 2 | | |
| | Winter | | |
| | Total | 11 | 2.66 |
| | Third grade | 4 | 1.85 |
| | Fourth grade | 7 | 3.54 |
| | Spring | | |
| | Total | 17 | 4.10 |
| | Third grade | 7 | 3.26 |
| | Fourth grade | 10 | 4.95 |
| Tier 2 | Students moved to Tier 1 | | |
| | Winter | | |
| | Total | 30 | .725 |
| | Third grade | 12 | 6.06 |
| | Fourth grade | 18 | 8.33 |
| | Spring | | |
| | Total | 9 | 2.16 |
| | Third grade | 2 | 0.99 |
| | Fourth grade | 7 | 3.26 |
| | Students moved to Tier 3 | | |
| | Winter | | |
| | Total | 8 | 1.93 |
| | Third grade | 4 | 1.85 |
| | Fourth grade | 4 | 2.02 |
| | Spring | | |
| | Total | 8 | 1.92 |
| | Third grade | 4 | 1.86 |
| | Fourth grade | 4 | 1.98 |
| Tier 3 | Students moved to Tier 2 | | |
| | Winter | | |
| | Total | 10 | 2.42 |
| | Third grade | 2 | 1.01 |
| | Fourth grade | 8 | 3.70 |
| | Spring | | |
| | Total | 3 | 0.72 |
| | Third grade | 1 | 0.50 |
| | Fourth grade | 2 | 0.93 |

more intensive tier (*n* = 54) during the school year. More students moved from Tier 2 back into Tier 1 (*n* = 39) than from Tier 3 to Tier 2 (*n* = 13). In addition, more students moved from Tier 1into Tier 2 (*n* = 28) than from Tier 2 to Tier 3 (*n* = 16).

Research Question 5: *Are students benefitting from instruction following a tier movement placement based on the rates of improvement (in subsequent progress monitoring assessments)?*

Outcomes of students who moved between tiers in the winter were examined by reviewing the student rate of growth and benchmark level in the spring assessment. Results presented in Table 9 show that 50 to 66% of all the students across grades who moved to more intense tiers (Tier 1 to 2 or Tier 2 to 3) showed adequate slope; however, less than 30% met benchmark in the spring assessment. Additionally, there were between 42% and 30% of students who moved to a less intensive tier (Tier 3 to 2 or Tier 2 to 1) who had an adequate slope in the spring. Seventy-eight percent of students in Tier 1 who had moved from Tier 2 met the spring benchmark whereas 10% of students in Tier 2 who had moved from Tier 3 met the spring benchmark.

Table 9

*Spring Outcome of Students Moved Between Tiers in Winter*

| | | Adequate slope | | Met benchmark | |
|---|---|---|---|---|---|
| Spring tier | Movement after winter | Frequency | % | Frequency | % |
| Tier 1 | Students who moved to Tier 1 from Tier 2 | | | | |
| | Total | 14 | 43.75 | 25 | 78.13 |
| | Third grade | 5 | 41.67 | 9 | 75.00 |
| | Fourth grade | 9 | 56.25 | 16 | 100.00 |
| Tier 2 | Students who moved to Tier 2 from Tier 1 | | | | |
| | Total | 5 | 50.00 | 3 | 30.00 |
| | Third grade | 1 | 16.67 | 0 | 0.00 |
| | Fourth grade | 4 | 100.00 | 3 | 75.00 |
| | Students moved to Tier 2 from Tier 3 | | | | |
| | Total | 3 | 30.00 | 1 | 10.00 |
| | Third grade | 1 | 50.00 | 0 | 0.00 |
| | Fourth grade | 2 | 25.00 | 1 | 12.50 |
| Tier 3 | Students who moved to Tier 3 from tier 2 | | | | |
| | Total | 4 | 66.67 | 1 | 16.67 |
| | Third grade | 2 | 66.67 | 0 | 0.00 |
| | Fourth grade | 2 | 66.67 | 1 | 33.33 |

# CHAPTER V

# DISCUSSION

Students must be able to fluently read in order to learn academic content throughout their school years. To support student reading outcomes, many schools are now implementing a RTI approach to provide systematic tiers of instructional options and a continuous monitoring system to identify deficits when they first emerge and attempt to remediate deficits before they become severe enough to warrant special education services. Although the research is still emerging on the effectiveness of RTI, results of studies examining academic outcomes of RTI programs show that reading performance increases for a higher percentage of student population within a school implementing RTI versus a school that is not implementing an RTI program (Berkeley et al., 2009; Marston et al., 2003; O'Connor et al., 2005; Restori et al., 2008; Simmons et al., 2008; Vellutino et al., 1996).

Reviewing student responses to intervention programs on a regular basis is a required element of RTI to enable educators to make timely decisions about student intervention needs throughout the academic year (Burns & Gibbons, 2008). CBM is currently the most researched continuous measurement system that provides useful progress monitoring information about the effects of reading programs and interventions in schools on reading performance (Christ, Scullin, Tolbize, & Jiban, 2008; Shapiro, 1996; Stecker et al., 2005). Patterns of student populations based on CBM data throughout a school year in a school setting implementing a RTI program for reading were explored in this study in terms of tier placements and tier effectiveness. The primary

purpose of the descriptive research reported in this article was to better understand the RTI process in an applied setting regarding decisions that are made based on CBM measures to determine if: (a) each tier is serving a reasonable percentage of students, (b) each tier is enhancing progress for most students (i.e., 80% or more), and (c) students are moved in and out of tier levels following review of CBM data, and (d) students are benefitting from instruction following a tier movement placement. This study reports on the progress of RTI program in two schools.

## Percentage of Students Served in Each Tier

Several researchers (Burns et al., 2010; Burns & Gibbons, 2008; D. Fuchs et al., 2004) stress that effective allocation of resources is an essential part of the success of the RTI process. Given the allocation of available time and teacher resources in school settings, advocates of RTI model propose that 80% of children within a school to be adequately progressing with the universal general education curriculum (Tier 1), 15% of children who may be at-risk for reading problems to be given more intensive small group instructional program (Tier 2) and 5% of students at high risk or not responding to Tier 2 students to be given more individualized instruction (Tier 3). These results in this present study showed that when ORF AIMS percentiles were used to determine tier placement, distribution of students more closely matched the standard 80%-15%-5% of the typical RTI model. The ORF percentages of student in each tieri matched placements by school personal although both schools reported using other data sources. Any school may use multiple formative assessments to monitor progress and make decisions. There are

generally two types of formative assessments that may be used: general outcome measures (GOM) and subskill mastery measurements (SMM). A GOM is a probe or progress monitoring measure that looks at multiple skills at a time (Christ & Vining, 2006). Measure of oral reading rate is a GOM that is also considered a strong predictor of performance on yearly reading achievement tests (Foegen & Deno, 2001; L. Fuchs & Deno, 1991). Measures of performance on a single skill, such as high frequency words or sounds, are considered SMM assessments. SMM measures reveal more rapid growth to instruction given within a short period of time relative to a GOM measure but are not predictive of student performance on multiple skill performance tests such as end of year tests. The AIMs ORF measure examined in this study is considered a GOM, thus, we expected other assessments to be considered in decision making in addition to the oral reading rates. In the two schools utilized in this study, additional sources or assessments utilized in making decisions on top of progress monitoring probe data included RAI and RLI results (reading level indicators), PALS data as well as input from teachers and other team members including school psychologist, principal, ESL teacher and educational diagnostician. However, the lack of variance between ORF levels and school placement suggests that the schools were substantially relying on AIMS outcome data or the additional data substantiated AIMS data for tier placement decisions. Some plausible explanatiosn in this finding is the other data may not have been as sensitive to change as CBM data, as valued by team members, or showed similar outcomes.

## Growth Rates Across Time and Grades

Overall, the mean ORF scores at both schools and in both grades showed an increase in school across time, but mean growth slowed down between Winter/Spring and Fall/Winter assessments. The results obtained through this research also showed differences in third- and fourth-grade growth with third graders tending to demonstrate more growth than fourth graders. This slowed pattern of growth was similar to a lower pattern in growth rates on CBM progress monitoring data in a study conducted by L. Fuchs and D. Fuchs (1993), which also examined patterns of student growth across time. Oral reading fluency rates of first through sixth-grade students ($n$ = 117) were monitored each week throughout one school year from October through April. Weekly monitoring revealed that progress for most students could be described as linearly increasing with time. Additionally, this study concluded that the magnitude, or rate of improvement tended to decrease gradually over the year and in more advanced grade levels. In study with a larger sample, Christ, Silberglitt, Yeo, and Cormier (2010) showed a similar steady decrease in weekly ORF growth based on AIMS reading probes between lower and upper grades from a large sample of general and special education students ($N$ = 3,808) from second to sixth grades. Mean performance levels as well as decrease in growth with increase in grade level obtained in this study were similar to the CBM data presented by Christ and colleagues. Moreover, results from both studies show that the ORF growth rates decrease between fall and spring. Christ and colleagues suggested the possibility that changes in teacher expectations and classroom behavior management may alter academic time and instructional intensity across the school year such that students

are making less progress in spring (Christ et al., 2010). More research needs to be done to determine possible reasons for this observed pattern; however one possible reason for this may be that younger children tend to grow more and faster than older children. Additionally, this pattern has the possibility of being due to differences in difficulty of probes, variability of probes, and differences in content overlap between the various probes for the different age groups. Other plausible explanations may include differences in emphasis in general curriculum between the two grades (i.e., focusing more on vocabulary development versus fluency versus comprehension), more difficult skills, change in instructional intensity or a focus on end of year testing. Although more complex reading skills are addressed each year, growth rates may be influenced by acquisition of basic skills in early months and years (e.g., decoding skills) followed by attention on more complicated aspects of learning (e.g., comprehension, vocabulary) . Further, oral passage reading is a direct measure of early skills such as decoding and fluency and may not be as sensitive to growth in more complex comprehension skills. Regardless of the reason why scores decline, it is important to note that this decline in ORF rates may also influence decisions about program effectiveness and student movement between tiers. If a decline in ORF rates is common in elementary school, future research may also identify potential strategies in each tier to maintain the higher Fall learning rates.

Another interesting finding in this study that warrants discussion is the percentage of students showing adequate growth within each tier in winter and spring when examining growth as the difference in AIMS 26$^{th}$ percentiles score between fall and

winter or between winter and spring benchmark data. When examining percentages of students showing adeqaute progress within each tier based on slope data in this study, there were lower percentages of students responding than the expected 80% and there were lower percentage of students showing adequate growth in the more intensive tiers. Based on this criterion, there were less than 63% of students responding in each tier in both winter and spring and the percentage of students identified as responding decreased with increased tier intensity. Moreover, the percentage of students responding improved in Spring in Tier 1 and 3 and decreased in Tier 2. This suggests, similar to conclusions from Burns and colleagues (2005), that the method educators utilize to determine student placement can affect decision making, which therefore can also influence need for program instructional changes or the amount of support a student may or may not receive. Thus, the method chosen to determine student need would impact the distribution of resources, time or program changes in a given school. Given the effect the method used to determine progress can have on a student's educational progress, as well as on the distribution of resources within a school, more research needs to be done to determine most effective and appropriate methods.

While the percentage of student responding was low, it is important to note that there is limited data on what is the best method to determine a particular student's placement (Johnson, Jenkins, Petscher & Catts, 2009). Some research results have indicated that using primarily a student's reading level is an effective method, while others indicated the addition of student's growth rate, or slope is a more accurate or useful measure to make decisions regarding screening and progress outcomes (Burns et

al., 2010; D. Fuchs et al., 2004). Based on a review of the literature on the utility of CBM as a decision-making tool, Ball and Christ (2012) conclude that the sensitivity to evaluate the instructional effects depends on the purpose of the assessment. Although results are mixed when reviewing individual data, there is some evidence (Ardian et al., 2005) that CBM reading data are sufficiently sensitive enough to evaluate whether or not a program is resulting in adequate change in reading fluency after a 3 month period of instruction. Burns and colleagues (2005), however, found that a comparison of decisions made with an aimline and a dual discrepancy method (i.e., review of slope and level) resulted in different decisions on program responsiveness for 40% of the students. This slope criterion was used to judge tier program effectiveness given that this is an easy calculation and comparison benchmark that schools can use to judge program outcomes. Future research may further investigate if these results are typical due to an increase in reading difficulty of probes or that different instructional strategies may increase growth.

## Flexible Movement of Student Tier Placement

As noted above, review of RTI data also allows decisions about the flexible movement in and out of instructional tiers for students with and without a disability (Burns & Gibbons, 2008). Consideration of tier movement can occur at any time data is reviewed during a school year. And given that data is typically reviewed at least three times a year to evaluate outcomes of the general education program and is reviewed more frequently for more intensive tiers, a student may be moved multiple times throughout the

school year. While group flexing has not been thoroughly researched, what research has been done has resulted in positive findings (Glover & DiPerna, 2007; Healy et al., 2005). Ideally, students are given higher or lower levels of services students when their performance indicates the need for a change in services. Within a successful Tier program, it would be expected that most children would demonstrate a quick rate of improvement towards benchmark level or show improved reading rates with additional Tier support while few others would struggle. Another goal of RTI is to decrease the time students wait to receive appropriate instruction. For example, severe deficits may require immediate movement from Tier 1 to Tier 3 to provide most appropriate instruction in a timely (Lemons et al., 2010). Frequent monitoring and flexible movement between tiers allows placement children in the most appropriate instructional programs at the right time throughout a school year. However, little is known about the degree that movement between tiers may impact the tiered instructional process (Compton et al., 2006). Further, there are few empirically supported guidelines in the literature on the timing of this type of decision making, duration of intervention or data that warrants change, and frequency of movements that produces optimal good outcomes.

Despite the potential for multiple movements between tiers throughout the school year, a pattern noted with this data set was that regardless of method used, there was a need for some of movement of students moving students to a lower or more intensive tier placement throughout the year. Additionally, results obtained in this study indicated that while flexible movement between tiers is beneficial for some students, this movement alone is not sufficient in helping students achieve desired growth; additional means of

determining and providing needed support are necessary. Plausible reasons why movement may occur between tiers in the applied settings may include resources, logistics, and perception that more time is needed to maintain gains.

## Limitations

There are several limitations of this study which in the future with additional studies can hopefully eliminated. One limitation is the small sample size used for the study. This study only included data from two schools and only 3rd and 4th grade. It may be that obtained results would have been different with a greater sample size and with a larger span of grades. Additionally, both schools utilized were Title I. Overall positive results were obtained, however it may be that schools that do not receive additional money and support as a result of being Title I may not achieve the same positive results due to more limited resources. Future research should look to explore these questions in non-Title I schools.

A second limitation of this study is that it only examined one school year. It may be that results would have been different if looked at data across years. This would allow an opportunity to see what changes schools and school systems make over time and what impact, if any, these changes have on student growth and improvement.

A third limitation was the use of archival data, which makes it slightly difficult to assess precise implementation of the CBM probes and decision making procedures used to place students. In the future, it would be interesting to collect and utilize data during the current school year to better assess program characteristics and integrity of

implementation.

Another limitation in this study is the difficulty in accessing needed information through the AIMSWeb website. While the benchmarks and percentiles used in this study were gathered directly from the AIMSWeb website based on information given to this researcher by district personnel, detailed information on source or rationale was limited due to the nature of the AIMSWeb system. As AIMSWeb is a paid service, much of the resources and information can only be accessed with payment and AIMSWeb has only published limited resources on their system that is available for public use. An outside person looking for information on percentiles, benchmark scores and rationale for decisions is unable to obtain this information without paying the AIMSWeb service. Again, this researcher did not have full access to the AIMSWeb system in order to better understand information received from district personnel.

Finally, this study did not examine how frequently data should be analyzed within each tier. Our results only touched the surface of this question and showed that while some frequent analysis and reassignment may be occurring, few students are being moved between tiers. Future studies are needed to determine ideal frequency of reassignments or readjustments that would lead to more optimal growth for the largest number of students. Additionally, as there currently are no clear guidelines on how long and when to change, future research would benefit from exploring this avenue further.

## Practical Implications

There are several practical implications for the educational field based on the

results of this study. For example, based on these results, younger children tend to show more growth than older children which speaks to the definite importance and need for early intervention. This is evidenced by the trend that third-grade student ORF increases were on average greater than fourth grade student ORF increases. If we can continue finding better ways to analyze and monitor not only need, but growth early, the more likely we will be able to optimally help students during times of "prime" ability to learn and grow. Another practical implication of this research is showing that having a system of intervention and monitoring does help provide support to aid growth and success for struggling students. A structured system helps educators gain a better understanding and insight into the appropriate allocation of precious educational resources. Further, this research suggests that school personnel are appropriately using data within a RTI model, including appropriate collection and use of data to determine effectiveness of instruction and interventions and student changing instructional reading needs. Data must not only be gathered and analyzed, but it must be reviewed and utilized on a frequent basis to modify all aspects of service delivery, including intervention type and frequency, frequency of movement between tiers, and allocation of resources including time, money and personnel.

In sum, while the current research project results indicate that RTI is a very promising model to help maximize student growth, there are also many areas where research needs to continue. Many interesting patterns were found in this study including noting that student placement within the tiers of RTI depends on method chosen to evaluate along with particular criterion chosen. This also indicates however, that more

research needs to continue to determine precisely what the most appropriate methods and criterion are. Further, this research indicates that flexible grouping, or movement between tiers, has the potential to help the most students possible, however how frequently movement should be evaluated still needs to be determined. Finally, this project further supported previous research noting that allocation of resources and a systematic and structured delivery method of RTI are important for success. Overall, with continued research and practice, an implemented RTI model will undoubtedly help many students learn and succeed in one of the most basic and essential areas of learning: reading.

## REFERENCES


Ardoin, S., Suldo, S., Witt, J., Aldrich, S., & McDonald, E. (2005). Accuracy of readability of estimates' predictions of CBM performance. *School Psychology Quarterly, 20*(1), 1-22.

Aud, S., Hussar, W., & Kenna, G. (2011). *The condition of education 2011: National Center for Educational Statistics*. Retrieved from http://nces.ed.gov/pubs2011/2011033.pdf

Baker, S. K., & Good, R. H. (1995). Curriculum-based measurement of English reading with bilingual Hispanic students: A validation study with second-grade students. *School Psychology Review, 24,* 561-578.

Ball, C., & Christ, T. (2012). Supporting valid decision making: Uses and misues of assessment data within the context of RTI. *Psychology in the Schools, 49,* 231-244.

Barnes, A. C., & Harlacher, J. E. (2008). Clearing the confusion: Response-to-intervention as a set of principles. *Education & Treatment of Children, 31,* 417-431.

Berkeley, S., Bender, W. N., Peaster, L. G., & Saunders, L. (2009). Implementation of response to intervention: A snapshot of progress. *Journal of Learning Disabilities, 42(1),* 85-95.

Boada, R., Riddle, M., & Pennington, B. (2008). Integrating science and practice in education. In E. Fletcher-Jazen & C. R. Reynolds (Eds.), *Neuropsychological perspectives on learning disabilities in the era of RTI: Recommendations for diagnosis and intervention* (pp. 179-191). New York, NY: Wiley.

Burns, M. K., Apppleton, J. J., & Stenhouer, J. D. (2005). Meta-analytic review of responsiveness to intervention research: Examining field-based and research-implemented models. *Journal of Psychoeducational Assessment, 23,* 381-394.

Burns, M. K., & Gibbons, K. A. (2008). *Implementing response to intervention in elementary and secondary schools*. New York, NY: Taylor & Francis Group.

Burns, M. K., Scholin, S., Kosciolek, S., & Livingston, J. (2010). Reliability of decision-making frameworks for response to intervention for reading. *Journal of Psychoeducational Assessment, 28,* 102-114.

Burns, M. K., & Ysseldyke, J. E. (2005). Comparison of existing response-to-intervention models to identify and answer implementation questions. *The California School Psychologist, 10,* 9-20.

Christ, T. J., Scullin, S., Tolbize, A., & Jiban, C. L. (2008). Implications of recent research curriculum-based measurement of math computation. *Assessment for Effective Intervention, 33,*198-205.

Christ, T. J., Silberglitt, B., Yeo, S., & Cormier, D. (2010). Curriculum-based measurement of oral reading: An evaluation of growth rates and seasonal effects among students served in general and special education. *School Psychology Review, 39,* 447-462.

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

Compton, D., Fuchs, D., Fuchs, L., & Bryant, J. (2006). *Integrating research with improved policies and practices.* Retrieved from http://www.nrcld.org/about/presentations/2006/NASPSelecting.pdf

Daniel, M. H. (2010). *Reliability of AIMSweb Reading Curriculum-Based Measurement (R-CBM) (Oral Reading Fluency).* Upper Saddle River, NJ: Pearson.

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

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

Fewster, S., & MacMillan, P. D. (2002). School-based evidence for the validity of curriculum-based measurement of reading and writing. *Remedial & Special Education, 23,* 149-156.

Fletcher, J., Lyon, G., Barnes, M., Stuebing, K., Francis, D., Olson, R., … Shaywitz, B. A. (2002). Classification of learning disabilities: An evidence-based evaluation. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 467-519). Mahwah, NJ: Erlbaum.

Fletcher, J. M., & Vaughn, S. (2009). Response to intervention: Preventing and remediating academic difficulties. *Child Development Perspective, 3*(1), 30-37.

Foegen, A., & Deno, S. L. (2001). Progress monitoring measures in mathematics: A review of the literature. *The Journal of Special Education, 41,* 121-139.

Foorman, B. R., Francis, D. J., & Fletcher, J. M. (1995). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*(1), 37-55.

Francis, D. J., Shaywitz, S. E., Steubig, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology, 88*(1), 3-17.

Francis, D., Santi, K., Barr, C., Fletcher, J., Varisco, A., & Foorman, B. (2008). Form effects on the estimation of students' oral reading fluency using DIBELS. *Journal of School Psychology, 46,* 315-342.

Fuchs, D., & Fuchs, L. S. (2005). Peer-assisted learning strategies promoting word recognition, fluency, and reading comprehension in young children. *The Journal of Special Education*, *39*(1), 34-44.

Fuchs, D., Fuchs, L. S., & Compton, D. L. (2004). Identifying reading disabilities by responsiveness-to-instruction: Specifying measures and criteria. *Learning Disability Quarterly*, *27,* 216-227.

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research & Practice, 18*(3), 157-171.

Fuchs, D., Stecker, P., & Fuchs, L. S. (2008). Tier 3: Why special education must be the most intensive tier in a standards-driven No Child Left Behind world. In D. Fuchs, L. S. Fuchs, & S. Vaughn (Eds.), *Response to intervention: A framework for reading educators* (pp. 71-74). Newark, DE: International Reading Association.

Fuchs, L. S., & Deno, L. S. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional children*, *57*, 488-500.

Fuchs, L. S., & Fuchs, D. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22*(1), 1-30.

Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research & Practice, 13,* 204–219.

Fuchs, L. S., & Fuchs, D. (2006). A framework for building capacity for responsiveness to intervention. *School Psychology Review, 35,* 621-626.

Fuchs, L. S., & Fuchs, D. (2007). A model for implementing responsiveness to intervention. *Teaching Exceptional Children, 39*(5), 14-20.

Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly*, 33-45.

Gettinger, M., & Stoiber, K. (2007). Applying a response to intervention model for early literacy development in low-income children. *Topics in Early Childhood Special Education, 27,* 198-213.

Glover, T. A., & DiPerna, J. C. (2007). Service delivery for response to intervention: Core components and directions for future research. *School Psychology Review, 36,* 526-540.

Good, R., & Jefferson, G. (1998). Contemporary perspectives on curriculum-based measurement validity. In M. Shinn (Ed.), *Advanced application of curriculum-based measurement* (pp. 61-88). New York, NY: Guilford.

Good, R., & Kaminski, R. (1998). Assessing early literacy skills in a problem-solving model: Dynamic indicators of basic early literacy skills. In M. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 113-142). New York, NY: Guilford.

Good, R., Gruba, J., & Kaminski, R. (2002a). Best practices in using Dynamic Indicators of Basic Early Literacy Skills (DIBELS) in an outcomes-driven model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (4th ed., pp. 699-720). Bethesda, MD: National Association of School Psychology.

Good, R., Simmons, D., Kame'enui, E., Kaminski, R., & Wallin, J. (2002b). *Summary of decision rules for intensive, strategic, and benchmark instructional recommendations in kindergarten through third grade.* Eugene, OR: University of Oregon.

Gresham, F. M. (2002). Responsiveness to intervention: An alternative approach to the identification of learning disabilities. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Response to treatment* (pp. 467–519). Mahwah, NJ: Erlbaum.

Gresham, F. M. (2007). Evolution of the response-to-intervention concept: Empirical foundations and recent developments. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 10-24). New York, NY: Springer.

Gresham, F. M., & Witt, J. K. (1997). Utility of intelligence tests for treatment planning, classification, and placement decisions: Recent empirical findings and future directions. *School Psychology Quarterly, 12,* 249-267.

Healy, K., Vanderwood, M., & Edelston, D. (2005). Early literacy interventions for english language learners: Support for an RTI model. *The California School Psychologist, 10,* 55-63.

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.

Hosp, M. K., & Fuchs, L. S. (2005). Using CBM as an indicator of decoding, word reading, and comprehension: Do the relations change with grade? *School Psychology Review, 34*(1), 9-26.

Howe, K. B., & Shinn, M. M. (2002). *Standard reading assessment passages (RAPs) for use in general outcome measurement: A manual describing development and technical features.* Eden Prarie, MN: EdFormation.

Jenkins, J., & Terjeson, K. J. (2011). Monitoring reading growth: Goal setting, measurement frequency and methods of evaluation. *Learning Disabilities Research & Practice, 26*(1), 28-35.

Jimerson, S. R., Burns, M. K., & VanDerHeyden, A. M. (2007). Response to intervention at school: The science and practice of assessment and intervention. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The Science and practice of assessment and intervention* (pp. 3-9). New York, NY: Springer.

Johnson, E. S., Jenkins, J. R., Petscher, Y., & Catts, H. W. (2009). How can we improve the accuracy of screening instruments? *Learning Disabilities Research & Practice*, *24*(4), 174-185.

Kennedy Krieger Institute. (2005). *Learning disorders*. Retrieved from http://www.kennedykrieger.org/index.jsp

Lemons, C. J., Key, A. P., Fuchs, D., Yoder, P. J., Fuchs, L. S., Compton, D. L., . . . Bouton, B. (2010). Predicting reading growth with event-related potentials: Thinking differently about indexing "Responsiveness." *Learning & Individual Differences, 20,* 158-166.

Marston, D. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it. In M. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18-78). New York, NY: Guilford.

Marston, D. (2005). Tiers of intervention in responsiveness to intervention prevention outcomes and learning disabilities identification patterns. *Journal of Learning Disabilities*, *38*(6), 539-544.

Marston, D., Fuchs, L. S., & Deno, S. (1986). Measuring pupil progress: A comparison of standardized achievement tests and curriculum-related measures. *Diagnostique, 11,* 71-90.

Marston, D., Muyskens, P., Lau, M., & Canter, A. (2003). Problem-solving model for decision making with high-incidence disabilities: The Minneapolis experience. *Learning Disabilities Research & Practice, 18*(3), 187-200.

McMaster, K., & Espin, C. (2007). Technical features of curriculum-based measurement in writing: A literature review. *The Journal of Special Education, 41*(2), 68-84.

National Center for Education Statistics. (2010). *The Nation's Report Card: Reading 2009.* Washington, DC: Institute of Education Sciences.

O'Connor, R. E., Harty, K. R., & Fulmer, D. (2005). Tiers of intervention in kindergarten through third grade. *Journal of Learning Disabilities, 38,* 532-538.

Office of Special Education Programs. (2007). *Children with disabilities receiving special education under Part B of the Individuals with Disabilities Education Act*. Washington, DC: U.S. Department of Education.

Reschly, A. L. (2005). Learning disabilities identification primary intervention, secondary intervention, and then what? *Journal of Learning Disabilities*, *38*, 510-515.

Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology, 47,* 427-469.

Restori, A. F., Gresham, F. M., & Cook, C. R. (2008*). "Old habits die hard:" Past and current issues pertaining to response-to-intervention*. Retrieved from http://www.thefreelibrary.com/"Old habits die hard:" past and current issues pertaining to...-a0197363403

Reynolds, C. R., & Shaywitz, S. E. (2009). Response to intervention: Prevention and remediation, perhaps. Diagnosis, no. *Child Development Perspectives, 3*(1), 44-47.

Shapiro, E. (1996). *Curriculum-based measurement: Assessing special children* (2nd ed.). New York, NY: Guilford.

Shinn, M. R. (1989). Identifying and defining academic problems: CBM screening and eligibility procedures. In M. Shinn (Ed.*), Curriculum-based measurement: Assessing special children* (pp. 90-129). New York, NY: Guilford.

Shinn, M. R. (1998). Best practices in using curriculum-based measurement in a problem-solving model. *Best Practices in School Psychology IV*, *1*, 671-697.

Shinn, M. R. (2007). Identifying students at risk, monitoring performance, and determining eligibility within response to intervention: Research on educational need and benefit from academic intervention. *School Psychology Review, 36,* 601-617.

Shinn, M. R., & Bamonto, S. (1998). Advanced applications of curriculum-based measurement: "Big ideas" and avoiding confusion. In M. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 1-31). New York, NY: Guilford.

Shinn, M. R., & Shinn, M. M. (2002). *AIMSweb training workbook: Administration and scoring of Reading Curriculum-Based Measurement (R-CBM) for use in general outcome measurement*. Eden Prarie, MN: Edformation.

Simmons, D. C., Coyne, M. D., Kwok, O.-m., McDonagh, S., Harn, B. A., & Kame'enui, E. J. (2008). Indexing response to intervention: A longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilities, 41*(2), 158-173.

Speece, D. L., & Shekitka, L. (2002). How should reading disabilities be operationalized? A survey of experts. *Learning Disabilities Research & Practice*, *17*(2), 118-123.

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning.* Mahwah, NJ: Erlbaum.

Stecker, P. M., Fuchs, L. S., & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of research. *Psychology in the Schools, 42,* 795-819.

Swanson, H. L. (2008). Neuroscience and response to instruction (RTI): A complementary role. In E. Fletcher-Janzen & C. B. Reynolds (Eds.), *Neuropsychological perspectives on learning disabilities in the era of RTI: Recommendations for diagnosis and intervention* (pp. 28-53). New York, NY: Wiley.

Tilly, W. D. (2003, December). *Heartland Area Education Agency's evolution from four to three tiers: Our journey —Our results*. Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.

Torgesen, J., Alexander, A., Wagner, R., Rashotte, C., Voeller, K., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes for two instructional approaches. *Journal of Learning Disabilities, 34,* 33-58.

Torgesen, J. K., Morgan, S., & Davis, C. (1992). Effects of two types of phonological awareness training on word learning in kindergarten children. *Journal of Educational psychology*, *84*(3), 364-370.

VanDerHeyden, A. A., Snyder, P. A., Broussard, C., & Ramsdell, K. (2007a). Measuring response to early literacy intervention with preschoolers at risk. *Topics in Early Childhood Special Education, 27,* 232-249.

VanDerHeyden, A. M., Witt, J. C., & Gilberston, D. (2007b). A multi-year evaluation of the effects of a response to intervention (RTI) model on identification of children for special education. *Journal of School Psychology, 45,* 225-256.

Vaughn, S. (2003). *How many tiers are needed for response to intervention to achieve acceptable prevention outcomes?* Kansas City, MO: National Research Center on Learning Disabilities.

Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice, 18*(3), 137-146.

Vaughn, S., Linan-Thompson, S., & Hickman, P. (2003). Response to instruction as a means to identifying students with reading/learning disabilities. *Exceptional Children, 69,* 394-409.

Vaughn, S., Wanzek, J., Murray, C. S., Linan-Thompson, S., & Woodruff, A. L. (2009). Response to early reading intervention: Examining higher and lower responders. *Exceptional Children, 75*, 165-183.

Vellutino, F. R., Scanlon, D. M., Small, S., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *Journal of Learning Disabilities, 39*(2), 157-169.

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

# APPENDIX

## Interviews with Elementary School Administrators

Program Interview questions:

1. Tell me about the Tier 1 (general education) process.
    a. Who administers universal benchmarks and how are they trained?
    b. How often are Progress Monitoring measures administered in this tier?
    c. Is there a specific curriculum or program implemented? If so, what is it?
    d. Number of hours for reading for each grade? Number of days per week?
    e. What does a general day look like?
    f. Who implements the reading Tier 1 program (if anyone) with the teacher?

2. Tell me about the Tier 2/Tier 3 process.
    a. Who administer Progress Monitoring measures for these tiers and how are they trained?
    b. How often is data evaluated?
    c. Do you use a particular curriculum or program(s) for intervention?
    d. Number of hours per day? Number of days per week?
    e. What does a general intervention session look like? Groups and group size?
    f. Who implements the program?
    g. Intervention targeted skills?

3. Tell me about the decision making and data analysis process.
    a. How do you identify the students for each tier based on progress monitoring and benchmark data?
    b. How often does your team meet to review data and make decisions about progress?
    c. Who makes decisions?
    d. How do students transition between tiers (what is the process)?
    e. Any other measures used to evaluate and make decisions?

**School 1 Interview:**

***1. Tell me about the Tier 1 (general education) process.***

***a. Who administers universal benchmarks and how are they trained?***

The teachers; they are trained typically by Reading Specialists or sometimes through county. All teachers are trained on AIMSWeb and this year they were also trained on Fountus & Pinnell. Teachers have been trained through reading specialists/county/developer training on DRA (Developmental Reading Assessment from Scholastic)

Teachers learn strategies through “Strategies that Work”- book by Stephanie Harvey and Anne Goudvia; these books have lessons/guidelines in them as well

***b. How often are Progress Monitoring measures administered in this tier?***
3 times a year

***c. Is there a specific curriculum or program implemented? If so, what is it?***
Different grade levels use different things or a combination of techniques. Frequently: Fountas & Pinnell Leveled Reading system, QRI (Qualitative Reading Inventory) Word List

Up to teacher discretion what specific lessons follow; a lot of the leveled reading programs/books come with lessons and tips on how to conduct guided reading for different levels

***d. Number of hours for reading for each grade? Number of days per week?***
2 hours of LA/Reading a day- every day with 45 minutes of Guided Reading imbedded into that period. Guided Reading is when class split into groups and receive instruction on their level. The students are broken into groups based on levels/needed skills.

***e. What does a general day look like?***
First thing is a "morning message" or some type of writing exercise/workshop for about 30 minutes; then the class has 10 minutes of read-aloud from a book followed by 20 minutes of discussion about book. There is always 10-15 minutes of word study (for fluency help) - sometimes that is embedded into 45 minutes of guided reading and sometimes it is separate. During 45 minutes of guided reading students are broken into small groups (Split for T2 and T3) and low students receive help from teacher or reading paraprofessional

***f. Who implements the reading Tier 1 program (if anyone) with the teacher?***
Paraprofessionals work with teachers for T1 and T2

***2. Tell me about the Tier 2/Tier 3 process.***

***a. Who administer Progress Monitoring measures for these tiers and how are they trained?***
T2- Teachers; T3- SPED Staff or reading specialist if needed

All participants are in constant communication about students and progress

Trained similar to T1

***b. How often is data evaluated?***
T2- PM every week to every other week
T3- PM 1-2 times a week

***c. Do you use a particular curriculum or program(s) for intervention?***

Leveled texts; A-Z readers; Teachers learn strategies through "Strategies that Work"-book by Stephanie Harvey and Anne Goudvia; these books have lessons/guidelines in them as well

***d. Number of hours per day? Number of days per week?***

T2/T3 have 45 min direct instruction/intervention every day through imbedded time in schedule

T2 gets targeted help in classroom during IE time (unless have other services such as Speech, ESL)

T3 is also pulled during "IE" time (Intervention or Extension). T3 are pulled by reading specialist or SPED teacher. This happens 4-5 days a week for 30 minutes

***e. What does a general intervention session look like? Groups and group size?***

T2: Typically 6-9 for IE and 10-15 for time within class
T3: 6-8

***f. Who implements the program?***

T2: teacher and reading paraprofessional

T3: SPED teacher and Reading Specialist

***g. Intervention targeted skills?***

Target strategies: Visualize, Look Backs, Inferences, Text Connection, Synthesize Comprehension
Fluency through Sight Words practice

***3. Tell me about the decision making and data analysis process.***

***a. How do you identify the students for each tier based on progress monitoring and benchmark data?***

Based on their Leveled reading placement, AIMSWeb results, RAI (Reading Assessment Inventory) benchmark scores

Essentially based on levels/cut-offs/ranges determined by programs using/county

***b. How often does your team meet to review data and make decisions about progress?***

Grade levels have 1/month meeting with the reading specialist
Grade levels have 1/week meetings as grade levels to discuss data
1/every other month the entire team comes together (1/2 day) for data review/discussion of placement

***c. Who makes decisions?***

The data/RTI team: ESL teacher, general education teacher, Reading Specialist, Technology specialist, Principal, Assistant Principal

***d. How do students transition between tiers (what is the process)?***

Each teacher keeps a "running record" of each student/class which contains progress, scores, break downs for each student.

Decisions are based on results from the running record/benchmark data

The running record helps teachers/specialists know what a student's exact deficit(s) is (i.e. processing, comprehension, decoding)

Results from AIMSWeb and DRA benchmarks/PM show where the mistakes are made

Teachers are asked to keep specific records of each student and individual reading "behavior" (Strengths/weaknesses)

Typically, students are moved depending on progress 1 time a quarter

Guideline: if they have 3-4 data points on current leveled reading that is 95% or above accuracy/overall score move to different tier;

If see a student struggling 3-4 consecutive times consider movement as well

**School 2 Interview:**

1. ***Tell me about the Tier 1 (general education) process.***

***a. Who administers universal benchmarks and how are they trained?***

Teachers: Reading Specialists, ESL, Coordinators train other teachers. For some programs used like SRA, LLI, SIOP- training is provided by the county; Some training is also done through observations; At beginning of the year they have a "data walk"- go through previous data, what it looks like, means, etc.

***b. How often are Progress Monitoring measures administered in this tier?***

3 times a year (Fall, Winter, Spring)

***c. Is there a specific curriculum or program implemented? If so, what is it?***

K-3: PALS
2-5: AIMSWeb (MAZE and Fluency)
3-5: RAI (Winter and Spring)

***d. Number of hours for reading for each grade? Number of days per week?***

90 minutes/day for T1 (General Education)

***e. What does a general day look like?***

Specifics change every day, but typically there is whole group instruction followed by

centers to target more specific skills. The primary teacher is supported by a paraprofessional. Each is in charge of one center- students rotate through centers. When divided into groups, T1, T2, T3 are typically separated. Two centers are teacher directed; two are "practice" times. There is a lot of whole group modeling and practice as well

***f. Who implements the reading Tier 1 program (if anyone) with the teacher?***

Paraprofessionals

***2. Tell me about the Tier 2/Tier 3 process.***

***a. Who administer Progress Monitoring measures for these tiers and how are they trained?***

T2- Teachers/Paraprofessionals; T3- Pulled out; SPED or Specialist (Even though a student is not officially SPED, receive SPED-like interventions/PM in T3.
Trained through district or other specialists
PM for T2 and T3 happens weekly through AIMSWeb Benchmarks

***b. How often is data evaluated?***

Officially, at least monthly; but happens informally every week to every 2-3 weeks. Data more thoroughly evaluated after each benchmark. Each benchmark also informs teachers/grades/school what need to target as a whole group. Results from benchmarks lead discussion about tier transitions etc. When evaluate data, look at each Quadrant: who is in it, what looks like and how can we change/help. Follow the guideline that need 6-7 data points to really know if a student is going to respond to a particular intervention/instruction

***c. Do you use a particular curriculum or program(s) for intervention?***

Overall, no 1 program used; teachers have flexibility to use/tweak what will work best for their student(s); Programs typically utilized in part or whole include: PALS (Phonological Awareness Literacy Screening) lessons; Florida Reading Research lessons (Florida Center for Reading Research), SOAR (Set goals; Organize, Ask Questions, Record Progress) to Success (Houghton Mifflin) strategies; Houghton Mifflin "Intervention Kit" and Intervention Reading Series; LLI (Leveled Literacy Intervention)

Teachers have flexibility to use what works for their students; can use parts of different strategies to find what works

In the coming years going to use SRA Reading Laboratory (McGraw Hill), SIOP (Sheltered Instruction Observation Protocol by Pearson; although typically utilized for ELL, same principles work for all struggling students and especially those with SES disadvantages), and Herman Reading Method (Sopris West)

d. ***Number of hours per day? Number of days per week?***

*T2 and T3 receive T1 instruction PLUS:

T2: every other day (goal is to imbed everyday reading class) through IE (Intervention or Expansion- Pull out time for targeted interventions for students T2-3 and "expansion" or building proficiency/skills in more advance reading skills for T1 students) for 45 minutes

T3: every day for 30 minutes

Can be done in 1 session, or if teacher decides certain students need 20 minutes 2 times a day, that can be done as well

***e. What does a general intervention session look like? Groups and group size?***

Depends on what targeting that day, what program using;
T2: 4-5 students
T3: 2-3 students

***f. Who implements the program?***

T2: Teacher w/in classroom
T3: Pulled out- Reading Specialist, ESL teacher, SPED teacher

***g. Intervention targeted skills?***

Individualized- based on student needs/deficiencies
Use results of the RAI to guide in beginning; RAI has 8 strands- look for students w/deficiencies in each area; group according to deficiencies so intervention can be better targeted to their needs

Fluency- can be done in literacy center w/in classroom; also can do "Readers Theater"

***3. Tell me about the decision making and data analysis process.***

***a. How do you identify the students for each tier based on progress monitoring and benchmark data?***

AIMSWeb benchmarks (ranges) are used as well as teacher/administration/reading specialist discussion to determine what would be a "fair" range. Students are identified Green (T1), Yellow (T2) or Red (T3) like AIMSWeb. Teacher input is also used and important. If a student shows up on the low end of green, but the teacher feels he or she needs to be in T2, then that student can be in T2.

***b. How often does your team meet to review data and make decisions about progress?***

1 to 2 times a month officially to discuss data, but teachers meet weekly as a grade level and data can be discussed. Each teacher has a data binder with PM/Benchmark data for each student/class. This binder also has the interventions in place for T2/T3 students.
1 time a month the teachers meet with the Reading Specialist, Educational Diagnostician, Psych, AP to discuss data; Grade level also meets 1 time a month officially to discuss data.

***c. Who makes decisions?***

RTI Team (Teacher, Specialists, Psych, Diagnostician, AP); Teachers have a lot of input because they know each student and strengths/weaknesses as well as what works and doesn't work. Although they have programs/lessons they follow, teachers have flexibility to change/modify according to students' needs; know there is not 1 "boxed" program that will work for every student.

***d. How do students transition between tiers (what is the process)?***

Student progress is discussed at meetings; if a student is on track or moving above their goal line can discuss switching tiers. If student is flat lining or decreasing can discuss moving tiers as well as what needs to change- is it a lack of response to the intervention or instruction? Do they need to have a different structure for when interventions take place? (Break into multiple sessions a day, or one extended time a day).
Heavy reliance on PM data as well; follow the 6 or 7 data points to know how progressing.
"Quadrant" the students: Q1= T3; Q2=T2; Q3=T1; Q4= Advanced

***e. Any other measures used to evaluate and make decisions?***

PALS benchmarks, AIMSWeb Benchmarks, RAI/RLI