DEVELOPMENT OF AN ASSESSMENT RUBRIC FOR THE IMPLEMENTATION

OF RESPONSE TO INTERVENTION (RTI) AT ELEMENTARY SCHOOLS

by

KIMY H. LIU

A DISSERTATION

Presented to the Department of Special Education
and Clinical Sciences
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

September 2009

University of Oregon Graduate School

Confirmation of Approval and Acceptance of Dissertation prepared by:

Kimy Liu

Title:

"Development of an Assessment Rubric for the Implementation of Response to Intervention (RTI) at Elementary Schools"

This dissertation has been accepted and approved in partial fulfillment of the requirements for the degree in the Department of Special Education and Clinical Sciences by:

Elizabeth Harn, Co-Chairperson, Special Education and Clinical Sciences
Gerald Tindal, Co-Chairperson, Educational Methodology, Policy, and Leadership
Deborah Olson, Member, Special Education and Clinical Sciences
Yvonne Braun, Outside Member, Sociology

and Richard Linton, Vice President for Research and Graduate Studies/Dean of the Graduate School for the University of Oregon.

September 5, 2009

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

An Abstract of the Dissertation of

Kimy H. Liu   for the degree of   Doctor of Philosophy

in the Department of Special Education and Clinical Sciences

to be taken   September 2009

Title: DEVELOPMENT OF AN ASSESSMENT RUBRIC FOR THE

IMPLEMENTATION OF RESPONSE TO INTERVENTION (RTI) AT

ELEMENTARY SCHOOLS

Approved: _____

Dr. Beth Harn, Co-Chair

_____

Dr. Gerald Tindal, Co-Chair

Response to Intervention (RTI) is a relatively new approach to identifying

students with learning disabilities that has impacted how schools coordinate resources

and implement their instructional support and assessment systems. Because RTI is a

federal initiative and relatively new approach, there is little consensus on specific

approaches and practices to guide school in implementation. This approach has

potentially great consequences in providing supports in a more efficient manner for at-

risk students as well as improving outcomes for all students. However, to meet this goal,

schools need a tool to help guide and evaluate their practices to ensure quality

implementation. In this dissertation, I describe the development of a rubric designed to meet this need. This *RTI Assessment Rubric* is intended to be a tool that practitioners and researchers can use to evaluate the implementation of RTI at the elementary school level. Using qualitative methodologies, the rubric was developed based upon a review of the literature and validated by known experts in the field in terms of its breadth and depth of content. Through this process six components were identified which include evaluating: Tier 1 instruction, universal screening procedures, Tier 2 and Tier 3 interventions, progress monitoring practices, evidence-based decision-making, and organizational supports. Next, case studies were created on 5 second grade students who are at risk for reading disabilities in two elementary schools implementing RTI practices. These case studies included classroom observations, interviews with teachers/administrators, record reviews, as well as evaluating student response to school implemented interventions.

When these case studies were examined by external reviewers using the RTI rubric, the device demonstrated initial evidence of inter-rater reliability and sensitivity in discriminating between the quality of RTI implementation in these two schools. Teachers reported that the rubric captured the RTI experience within their building. This rubric appears to be useful for both researchers and practitioners as they begin implementation of RTI in schools because it appears to provide relevant, practical, and useful information to guide schools' next steps in RTI implementation.

CURRICULUM VITAE

NAME OF AUTHOR: Kimy H. Liu

PLACE OF BIRTH: Taiwan

DATE OF BIRTH: April 11, 1967

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

    University of Oregon, Eugene, OR
    Simmons College, Boston, MA
    National Tsing-Hua University, Hsin-chu, Taiwan

DEGREES AWARDED:

    Doctor of Philosophy, Special Education, 2009, University of Oregon
    Master of Education, 2004, University of Oregon
    Master of Arts in Teaching, 1994, Simmons College
    Bachelor of Arts, Foreign Language, 1989, National Tsing Hua University

AREAS OF SPECIAL INTEREST:

    Learning Disability
    Reading

PROFESSIONAL EXPERIENCE:

    Research Assistant, University of Oregon, College of Education,
        September 2006 - June 2009

    Research Teaching Fellow, University of Oregon, College of Education,
        September 2005 – June 2006; June 2004 – August 2005

PUBLICATIONS:

Alonzo, J., Liu, K., Tindal, G. (2008). Examining the Technical Adequacy of Second-Grade Reading Comprehension Measures in a Progress Monitoring Assessment System (Technical Report #0808). Eugene, OR: Behavioral Research and Teaching.

Alonzo, J., Liu, K., & Tindal, G. (2008). Examining the Technical Adequacy of Second-Grade Reading Comprehension Measures in a Progress Monitoring Assessment System (Technical Report # 44). Eugene, OR: Behavioral Research and Teaching.

Jung, E., Liu, K., Ketterlin-Geller, L., Tindal, G. (2008). Instrument Development Procedures for Mathematics Measures (Technical Report #0802). Eugene, OR: Behavioral Research and Teaching.

Liu, K., Carling, K., Ketterlin-Geller, L., Tindal, G. (2008). Instrument Development Procedures for Rapid Reading Rate Measures (Technical Report #0805). Eugene, OR: Behavioral Research and Teaching.

Liu, K., Ketterlin-Geller, L., Yovanoff, P., Tindal, G. (2008). Examining Item Functioning of Math Screening Measures for Grades 1-8 Students (Technical Report #0804). Eugene, OR: Behavioral Research and Teaching.

Liu, K., Sundstrom-Hebert, K., Ketterlin-Geller, L., Tindal, G. (2008). Instrument Development Procedures for Silent Reading Measures (Technical Report #0803). Eugene, OR: Behavioral Research and Teaching.

Liu, K., Sundstrom-Hebert, K., Ketterlin-Geller, L., Tindal, G. (2008). Instrument Development Procedures for Maze Measures (Technical Report #0806). Eugene, OR: Behavioral Research and Teaching.

Alonzo, J., Liu, K., & Tindal, G. (2007). Examining The Technical Adequacy of Reading Comprehension Measures in a Progress Monitoring Assessment System (Technical Report # 41). Eugene, OR: Behavioral Research and Teaching.

ACKNOWLEDGMENTS

"The bridge between a promising idea and the impact on

students was implementation, but innovations were seldom implemented

as intended." (Berman and McLaughlin, 1976, p. 349).

This dissertation was about developing an assessment rubric for the

implementation of the Response to Intervention (RTI) model and documenting two

schools' implementation of RTI. Implementation of a brilliant idea involves a willful and

intentional change. It requires the readiness to embrace the change, the skills and

knowledge to sustain the change, the leadership to prioritize the tasks and negotiate with

competing contingencies, and the organizational supports to facilitate the persistent

efforts. In a strange way, this dissertation mirrored my academic journey at University of

Oregon. My academic journey at University of Oregon started with a promise, a brilliant

idea, to redefine my life in a time of crisis. This promising idea could not have come to

fruition without the guidance and support from my teachers, colleagues, family and

friends.

First, I would like to thank my advisor Dr. Beth Harn for shepherding me through

the program and the preparation of the manuscripts. Second, I would like to thank Drs.

Deborah Olson and Yvonne Braun for their insightful comments and feedbacks. I want

to thank Drs. Ed Shapiro, Doug Marston and Terri Wallace for reviewing the content of

my rubric; and Dr. John Tenny, the developer of ECOVE for invaluable technical

support. I would like to thank Drs. Larry Sullivan, Yvonne Curtis, and Julie Alonzo of

the UO-4J National Model Demonstration Project of RTI Study for their support of my

dissertation; and principals and teachers of Gilham and Edgewood Community

To My Guiding Light, Dr. Barbara Bateman

To My *Godparents*, Karen and Carl

TABLE OF CONTENTS

Chapter   Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Controversy has surrounded the reliability and validity of identifying students with

specific learning disability (SLD) since the inception of the category in PL 94-142 (Kavale

& Forness, 2000). Dissatisfaction with both the SLD identification process and the delayed

delivery of academic intervention services for "at-risk" students has prompted researchers

and practitioners to seek alternative ways to identify SLD and to provide timely

interventions to low-performing students. IDEA 2004 recognized "scientific, research-

based interventions" such as Response to Intervention (RTI) as one of the alternative

approaches to replace the traditional IQ-achievement discrepancy model for identifying

students with SLD (IDEA 2004 Final Rules, 34 CFR §§300.307, 300.309 and 300.311).

The need to improve all students' learning outcomes and the recognition of RTI as a

legitimate approach for identifying students with learning disability fueled interest in

studying RTI implementation.

What Is RTI?

Response to Intervention (RTI) is a complex process for addressing several

components of education service delivery for low-performing students, including:

diagnosing, labeling, instructing, and evaluating. RTI is anchored upon a cycle of

problem-solving and multi-tiered service delivery approach, with the three-tiered model as the most popular version (Denton, Vaughn, & Fletcher, 2003; Fuchs & Deshler, 2007; Fuchs, Mock, Morgan, & Young, 2003; Fuchs, 2003; Marston, Muyskens, Lau, & Canter, 2003). The cycle of problem-solving illustrates a behavioral approach highlighted by its emphasis on a data-driven decision-making process that includes identifying the problems, planning the intervention, implementing the intervention, assessing the student learning outcomes, and using this assessment data to improve instructional design and delivery (Burns, Appleton, & Stehouwer, 2005; Fuchs et al., 2003; Marston et al., 2003; Troia, 2005).

Originally developed in public health service, the three-tiered model was subsequently applied in the school setting to remediate students' antisocial behaviors (Larson, 1994; Simeonsson, 1994; Walker, Horner, Sugai, Bullis, Sprague, Bricker, & Kauffman, 1996). In this model, service providers coordinate Tier 1, Tier 2 and Tier 3 Interventions within a school: Tier 1 interventions focus on Universal Screening and prevention; Tier 2 interventions focus on high-risk groups and intervening early before problem behaviors or skill deficits become more serious; Tier 3 interventions focus on individuals who do not respond to Tier 2 interventions and provide them with even more targeted and intensive interventions (Grimes & Kurns, 2003; Kovaleski, 2003; Lau, Sieler, Muyskens, Canter, VanKeuren, & Marston, 2006). The three-tiered model provides educators a logical way to allocate educational resources and to improve the efficiency of the instructional delivery system (Gresham, 2002).

RTI has been defined as the practice of (a) providing high quality instruction or intervention matched to student needs and (b) using learning rate over time and level of

performance to (c) make important educational decisions (Batsche, Elliott, Graden, Grimes, Kovaleski, & Prasse, 2005). The purposes of RTI include providing effective early intervention to at-risk or low-achieving students and creating better identification systems for students who are suspected of having specific learning disabilities (Burns et al., 2005; Burns & Ysseldyke, 2005; Coyne, Kame'enui, & Simmons, 2001; Gresham, 2002; Justice, 2006). RTI is designed to coordinate the resources within the school building to improve the effectiveness and efficiency of instructional delivery (Lau et al., 2006).

Definitions of RTI, although providing an overview of what the approach is supposed to include, do not provide the specific procedural guidelines needed to evaluate the implementation of RTI in schools. For example, how do schools determine if students have received "high quality instruction or intervention designed to match their needs"? How do schools determine if the "learning rate" of the given student was adequate? How do schools know that the information they have collected was *the* "right" data and will assist in making "important educational decisions"? How one operationalizes these concepts depends, to a great extent, on one's philosophy about the approach.

When viewed from the systems perspective, RTI can be seen as a logical way to coordinate the resources of general education, special education, Title One, and English Language Learner programs (Fuchs & Fuchs, 2006; Johnson, Mellard, Fuchs, & McKnight, 2006; Marston, Lau, & Muskens, 2007a; Peterson, Prasse, & Shinn, 2007). Proponents of this view suggest RTI can be used to improve the effectiveness and efficiency of the service delivery system in schools, thereby improving all students' learning outcomes. Proponents of this view stress the importance of capacity building and

sustained practice. When viewed from a narrower perspective, RTI can be thought of as a process to provide at-risk students early interventions (Tilly, 2003; Vellutino, Scanlon, Sipay, Small, Pratt, & Chen, 1996; Vellutino, Scanlon, Small, & Fanuele, 2006), or as an alternative approach for identifying students with specific learning disability (Gresham, 2002; Kovaleski, 2003; Speece, Case, & Molloy, 2003).

How an RTI approach is enacted in a school system may well depend on the perspective about the approach held by key decision makers. While some (Grimes & Kurns, 2003; Lau et al., 2006; Marston et al., 2007a; Peterson et al., 2007; Tilly, 2003) design their RTI system as a problem-solving model, others (Fuchs & Deshler, 2007; O'Conner, Harty, & Fulmer, 2005b; Torgesen, Alexander, Wagner, Rashotte, Voeller, & Conway, 2001; Vaughn, Linan-Thompson, & Hickman, 2003) focus more on an approach that emphasizes a standard protocol, and still others (VanDerHeyden, Witt, & Gilberton, 2007) use a hybrid of the two models in designing their RTI programs. If those designing the RTI process are more in favor of a problem-solving model, the intervention provided can be more tailored to individual students' instructional needs. In contrast, those who focus on RTI from a standard protocol model perspective emphasize on adhering standardized administration guidelines and on providing proven effective intervention programs. The differences between these approaches have led to confusion over how schools are to implement the process of RTI (Burns et al., 2005; Christ, Burn, & Ysseldyke, 2005; Reschly, Coolong-Chaffin, Christenson, & Gutkin, 2007). This confusion poses a challenge for schools. Teachers and administrators wish to implement the advocated federal policy, yet have no clear guidelines for how to determine if they are implementing it appropriately. Because the goal of RTI is to coordinate the resources in

the school system to improve student learning outcome, practitioners need a method for evaluating the quality of RTI implementation at both the school and student level.

In this dissertation, I describe the development of a rubric designed to respond to this identified need. This *RTI Assessment Rubric* is intended to be a tool that teachers and administrators can use to document and evaluate their implementation of RTI at the elementary school level. It is based on the essential components of RTI identified through a literature review. Although the RTI Assessment Rubric is not a test *per se*, it is designed to be an assessment instrument. Thus, in designing the rubric and gathering reliability and content related validity evidence related to its use, I followed typical test development procedures recommended by Downing (2006), supplemented with Yin's (2003) case study methodology to provide data by which to evaluate the sensitivity, reliability, and utility of the rubric. The next sections provide a general overview of this process to contextualize my research questions.

Conceptualizing the Process of RTI Implementation from a Student Perspective

This study focused on the implementation of RTI at the elementary school level, specifically addressing the content area of reading. Because an expected outcome of implementing RTI is to improve students' learning outcomes, the RTI assessment rubric must provide a mechanism for evaluating the design and delivery of instructional service provided to individual students. Based on my review of the literature (which will be presented in Chapter 2), a student with severe reading skill deficits goes through the following process to receive different types of interventions based on an RTI model:

1. A student receives Tier 1 instruction, which is provided to all students.

2. After a few weeks of Tier 1 Instruction, the student and all his/her classmates receive screening to determine whether they need Tier 2 intervention.

3. If the student scores below a pre-determined cut-score, he/she is assigned to receive targeted Tier 2 intervention in a small-group setting.

4. The student is monitored for his/her progress in learning while receiving Tier 1 instruction and Tier 2 intervention.

5. Teachers use the data gathered from progress monitoring to decide whether to continue, discontinue, or change the interventions the student is receiving.

6. The student who does not respond to Tier 2 intervention receives more explicit, intense, and targeted Tier 3 interventions.

7. The student receiving Tier 3 interventions continues to be monitored.

8. If the student does not respond to the most intensive intervention provided within the RTI program, he/she is referred for an evaluation of his/her eligibility for special education.

All of these steps are related and will be discussed in greater detail in the literature review. Based on the identified RTI components and their operational definitions, I developed criteria to evaluate the degree to which a school had implemented each component. The rubric is intended to encompass all of the important components of RTI implementation and to be sensitive enough to distinguish the qualitative differences that exist between schools at different stages of RTI implementation.

Creating a Rubric for Documenting RTI Implementation

The process I used in the development of the RTI Assessment Rubric was adapted from standard procedures for developing test instruments, an approach selected with a

goal to develop an instrument that could be used to assess RTI implementation. The framework of test design served as a guide throughout my work. Downing (2006) recommends a twelve-step process for creating an effective assessment tool that meets the standards established by the joint recommendations of the American Educational Research Association (AERA), the American Psychological Association (APA) and the National Council on Measurement in Education (NCME). Half of the steps in Downing's model specifically address tasks related to student assessments with individual items and scores. The other six are applicable to the process of developing the RTI Assessment Rubric. These six steps include: (a) identifying the important components of the targeted domain; (b) operationalizing these important components and providing descriptors for each component at different levels of understanding; (c) creating test items based on the descriptors listed in the test blueprint; (d) gathering content related validity evidence; (e) conducting a pilot study on the assessment to gather evidence about the stability of the instrument, and (f) creating a technical report and documenting validity evidence (Downing, 2006).

These procedures were used to construct the rubric. First, I focused on identifying the essential components of RTI through a synthesis of the literature. Second, I provided operational definitions for each component, categorizing implementation into three levels of implementation: *fully*, *partially* and *not at all*. Third, for each level of implementation, I provided a set of descriptors as observable or measurable criteria, organizing these into a rubric. Fourth, experts in the field of RTI reviewed and provided content validity evidence for the rubric. Fifth, two independent raters used the rubric to evaluate the implementation of RTI in two elementary schools. The inter-rater agreement between

raters was established as well as evidence for the stability or reliability of the rubric. Finally, the raters' and the participating teachers' feedback on the relevance and ease of use of the rubric was gathered to document its utility and social validity. Combined, these steps provide information related to the validity and reliability of the rubric for use as a tool to evaluate RTI implementation. Ultimately, then, in this dissertation I present the results of an instrument development and validation study, while concurrently reporting on a case study of RTI implementation at two schools in the Pacific Northwest. This dissertation focusing on addressing the following research questions:

Research Questions

1. To what extent does the RTI Assessment Rubric appropriately measure the construct of RTI implementation?

2. What evidence substantiates the reliability of the RTI Assessment Rubric for evaluating the implementation of each identified RTI component?

3. What evidence substantiates the sensitivity of the RTI Assessment Rubric for measuring qualitative differences in schools' RTI implementation?

4. Do differences in RTI implementation, as measured by the RTI Assessment Rubric, correspond with different student outcomes?

CHAPTER II

REVIEW OF RELEVANT LITERATURE

Response to Intervention (RTI) is a relatively new process schools can use to identify students with possible learning disabilities. The challenge that schools face in using this process is that the procedures for implementing it are neither clear nor widely accepted.

Process in Identifying Seminal Writings

First, I conducted a literature review using three computer databases: ERIC, Education Abstracts, and PsychINFO. Because I was interested in current RTI practices, I focused my search on articles published after the year 2000. I conducted the initial search using combinations of key researchers and key words or phrases. Examples of key researchers in the reading intervention studies include Deno, L. Fuchs, D. Fuchs, Foorman, Good, Grimes, Kame'enui, Marston, Reschly, Tilly, Vellutino, and Vaughn. The key words and phrases I used to locate the RTI articles include response (or responsiveness) to intervention, response (or responsiveness) to instruction, problem solving model, Heartland or Heartland Model, intervention-based assessment, instructional support team, Minneapolis problem solving model, standard protocol model, early reading intervention, and three-tiered model.

All key words, with the exceptions of the last three, were the key words used in a recent RTI meta-analysis (Burns et al., 2005) that focused on the problem-solving models with RTI practice. To broaden the search, I added other three sets of key words: (a) standard protocol model vs. problem-solving model, (b) three-tiered model, and (c) reading interventions. I added the standard protocol model and problem-solving model because they represent two different ways to conceptualize the RTI implementation process. I added the three-tiered model because it is a distinctive feature of RTI. I added the search term reading intervention because it was related to the content of this study. I conducted a secondary search by examining the references of relevant articles to look for recurring references that were not found in the primary search. I included Google Scholar as an additional database in the secondary search.

Second, I reviewed the abstract of each selected article to determine its relevance. I used the following inclusion/exclusion criteria to refine my selection of articles. I selected articles directly addressing (a) the RTI implementation process, (b) elementary school students, (c) reading instruction or interventions, and (d) RTI, one of the four problem solving models, or the standard protocol model. The first three criteria were aligned with the purpose of this study The last criterion was selected to include a wide range of RTI implementation approaches. Next, I eliminated the articles focusing on English Language Learners; behavior modification; math or science instruction; and identifying students as having learning disabilities because these emphases were not aligned with the purpose of this study.

Third, I set a benchmark for identifying essential components of RTI implementation. I was cognizant that articles took varied perspectives in conceptualizing

RTI, so some articles focused on certain aspects and did not discuss other aspects. Because of this variety, and to enable a broader understanding of these components for this project, I used a modest *a priori* standard that 40% of the reviewed articles had to address a component for that particular component to merit further analysis. I chose to set this standard at this level to reduce the chance of committing a Type 2 error, failing to identify a component that warranted further analysis.

Next, I reviewed the articles and coded them both by the way they conceptualized the process of RTI and by the components of RTI they specified. Summaries of these coding analyses are displayed in Tables 1 and 2. (See Appendix A for all the Tables).

Last, I compared my list of essential components with two authoritative handbooks designed to guide practitioners implementing the RTI model. The two handbooks were *Response to Intervention: Policy considerations and implementation* (hereafter referred as NASDSE, 2005), published by the National Association of State Directors for Special Education (Batsche et al., 2005), and *Response to Intervention: How to do it?* (hereafter referred as the RTI Manual), published by the National Research Center on Learning Disability (NRCLD) in 2006 (Johnson et al., 2006). The purpose of this comparison was to validate the six essential components through triangulation of the selected articles, the NASDSE 2005 and the RTI Manual.

In a typical instrument validation process, test developers compare students' responses on a newly developed instrument to their responses on an instrument with known technical adequacy to gather evidence of concurrent validity. In this study, I used the components identified in the two widely-cited handbooks as a proxy for a validated test instrument or an established criterion for the rubric. The results of this comparison

are displayed in Table 3. Comparing the components identified in the literature review with those identified in the two handbooks yielded content validity evidence for the components included on the RTI Assessment Rubric.

Results of Literature Review

In all, 26 articles met the selection criteria for this literature review. Of these, 21 were located through the primary search and 5 were located through the secondary search. In the following sections, I report the findings of my literature review and the comparison of the components identified by the literature review, NASDSE (2005), and the RTI Manual. I began by synthesizing the literature on the procedure of implementing an RTI approach and then move to a synthesis of the literature on the essential components of RTI.

What Is the Procedure of Implementing RTI?

All of the selected articles mentioned multi-tiered models, including a variation of a three-tiered model (see Table 1. For all the tables, see Appendix A). Throughout these articles, Tier 1interventions typically included universal screening and instruction in general education classes. Tier 2 interventions typically included progress monitoring and small group instruction, and Tier 3 interventions usually included long-term program change and more intensive instruction, often in a very small group, or 1:1 instruction (Denton et al., 2003; Fuchs et al., 2003; Kamps & Greenwood, 2003; Kovaleski, 2003; Marston et al., 2003; Tilly, 2003). Students receiving special education services were typically discussed as receiving Tier 3 supports, yet some researchers suggested that these services indicated a student who had moved beyond the tiers of RTI altogether (Burns & Ysseldyke, 2005; Fuchs et al., 2003; Justice, 2006; Kamps & Greenwood,

2003; O'Connor, 2000; Tilly, 2003). The number of tiers identified in the different studies varied from two to four, depending on whether or not special education and instruction provided to all students were included as part of the multi-tiered intervention model (Fuchs et al., 2003; Grimes & Kurns, 2003; Kamps & Greenwood, 2003; Kovaleski, 2003; Kraochwill, Clements, & Kalymon, 2007; Marston, 2005; O'Conner, Fulmer, Harty, & Bell, 2005a). Regardless of number of tiers, the procedures through which students moved through the different tiers described in these studies paralleled the description I provided in Chapter 1. The consistency of the tiered approach to instruction across all studies is worth noting, as this was the only component of RTI that was found uniformly in all cited work.

In contrast, only eight articles explicitly mentioned a cycle of problem-solving (Burns et al., 2005; Fuchs et al., 2003; Grimes & Kurns, 2003; Kovaleski, 2003; Lau et al., 2006; Marston et al., 2003; Tilly, 2003; VanDerHeyden & Jimerson, 2005). This lack of consistency across studies suggests that use of a problem-solving cycle might not be recognized as an overt feature of RTI, even though the concept appears to be closely . linked to evidence-based decision-making. Relevant to my study, the structure of the tiers and the decision rules governing student movement between tiers are reported as two main sources of confusion for practitioners implementing RTI (Lau et al., 2006; Tilly, 2003). These concerns were echoed by other researchers (Burns et al., 2005; Fletcher, Francis, Morris, & Lyon, 2005; Fuchs & Deshler, 2007; Gerber, 2005).

What Are the Components of RTI?

The consistency with which different components of RTI were addressed in the articles varied. All 26 reviewed articles mentioned progress monitoring and Tier 1, Tier 2

and Tier 3 interventions, or three-tiered models or multi-tiered interventions. Of the 26 selected articles, 24 mentioned evidence-based decision-making or data-driven decision-making; 16 mentioned organizational support; and 12 mentioned universal screening or school-wide screening (see Table 2).

These results suggest a consensus on the inclusion of multi-tiered intervention models, with three-tiered models as the most prevalent variations, and progress monitoring in an RTI approach. These two components are aligned with one of the dual purposes of RTI: using prevention and early intervention to improve students' learning outcomes. Articles (such as (Fletcher et al., 2005; Gresham, 2002; Kavale & Forness, 2000; VanDerHeyden et al., 2007; Vaughn et al., 2003)) addressing the other oft-cited purpose of RTI, providing a better LD identification process, were not included as a component of this study because they did not align with the intended purpose of this dissertation.

Two articles (Foorman & Torgesen, 2001; VanDerHeyden & Jimerson, 2005) mentioned the act of progress monitoring but did not extend the discussion of progress monitoring to evidence-based decision-making. Screening was sometimes implied within the context of selecting students into the Tier 2 intervention, but no actual procedures were described. In those instances, I did not indicate that the component of screening was present in the articles. The results of this literature review suggest that screening might not be as widely mentioned as some of the other components of RTI. However, it is important to keep in mind that studies that dealt specifically with developing the screening measures, e.g., VanDerHeyden, Witt, and Naguin, (2003), were excluded from

this literature review because their focus was on only one component of RTI, not the entire process.

Although 16 articles mentioned some type of organizational support, few provided enough detail about what was meant by the term to allow for proper evaluation. Thus, I turned to other literature for guidance (Rhoades & Eisenberger, 2002; Setton, Bennett, & Linden, 1996) when developing the operational definition and descriptors for organizational support. Based on the *a priori* criteria set for inclusion of a component as an essential feature of RTI, all six identified components mentioned above were deemed essential components of RTI (see Table 2). Each of these components is discussed in detail in the next section. This discussion is used to formulate the operational definitions of each identified RTI component. Before I start the detailed description of the RTI components, I answer this question: How do I decide how many components are needed to cover the entire construct of RTI?

Why Six Components?

Through the process of synthesizing the literature on RTI, I identified several important components of RTI. Some of them, such as multi-tiered intervention, effective instruction, and research-based assessments, are relatively broad. Some of them, such as direct instruction, explicit instruction, effective instruction, differentiated instruction, and research-based curriculum, overlapped each other. Above all, the concept of fidelity implementation (Gresham, MacMillan, Beebe-Frankenberger, & Bocian, 2000) or treatment fidelity (Fuchs, 1998) permeates all aspects of RTI process, but the term means different things in different contexts. For example, fidelity of implementation in the context of Tier 2 intervention is very different from fidelity of implementation in the context of

instructional leadership.

How many components are enough to encompass all of the most important elements of the entire construct of RTI? This is not an insignificant question. If I strove for parsimony at all costs, I could have identified instruction and assessment, or multi-tiered model and evidence-based decision-making as the two essential components for RTI. These two pairs of components are all grounded in the literature, but they are too broad to be practical in developing an action plan for the improvement of RTI practice.

In my initial review of the literature, I identified 10 components of RTI. However, some of these components overlapped and overlapping components can create confusion in organizing data. To rectify this flaw, I organized the identified components in such a way that they are not only closely related to each other, but also distinctly different. For example, I combined instructional leadership, administrative support, professional development and professional collaboration into one component, organizational support. The elements of organizational support have one shared purpose: facilitating the implementation of research-based instruction and assessment practices.

Similarly, I reclassified research-based curriculum, effective instruction, and tiers of instruction and interventions into two components, Tier 1 instruction and Tier 2 and Tier 3 interventions, because they differ in purpose, instructional emphasis, frequency, duration, and intensity. Research-based curriculum in Tier 1 instruction differs from the research-based curriculum in Tier 2 and Tier 3 intervention. The former aims to provide comprehensive literacy instruction and the latter to provide targeted and focused remediation in reading (Chard, Stoolmiller, Harn, Wanzek, Vaughn, Linan-Thompson, & Kame'enui, 2008) Tier 2 and Tier 3 Interventions were clustered as one component because

they shared similar purposes and could be evaluated using the same approach.

I separated universal screening from progress monitoring because these two components serve different purposes in the RTI process. The purpose of universal screening is to identify students who need additional instructional support. The purpose of progress monitoring is to track how well these students respond to additional intervention supports (Tier 2/3). I also separated evidence-based decision-making from progress monitoring because progress monitoring focuses on the test administration and collecting assessment data, while evidence-based decision-making focuses on the interpretation and use of these assessment data to make decisions (Hosp, Hosp, & Howell, 2007; Howell & Nolet, 2000).

The Six Essential Components of the RTI Model

These six components are distilled from the relevant research literature of RTI. I will describe each of the six identified components in the following section, with an emphasis on how each is implemented in schools to influence students' learning outcomes. I will also justify the choice of the criteria for evaluating these components. These criteria were included in the initial draft of RTI Assessment Rubric.

Effective Tier 1 Instruction

Tier 1 Instruction occurs in general education classes and is intended for all students. Its major functions are to enhance all students' academic learning outcomes (Batsche et al., 2005) and to prevent students from developing reading skill deficits (Fuchs et al., 2003; Gresham, Reschly, Tilly, Fletcher, Burns, & Prasse, 2005; Grimes & Kurns, 2003; Kovaleski, 2003). Tier 1 reading instruction gives students access to grade-level core curriculum; its instructional emphases encompass all the grade-level literacy skills as

defined by the state-mandated academic standards (Foorman & Torgesen, 2001; Grimes & Kurns, 2003; Marston et al., 2003; Mellard, Byrd, Johson, Tollefson, & Boesche, 2004; O'Connor, 2000). Research-based core reading curriculum used in Tier 1 instruction is supposed to align the content with grade-level benchmarks (National Center on Education and the Economy, 1999; Simmons & Kame'enui, 2006; Speece et al., 2003). Instruction provided in Tier 1 should be explicit and systematic (Justice, 2006; Kamps & Greenwood, 2003; Vaughn et al., 2003). Aligned with the professional consensus in the field of reading instruction (Denton et al., 2003; Elmore, 1996; O'Conner et al., 2005a; Simmons & Kame'enui, 2006; Snow, Burns, & Griffin, 1998; Torgesen et al., 2001), Tier 1 instruction should emphasize the "five big ideas" of reading instruction: phonemic awareness, phonics, fluency, vocabulary, and comprehension (National Reading Panel, 2000). Within an RTI approach, the quality of Tier 1 Instruction is judged by how groups of students respond to the provided instruction. To be deemed effective, Tier 1 Instruction alone should meet the instructional needs (i.e., meeting grade-level expectations) of approximately 80% of students without additional instructional support (Foorman & Ciancio, 2005; Foorman & Moats, 2004; Fuchs & Deshler, 2007; Good, Simmons, Kame'enui, & Chard, 2003; Simmons & Kame'enui, 2006; Torgesen, Rashotte, Alexander, Alexander, & MacPhee, 2003; Vaughn & Fuchs, 2003). If a significant percent of students in the class are identified as needing supplementary and intensive instruction, the adequacy of Tier 1 instruction must be investigated (Foorman & Ciancio, 2005; Foorman & Moats, 2004; Fuchs & Fuchs, 2007).

Effective Tier 1 Instruction should be ensured before considering students' eligibility for special education (Foorman & Schatschneider, 2003; Fuchs et al., 2003;

Fuchs, 2003; Heller, Holtzman, & Messick, 1982; Kamps & Greenwood, 2003; Speece et al., 2003) and this requirement is reiterated in IDEA 2004 (IDEA Final Rules, 2006, 34 CFR §300.306). To facilitate general education teachers delivering generally effective Tier 1 instruction to all students, professional development and consultation from district personnel, veteran general education teachers, and specialists should be put in place (Denton et al., 2003; Fuchs et al., 2003; Lau et al., 2006; Marston, 2005; Marston et al., 2003; Troia, 2005; Vaughn et al., 2003). To evaluate the quality of Tier 1 Instruction, Foorman (2005) and L. Fuchs (2003) recommend using class-wide assessment to determine the students' overall responsiveness to the instruction. If the responsiveness of one class is significantly lower than that of other classes in the building, then school administrators must explore the possibility of inadequate instruction in relation to students' low-performance. The purpose of evaluating quality of the provided instruction is to ensure the instruction is implemented as intended, this concept is often referred in the RTI literature as treatment fidelity (Fuchs, 1998) or fidelity of implementation (Gresham et al., 2000).

In this dissertation study, Tier 1 Instruction is defined as the comprehensive core reading curriculum instruction provided to all students in general education classes (Burns & Ysseldyke, 2005; Denton et al., 2003; Fuchs et al., 2003; Graden, Stoller, & Poth, 2007; Marston, Pickart, Reschly, Heistad, Muyskens, & Tindal, 2007b; McMaster & Wagner, 2007; Peterson et al., 2007) for a sufficient amount of time to meet grade-level instructional goals (Harn, Kame'enui, & Simmons, 2007; Justice, 2006; McMaster & Wagner, 2007; Simmons & Kame'enui, 2006). Teachers should teach all "five big ideas" specified in National Reading Panel Report (2000) (Foorman & Torgesen, 2001; O'Conner et al.,

2005b; Snow et al., 1998), or relevant ones to specific grade-level (Good & Kaminski, 2003; Good, Simmons, & Kame'enui, 2001; Marston, 2005), using evidence-based strategies (National Reading Panel, 2000). Because the quality of Tier 1 instruction is the first step in identifying students at-risk of developing reading problems (Speece et al., 2003), the RTI Assessment Rubric should include criteria that are designed for teachers to evaluate the provided instruction.

1. Content of instruction. Following the suggestions of researchers in the field of reading instruction, the rubric explicitly states that the instruction needs to include phonics, fluency, vocabulary, and comprehension in Tier 1 instruction. While phonemic awareness is important in first years of literacy acquisition (Vellutino et al., 1996; Vellutino, Scanlon, & Tanzman, 1998), students are expected to master phonemic awareness by the end of first grade. As such, I do not expect teachers teach phonemic awareness beyond Grade 2 in Tier 1 instruction. This expectation is confirmed by examining the state content standards in reading (e.g. Oregon, New York, Wyoming) as well as the widely used reading assessment designed to measure early literacy skills (such as DIBELS) (Good, Gruba, & Kaminski, 2002a; Good & Kaminski, 2003).

2. Instructional time. In the mental health service delivery system, cumulative hours of service are used as a quality indicator of service rendered (McGrew, Bond, Dietzen, & Salyers, 1994). Caroll (1963) posits that school learning is a function of time. Multiple studies (e.g., (Bollman, Silberglitt, & Gibbons, 2007; Callender, 2007; Foorman & Moats, 2004; O'Conner et al., 2005a; O'Conner et al., 2005b; Peterson et al., 2007; Pikulski, 1998; Treptow, 2006; Vaughn et al., 2003)) recommend that schools allot at minimum 90 minutes each day for reading, writing, and language arts. This 90 minute

block is designed to ensure sufficient time is reserved to teach and develop reading. Because this duration of instruction devoted to reading is so prevalent in the research literature, I used the minimum of 90 minutes as the second indicator for instructional quality in Tier 1 instruction.

3. Use of instructional time. In addition to the amount of the time, the use of instructional time as well as teacher behaviors are indicative of the quality of instruction provided (Allington, 1983; Allington, 2002; Gambrell, Wilson, & Gantt, 1981; Gickling & Armstrong, 1978). All of the selected articles emphasize focusing on the five big ideas and providing students systematic and explicit literacy instruction, which suggests the connection between the quality of provided instruction and the use of instructional time and explicit and systematic literacy instruction. However, these selected articles did not provide observable and measurable criteria of how to evaluate these two factors. Thus, I consulted research on student engagement, academic learning time, and study of effective teachers in developing this criterion.

Extended from Caroll's (1963) school learning model, Fisher (1981) further ties the teacher behavior, academic learning time, to student learning outcomes because these two factors influence the quality of instruction. Hollywood, Mott and Dodson (1995) reported that estimated engagement time in class ranged from 50-90%. Allington (2002) reports that exemplary elementary school teachers spend at least 50% of instructional time throughout the school day on reading and writing. They teach reading using various materials, including curriculum from other content areas (such as math and science). These effective teachers are often engaged in active teaching, using explicit explanation and direct instruction, and modeling useful comprehension strategies employed by

proficient readers. Moreover, there is a growing body of literature (Edmonds & Briggs, 2003; Foorman & Moats, 2004; Foorman & Schatschneider, 2003; Graves, Gersten, & Hagger, 2004; Greenwood, Abott, & Tapia, 2003; Grek, Mathes, & Torgesen, 2003; Hagger, Gersten, Baker, & Graves, 2003; Kim, Briggs, & Vaughn, 2003; Klingner, Sturger, & Harry, 2003) on classroom observation that links the use of instructional time and effective teaching behaviors. These studies provide evidence that effective use of instructional time increases the percentage of time students are engaged in active reading. Presence of those effective teaching behaviors can lead to increasing learning opportunities and facilitate active learning. Documenting the presence of effective teaching behaviors is a reasonable way to substantiate the effectiveness of the teaching (Graves et al., 2004; Hagger et al., 2003). For these reasons , I chose the use of instructional time and teacher behaviors as indicators for quality of Tier 1 instruction. I also chose to use a checklist to document the presence of effective teaching behaviors (Foorman & Schatschneider, 2003; Hagger et al., 2003).

4. Student time on task. As stated before, learning is a function of time (Caroll, 1963). Instructional time is the time marked on the school schedule, indicating who is going to teach what grades, what subject matter, during what time each day. Academic learning time is the portion of instructional time where students are actively engaged in learning Gambrell, Wilson and Gantt (1981) reported that, on average, good readers maintained higher engagement rate (as indicated by percent of time on task) than poor readers during reading instruction. These findings confirmed Vygostsky's theory that students learn more and learn more effectively when they are learning at their instructional level. It also justifies the use of student on task behaviors as an indicator for

the quality of instruction. In this rubric, in the context of reading instruction, "on task" is defined as students sitting in their seats properly, eyes looking either at teachers or at the instructional materials attentively, and focusing on assigned reading or writing activities.

5. *Satisfaction.* Satisfaction is a key component for the continued practice of the chosen intervention (Carnine, 1997). Teachers' beliefs have a powerful impact on the styles and effectiveness of their teaching (Kagan, 1992). Teachers' perception of the curriculum can influence their use, delivery, and the quality of instruction. Similarly, students' beliefs on their ability and their perception of the curriculum also influence their buy-in for the curricular programs and thereby influence the quality of learning (Blumenfeld, Soloway, Marx, Krajcik, Guzdial, & Palincsar, 1991). With this in mind, I included a criterion to solicit teachers' and student's opinions to assess their impression of the instructional experience and understanding of RTI implementation.

Universal Screening

The purposes of universal screening are three-fold: to determine the quality of general education (Tier 1) instruction, to identify students who are at-risk for long-term difficulties in learning to read, and to cluster students to form homogenous intervention groups for targeted and focused differentiated instruction (Foorman & Ciancio, 2005; Kamps & Greenwood, 2003). Practitioners use universal screening to determine whether the instructional quality in general education classes is sufficient to bring most students to grade-level benchmarks and what progress rate can be reasonably expected of individual students (Foorman & Ciancio, 2005; Fuchs, 2003; Jenkins, 2003; VanDerHeyden et al., 2003). The selected articles were consistent in recommending that universal screening be

administered to all students three times a year to measure the students' progress against given criteria, such as a grade-level benchmark standard (Jenkins, 2005; Vaughn, 2003; O'Conner, Fulmer, & Harty, 2003). Four types of screening measures were typically referenced in articles about RTI implementation: curriculum-based measurement (CBM), informal inventory of students' academic skills, high-stake state or district assessment and norm-referenced standardized achievement tests. One of the most common used CBMs is the Dynamic Indicators of Basic Early Literacy Skills,(DIBELS), which is designed to evaluate key components skills in the early literacy development (Denton, Fletcher, Anthony, & Frances, 2006; Fletcher et al., 2005; Fuchs et al., 2003; Fuchs, 1995; Good et al., 2002a; Marston et al., 2007b; Shinn, 2002). Texas Primary Reading Inventory is an example of informal inventory of students' academic skills used in the RTI Study (Marston, 2005). Woodcock-Johnson Achievement Battery Achievement Battery-III and Whechsler Individual Achievement Test II are two examples of norm-referenced standardized achievement used in the RTI studies (Fletcher et al., 2005).

In RTI, most of the screening measures are curriculum-based measurements (CBMs). The validity and reliability of CBM as screening and progress monitoring assessments have been well documented (Crawford, Tindal, & Stieber, 2001; Deno, Fuchs, Marston, & Shin, 2001; Hosp & Fuchs, 2005; Stage & Jacobsen, 2001), including its predictability on students' performance in high-stake tests (Marston et al., 2007b; Tindal, Yovanoff, & Alonzo, 2006). Its use in the problem-solving model, one of the popular approaches in RTI implementation, to assess students' acquisition of reading and math skills is also well documented (Shinn, 2002).

Although RTI is designed to identify students who are at risk of failing to meet grade-level performance benchmarks, researchers appear to be in agreement that a low score on a screening measure should not be considered sufficient to qualify students for differentiated interventions. If a significant number of students in class all scored below the cut-score on the screening measure, then improving Tier 1 Instruction should take priority over providing students with additional interventions (Fuchs et al., 2003; Speece et al., 2003).

In this dissertation study, universal screening is defined as using a measure to assess all students to identify those who score below the pre-determined benchmark and should receive additional instructional support. To justify its proposed use, a universal screener must have appropriate documentation and evidence of validity and reliability. Such documentation is used to support the technical adequacy of the chosen screening measures. The technical adequacy of the measures can be found in technical reports and independent reviews such as Kratochwill, Clement and Kalymon (2007).

In addition to technical adequacy, the implementation of universal screening is judged by frequency of test administration and data presentation, targeted students, and uses of the screening data. With the exception of Shinn (2002), most selected articles endorse administering universal screening measures three times a year, this practice is called benchmarking. To avoid selection bias, the screening measures must be administered to all students. To make the presentation of the screening assessment data user-friendlier, the data must be arranged in a way that is easy to identify low-performing students and to create homogenous intervention groups.

I proposed to evaluate the quality of universal screening by the technical adequacy of the chosen screening measures, frequency of administration, and data presentation. I also proposed to obtain evidence for the implementation of universal screening by examining the chosen measure's technical adequacy and reviewing the accuracy of the test dates and screening assessment data.

Effective Tier 2 and Tier 3 Interventions

Consistently across the different studies reviewed, Tier 2 and Tier 3 Interventions were identified as appropriate for students identified as at-risk of developing reading difficulty or who did not respond to Tier 1 Instruction, as indicated by their level and rate of progress both being lower than that of their peers (Fuchs, 1995; Fuchs, 1998; Harn et al., 2007; Kamps & Greenwood, 2003). In a well functioning school, it would be expected that approximately 15% of the student population would need Tier 2 interventions and only 5% of students would need Tier 3 interventions (Simmons, Kame'enui, Beck, Brewer, & Fien, 2003). To accelerate these students' learning, teachers must change the learning conditions to intensify instruction (Denton et al., 2006; Fuchs et al., 2003; Gerber, 2005; Gresham, 2002; Kovaleski, 2003; Marston, 2005; Torgesen et al., 2001). One way to help struggling readers accelerate their learning would be to provide purposeful and strategic intervention (Foorman & Torgesen, 2001; Kamps & Greenwood, 2003; O'Connor, 2000; Torgesen et al., 2001; Vellutino et al., 1996; Vellutino et al., 1998), using research-based supplemental programs for Tier 2 students and research-based intervention programs for Tier 3 students (Coyne, Kame'enui, Simmons, & Harn, 2004; Engelmann & Carnine, 1991; Good et al., 2003). Another way

to meet these students' needs would be to modify or re-teach the selected sections of research-based core curriculum with these students.

Tier 2 and Tier 3 interventions differ from Tier 1 instruction in their targeted population, purpose, instructional emphases, frequency, duration and intensity of instruction (Gresham, 2002; Harn et al., 2007; Marston, 2005; O'Conner et al., 2005b; Tilly, 2003). Tier 2 and Tier 3 interventions are provided to students who cannot meet grade-level benchmark with Tier 1 instruction alone (Dickson & Bursuck, 1998; Harn et al., 2007; Kamps & Greenwood, 2003; McMaster, Fuchs, Fuchs, & Compton, 2005; O'Conner et al., 2005a; O'Conner et al., 2005b; Vaughn et al., 2003; Vaughn, Wanzek, Linan-Thompson, & Murray, 2007).

The intervention programs at Tier 2 and Tier 3 are supplementary to the core curriculum (Simmons et al., 2003). The purpose of Tier 2 and Tier 3 interventions is to remediate specific areas of skill deficits (Coyne et al., 2004; Foorman & Ciancio, 2005; Harn et al., 2007; Kamps & Greenwood, 2003), therefore, they are targeted, purposeful, and focused on only one to three of the five critical aspects of reading as specified by the report of the National Reading Panel (2000) (Foorman & Ciancio, 2005; O'Conner et al., 2005a; O'Conner et al., 2005b; VanDerHeyden & Jimerson, 2005; Vaughn et al., 2003).

To accelerate at-risk students' learning and to prevent future performance deficits, teachers provide these low-performing students with scaffolding and immediate feedback during additional practice on essential skills (e.g. alphabetic principles and decoding) so that students can achieve mastery in basic reading skills (Engelmann & Carnine, 1991; Foorman & Moats, 2004; O'Connor, 2000; Torgesen et al., 2001; Torgesen et al., 2003; Vellutino et al., 1996). To increase students' engagement rate and intensity of instruction,

these interventions should be delivered in small-group settings (Bollman et al., 2007; Foorman & Ciancio, 2005; Grimes & Kurns, 2003; Harn et al., 2007; Kamps & Greenwood, 2003). The authors of these selected articles all agreed that the purposeful, targeted and differentiated Tier 2 and Tier 3 instruction can improve the struggling readers' learning outcomes.

To ensure that Tier 2 and Tier 3 Interventions are implemented with fidelity, the interventions should be monitored to ensure they are taught as they are intended (Fuchs, 1998; O'Conner et al., 2005a; Torgesen et al., 2001). In selecting and developing the descriptors for Tier 2 and Tier 3 interventions, I used the same research literature and structure to set the criteria to judge the level of implementation. To avoid redundancy, that information is not repeated here.

The evidence of implementing effective Tier 2 and Tier 3 Interventions could be obtained by a combination of observing Tier 2 and Tier 3 Interventions, attending school-based meetings (to observe how decisions are made in configuring Tier 2 and Tier intervention groups), reviewing school schedules and chosen intervention program materials, and conducting teacher and student interviews.

Progress Monitoring

In the literature reviewed, progress monitoring is defined as the regular collection of students' responses to a chosen assessment. The purpose of progress monitoring is to document students' incremental change on the targeted early literacy skills (Good & Kaminski, 2003) and to gather evidence on whether students have responded to additional instructional supports provided within an RTI approach (Burns & Ysseldyke, 2005; Christ & Hintze, 2007; Gresham et al., 2005; Justice, 2006; Marston et al., 2007b;

Tindal et al., 2006; VanDerHeyden & Jimerson, 2005; Vaughn et al., 2003). In documenting progress monitoring, some researchers measure students' functional behaviors linked to instruction (Grimes & Kurns, 2003) while others use general outcomes measures such as curriculum based measurement (CBM) to track the development of students' basic reading skills (Foorman & Torgesen, 2001; Fuchs, 1995; Fuchs, 2003; O'Connor, 2000; Peterson et al., 2007; Shinn, 2002; Vaughn & Fuchs, 2003; Vaughn et al., 2003).

Across the studies synthesized, students' classification as Tier 1, Tier 2 or Tier 3 changes as a function of instruction and learning (Fletcher et al., 2005). The frequency of progress monitoring reported in the articles reviewed varied, depending on tier of intervention. In general, however, Tier 3 students were monitored more often than Tier 2 students. Vaughn (2003) recommended Tier 2 students be monitored twice per month and Tier 3 students weekly which is supported by others (Good & Kaminski, 2003; Good, Simmons, Kame'enui, Kaminski, & Wallin, 2002b).

For the purpose of this study, progress monitoring is defined as the use of valid and reliable repeated measurement over time to assess students' specific reading skills appropriate to their instructional program. Similar to universal screening, progress monitoring measures must have documented evidence of validity and reliability for their proposed uses: to capture incremental growth and to judge whether students are making good progress toward their instructional goal based on their levels and rate of progress (Fuchs, 2003).

In addition to the evidence for reliability and validity, test developers must provide information about test administration and scoring (Good & Kaminski, 2003). To

ensure the accuracy of scoring, the test administration must be standardized and the training of test administrators and scorers must be reported. To increase the utility of measures, progress monitoring data are often presented in graphs so that progress patterns are visible and practitioners can examine the relative position between the aimline (expected performance) and trendline (the observed performance) and determine if students are making good progress.

In the RTI Assessment Rubric, the implementation of progress monitoring is judged by the technical adequacy of the measures, the quality of data collection, and the presentation of data (Good et al., 2002a; Good & Kaminski, 2003). Without evidence of technical adequacy, the use of progress monitoring measures might not be justifiable. If the tests are not administered and scored according to standardized protocols and the data are not gathered systematically and regularly, the inferences made from the data being gathered might be invalid. If the data are presented in a way that teachers cannot easily differentiate students who make good progress from students who do not, the utility of the test would be in question.

Evidence-based Decision-making

Evidence-based decision-making is the logic and rationale behind RTI (Reschly et al., 2007). By practicing evidence-based decision-making, teachers and administrators use scientific inquiry and empirical evidence to guide the decision-making process in the design and delivery of instructional service (Stoner & Green, 1992). This approach is best illustrated in the problem-solving model within the RTI framework (Bollman et al., 2007; Callender, 2007; Ervin, Schaugency, Goodman, & McGlinchey, 2007; Graden et al., 2007; Grimes & Kurns, 2003; Kamps & Greenwood, 2003; Marston et al., 2007a; Tilly,

2003; VanDerHeyden et al., 2003). In the articles I reviewed, evidence-based decision-making is often mentioned in the context of progress monitoring and consulting with a multi-discipline team (i.e., special and general education teachers, specialists, administrators) to evaluate how well students were responding to instruction (Bollman et al., 2007{Burns, 2005 #17; Grimes & Kurns, 2003; Kamps & Greenwood, 2003; Shinn, 2002; Tilly, 2003). Evidence-based decision-making is a driving force that directs actions through the assessing-planning-teaching cycle of instruction, transforming progress monitoring data into information helpful for instructional planning (Marston et al., 2003; Tindal et al., 2006).

Good, Simmons, Kame'enui and Chard (2003) advocate that the evidence-based decision-making process and outcomes should be not only be reliable and valid, but also *manualized*, meaning that other qualified educators reviewing the same set of data, using the same decision-rules, would arrive at the same conclusion. For evidence-based decision-making to be manualized, all instructional support team members (i.e., general and special education teachers, administrators, etc.) should be knowledgeable about the assessments, the curriculum, and the instructional delivery. More importantly, all team members should have a thorough understanding of the decision rules and how to use the data to modify instruction.

The *Handbook of Test Development* (2006) provides an authoritative explanation on how to use test scores to make valid inferences. To justify the proposed inferences, the interpretation of the test scores must meet the following five assumptions. First, the test instrument used must be technically adequate and aligned with the purpose of the assessment. Second, it must be fair to all of the targeted student population. Third, it must

be administered and scored by qualified personnel, using standard test administration and scoring protocols. Fourth, the judges who interpret the scores must possess adequate technical knowledge about the chosen instrument; know the intended use of the measures, the administration and scoring process, the decision rules, and understand how to sensibly interpret the scores. Last, but not least, the professionals who participate in the decision-making process must first evaluate the trustworthiness and accuracy of their interpretation of the data before making educational decisions. Educators must not make an important educational decision based on one test score. Instead, before making a decision, they should compare the likely result of the decision against the consequence of the decision and determine whether the totality of accumulated evidence is sufficient to make a reasonable judgment (Downing & Haladyna, 2006). Although these guideline are often used to evaluate the validity of test for its proposed use (AERA, APA & NCME, 1999), teachers should adopt these guidelines to judge the quality of empirical data and evaluate their own decision making process, especially as it relates to decisions with a large potential impact on students, such as instructional grouping.

For the purpose of this study, evidence-based decision-making is defined as judging the effectiveness of interventions using specific metrics from progress monitoring and program-specific measures and using professional judgment to determine if intervention modifications are warranted. In the RTI Assessment Rubric, the implementation of evidence-based decision-making is judged by what data are used to make the decisions, how the decisions are made, and what decisions are being made.

Organizational Support

The effectiveness of RTI in the studies I synthesized is influenced by both the coherence of instructional support across different tiers of interventions and the dynamic feedback between instruction and assessment. Organizational support is instrumental to attain these goals. Organizational support can be found through examples of strong leadership that organizes administrative support (Justice, 2006; Lau et al., 2006) and provides targeted and meaningful professional development for teachers (Batsche et al., 2005; Justice, 2006; Lau et al., 2006; Marston et al., 2003). Administrators with strong instructional leadership secure the necessary funding, time, and human resources to accomplish these tasks (Fuchs & Fuchs, 2006). In addition, they set clear standards or expectations for professional training and program implementation as well as help teachers prioritize the activities pertinent to the design and implementation of RTI (Justice, 2006; Lau et al., 2006; Marston et al., 2007a; Marston et al., 2003).

To implement RTI practice effectively, teachers need different sets of skills. Although teachers in general support high standards of teaching and learning, many teachers are not readily prepared to implement best practice as suggested by empirical research (Garet, Porter, Desimone, Birman, & Kwang, 2001) nor are they offered sufficient opportunities to learn new teaching strategies (Chard, 2004). Teachers are more likely to implement a new practice if the professional training they receive includes presentation of the theory behind the practice, demonstration and opportunity for hands-on practice during the training, and on-going prompts and feedback to the teachers as they engage in the practice (Showers, Joyce, & Bennett, 1987). To improve the efficacy of professional training, those designing it should consider the content of the training, the

format and method of delivery, and the relevance of the topic to the audience (Chard, 2004; Denton *et al.*, 2003; Gersten, Chard, & Baker, 2000).

According to Hierbert (1999) and Friend (2000), opportunities for teachers to learn new teaching methods should be sustained by collegial support and on-going collaboration with colleagues for the purpose of enhancing students' learning. These learning opportunities are often anchored by shared visions and explicit goals that are related to students' cognitive process, curriculum, and pedagogy (Friend, 2000; Garet et al., 2001). Professional collaboration appears to facilitate use of collective knowledge and skills. Teachers benefit from observing other effective teachers and reflecting on the reasons for their effectiveness (Ervin et al., 2007; Garet et al., 2001).

In contrast, insufficient professional training often leads to ineffective teaching and weak support from teachers. General education teachers often have not received adequate pre-service training on how to differentiate instruction and use data to inform instruction (Lyons, Fletcher, & Barnes, 2002). To avoid the problems associated with poor training, leadership must include professional development opportunities, including coaching and time to collaborate with colleagues (Callender, 2007). In RTI-reading interventions, teachers need professional training on how to implement research-based reading instruction and progress monitoring effectively (Johnson et al., 2006; Mellard, 2004; O'Conner et al., 2005a). For example, O'Connor and her colleagues (2005) provided their participating teachers multiple sessions of professional training sessions across the school year. Following each session, grade-level teachers discussed instructional activities, adoption of instructional programs, and the timeline for implementing these programs. They also discussed how to interpret students' progress

monitoring scores using benchmarks guidelines provided by researchers. This format combined professional training and professional collaboration into a coherent unit. This format sounds intuitively logical, yet O'Connor and her colleagues did not report the amount of time teachers spent in professional training and collaboration or how they evaluated the effectiveness of these practices.

In fact, the amount of teacher training provided by researchers and districts was rarely quantified in field study reports of RTI (Gerber, 2005). In the few exceptions, the amount of preparation provided by researchers varied widely. Vaughn and her colleagues reported 20 hours of preparation for four tutors and weekly meetings (Vaughn & Fuchs, 2003; Vaughn et al., 2003). Torgesen cited 40 hours of baseline training (Torgesen et al., 2003). In neither article did researchers discuss content, format, or duration of the training. Variation in the levels of professional training provided for teachers could account for the confusion teachers experience in implementing RTI practice. More research on professional training for RTI practice is needed, particularly on how to design and evaluate effective professional training (Fuchs & Fuchs, 2006).

Because most of the selected articles did not specify the format and content of professional development and professional collaboration for RTI, I turned to the literature on service delivery and professional training for guidance. For example, the implementation of mental health service delivery is judged by staffing, organization and service (McGrew et al., 1994). Staffing is judged by client to staff ratios, the maximum of the case loads, and the number of available staff at any given time. Organization is judged by the division of labors, extent of professional collaboration, and the supporting network for the staff and students. In this rubric, I included the evaluation of staffing and

organization in the organization support; addressed the evaluation of the service or instructional delivery earlier in the Tier 1 instruction and Tier 2 and Tier 3 interventions.

Little (1993) stated that teachers' professional development must center on academic standards, curriculum and pedagogy, and the extent, nature and use of student assessments. Darling-Hammond and McLaughlin (1995) further elaborated on the features of effective professional training. They stated that effective professional development must engage teachers in concrete tasks of teaching, assessment, reflection on the processes of learning and development. It must be designed to improve teachers' professional competency so as to improve student learning outcomes. Professional training should not be episodic, but rather sustained, supportive, and on-going, including various proven effective teaching strategies, such as modeling, coaching, and problem solving. Most importantly, it must be integrated into the school-wide improvement plan (Darling-Hammond & Mclaughlin, 1995; Garet et al., 2001).

Smith and Andrew (1989) and Glickman, Gordan and Ross-Gordan (2001) both emphasized the importance of the principal asserting instructional leadership. An effective principal can transform visions to clear goals and objectives. By communicating these expectations with the staff, soliciting their feedback in creating an action plan, allocating resources for program change, and monitoring the progress, principals can be forceful leaders who promote school-wide instruction reforms. Organizational support lays the foundation for any program change in school settings. Instructional leadership and administrative support helps setting the agenda, locating resources, and implementing the program changes. Professional development and professional collaboration are keys to increase teachers' capacity to improve all students' learning outcomes (Borman,

Hewes, Overman, & Brown, 2003; Desimone, 2002).

Within an RTI approach, professional collaboration and administrative support were reported as being important in designing and implementing research-based reading instruction and progress monitoring (Grimes & Kurns, 2003; Kamps & Greenwood, 2003; Kovaleski, 2003; Marston, 2005; Troia, 2005). The RTI Manual provides teachers, administrators, and other professionals specific to-do lists pertinent to professional collaboration (Johnson et al., 2006). These to-do lists allow teachers and administrators to determine if some of the practices are already in place in their schools. If such practices are not in place, the lists prompt teachers to prioritize them for future implementation.

For the purpose of this study, organizational support is defined as strong leadership that organizes administrative support to provide effective professional training and encourage professional collaboration. This operational definition is distilled from the studies mentioned in the previous paragraphs, including Little (1993), Darling-Hammond and McLaughlin (1995), and Desimone (2002). Principals are responsible for guiding the school team to set the goals and objectives, allocating resources, coordinating administrative supports to facilitate implementing research-based reading instruction and assessments, and monitoring the progress of program change. The administrative staff are responsible for monitoring students' attendance, ensuring the instructional and assessment materials and necessary equipment are in place, and coordinating times for instruction and assessments. Teachers must be provided with on going training and support so they have the knowledge and skills needed to implement research-based instruction and assessment, identify students' instructional needs, and tailor the instruction to remediate these identified skill deficits. Teachers also need to have on-going professional collaboration to enhance the

effectiveness and efficiency of instructional delivery. Therefore, it is appropriate to judge the quality of organization support to facilitate the implementation of RTI by these four indicators.

Validate the Identified Essential Components

I compared the identified six essential components of RTI to both the NASDSE (2005) and the RTI Manual, two authoritative documents on RTI implementation. The NASDSE (2005) gave recommendations regarding policy considerations and implementation of RTI; the RTI Manual aimed to help practitioners understand, design, and evaluate the features of an RTI program prior to implementation (Johnson et al., 2006). All six identified components are either emphasized or explicitly mentioned in these two documents (see Table 3).

Summary

The construct of RTI is comprised of six important components. Universal screening aims to identify students needing additional instructional support and determine the severity of their needs. Effective Tier 1 instruction in the general education setting is designed to facilitate most students making sufficient progress toward grade-level expectations and to prevent students from needing additional supports by teaching them well. Effective Tier 2 and Tier 3 interventions provide additional support on specific reading skills to students who cannot meet grade-level expectations with Tier 1 instruction alone. Progress monitoring tracks the learning progress of students who receive Tier 2 and Tier 3 Interventions on these specific skills. Evidence-based decision-making helps teachers using assessment data to improve instruction. Organizational support provides teachers with administrative supports, professional training, and

opportunity for professional collaboration so they can optimize the design and delivery of instruction to improve student learning outcomes.

All of these components are identified as essential to effective RTI implementation. Without effective Tier 1 instruction, some students might be identified as needing Tier 2 or Tier 3 interventions because of inadequate instruction. Without universal screening, teachers could not effectively identify students who need Tier 2 or Tier 3 interventions. Without effective Tier 2 and/or Tier 3 interventions, low-achieving students' skill deficits could not be remediated. Without progress monitoring, teachers could not determine whether students were responding to the provided interventions. Without evidence-based decision-making, the data gathered in progress monitoring could not be used to improve instruction. Without organizational support, teachers might not have the necessary training to implement research-based instruction, progress monitoring, and evidence-based decision-making effectively. Nor could students who need Tier 2 or Tier 3 interventions receive targeted interventions from effective instructors in small-group settings. Missing any of these components would undermine the integrity of an effective RTI implementation. Schools that wish to implement RTI effectively must address each of these components in meaningful ways.

In the next chapter, I will describe the criteria for each descriptor, the process of obtaining content-related evidence, and evidence for utility of the rubric

CHAPTER III

METHODS

Researchers and educators need a valid and reliable way to evaluate the implementation of RTI. The approach is relatively new, and schools lack an objective means of measuring the degree to which their implementation of the approach aligns with components of RTI identified by experts in the field as essential for successful implementation. To address this need, I focused my dissertation on the creation of the RTI Assessment Rubric, an instrument designed to provide researchers and educators with a tool by which to assess the implementation of RTI at elementary schools.

Because the decision regarding students' responsiveness to instruction is determined at the student level, a useful RTI rubric must document and evaluate the instructional support and assessment provided to individual students. At the same time, RTI is consistently identified in the literature as a systems-wide approach to service delivery. Thus, to be most useful, an RTI rubric must also be able to capture systems-level information. The purpose of this study was to develop a valid, reliable, and useful rubric for evaluating the implementation of the RTI approach. With the goal of developing an RTI Assessment Rubric in mind, four research questions are posed:

1. To what extent does the RTI Assessment Rubric appropriately measure the construct of RTI implementation?

2. What evidence substantiates the reliability of the RTI Assessment Rubric for evaluating the implementation of each identified RTI component?

3. What evidence substantiates the sensitivity of the RTI Assessment Rubric for measuring qualitative differences in schools' RTI implementation?

4. Do differences in RTI implementation, as measured by the RTI Assessment Rubric, correspond with different student outcomes?

In the remainder of this chapter, I describe the process of developing the initial draft of the RTI Assessment Rubric and conducting the content review and field study.

Development of the RTI Assessment Rubric

In developing the RTI Assessment Rubric, I followed a test development procedure advocated by Downing and Haladyna (2006). As explained in Chapter 1, the rationale for using this procedure was that the rubric could be considered a test of a sort, albeit an atypical one. As mentioned earlier, of the 12 test development steps, only six were relevant to the development of the rubric. In Chapter 2, I described the first two steps of this instrument development, using the literature review to identify these six essential components and to compose the operational definitions and descriptors for the initial draft of the rubric. In the next section, I will describe the process of setting the criterion for each descriptor.

Create an RTI Assessment Rubric

The RTI Assessment Rubric is comprised of six sub-tests, one for each of the six identified RTI components. For each component, an operational definition and three sets of descriptors are provided to describe the implementation of the components when being implemented fully, partially, or not being implemented at all. To be fully implemented all

of the specified sub-components must be in place. A rating of partial implementation

indicates that most of the subcomponents were in place and a rating of "not

implemented" indicated that none of the subcomponents were in place. (See Appendix B

for the design of the RTI Assessment Rubric.) By consulting the relevant literature, I

selected the following criteria for each identified RTI component.

Tier 1 instruction. In this rubric, Tier 1 instruction is judged by (a) alignment

between the instructional goal and instructional emphases, (b) amount of instructional

time, (c) use of instructional time, (d) students' percentage of time on task during the

reading instruction time, and (e) the satisfaction of teachers and students about the

instruction provided. Tier 1 Instruction is considered fully implemented if the following

criteria are met:

1. Teachers explicitly teach all "five big ideas" specified by the National Reading

Panel (2000) or relevant to specific grade-level. This standard is supported by the

consensus of all of the reviewed articles.

2. Teachers teach all students Tier 1 reading instruction daily for 60-90 minutes

with most of the time spent engaged in reading and writing activities. This standard is set

by considering the common practice of 90 minutes reading block as reported in many

articles e.g., (Bollman et al., 2007; Callender, 2007; Foorman & Moats, 2004; O'Conner

et al., 2005a; O'Conner et al., 2005b; Peterson et al., 2007; Pikulski, 1998; Treptow,

2006; Vaughn et al., 2003) and by considering the feasibility of having a 60-minute rather

than 90-minute reading block in a half-day kindergarten program.

3.The criterion for use of the instruction is set by considering the estimated use of

time in a typical class (Hollywood et al., 1995), in the effective teachers' classroom

(Allington, 1983; Allington, 2002), and the data gathered from classroom using various observation tools (Edmonds & Briggs, 2003; Foorman & Moats, 2004; Foorman & Schatschneider, 2003; Gambrell et al., 1981; Graves et al., 2004; Grek et al., 2003; Hagger et al., 2003; Kim et al., 2003; Klingner et al., 2003 Greenwood, 2003 #157). These studies all confirmed that effective teachers spend at least half of their reading instruction time teaching phonemic awareness, phonics, fluency, vocabulary, and comprehension in early elementary grades. Based on the consensus of these empirical studies, I set the standard that to be deemed effective Tier 1 instruction, teachers should spend 50% of instructional time teaching the 'five big ideas.'

4. Gickling and Armstrong (1978) reported that when students are being taught at their instructional level, they average on task behavior 80 - 90% of the time. In contrast, when students are taught at the frustration level, they average only 45% on task behavior, and when students are taught at the independent level, they are on task on average 53% of the time. Gambrell, Wilson, and Gantt (1981) reported that proficient readers engaged in active learning 92% of the time when working with teachers and 89% of the time when working independently. Poor readers engaged in active learning 83% of the time under both conditions. Considering both studies, I set the benchmark of percentage of time on task at 80%.

5. I chose to include teachers' and students' opinion about provided instruction and interventions as one of the quality indicators, because their opinions about the relevance and importance of provided reading instruction might influence their willingness to participate in the active learning process.

Universal screening. Based on the synthesis of the literature, universal screening is

considered fully implemented if all three features are in place:

1. The chosen screening measure has explicitly documented validity and reliability evidence aligned with the purpose of assessment (AERA, APA, & NCME, 1999; Downing & Haladyna, 2006). This standard is set to ensure that the chosen instrument is technically adequate for this proposed use. AND

2. The screening measure is administered to all students three times a year to identify students who need additional instructional support (Jenkins, 2003; VanDerHeyden et al., 2003). This standard is set to ensure that schools have multiple opportunities to systematically identify students who need additional instructional support over the course of the school year. AND

3. Data were consistent with the appropriate unit of analysis (teacher roster, grade-level placement). The alignment between the data presentation and its intended use can enhance the utility of the universal screening.

Tier 2 and Tier 3 interventions. In developing the descriptors for Tier 2 and Tier 3 interventions, I used the same literature reviewed in discussing Tier 1 instruction provided direction and insights. To avoid redundancy, that information is not repeated here. Readers are referred to the previous section for this information.

Progress monitoring. In this rubric, progress monitoring is considered fully implemented if it met the following criteria:

1. The measures have documented validity and reliability evidence for monitoring progress on a specific skill that matches students' instructional needs. This standard is set to ensure the technical adequacy of the chosen progress monitoring measures for their proposed use.

2. Teachers administer repeated measures of alternate forms for student receiving additional support weekly or biweekly for at least three times. This standard is set to ensure that data are gathered in reasonable intervals and no important educational decisions are made based on a single test score. The standard of "at least three data points" is set to ensure a stable trend of student progress is established.

3. Progress monitoring data are collected and displayed in appropriate graphs so that students' progress can be evaluated through visual inspection by comparing the relative position between the aimline and trendline. The standard related to data display is set to improve the utility of the progress monitoring data.

Evidence of teacher implementation of progress monitoring can be obtained by reviewing the progress monitoring data and the technical report of the chosen progress monitoring measures.

Evidence-based decision-making. Within an RTI approach, evidence-based decision-making is instrumental in helping teachers improve the effectiveness and efficiency of instructional delivery, so that teachers can help struggling readers progress toward grade-level proficiency. As stated in Chapter 2, evidence-based decision-making focuses on the interpretation of the assessment data and the actions to be made based on that interpretation. The quality of evidence-based decision-making is judged by these two elements, as well as the thoughtfulness in the decision-making process. In this rubric, evidence-based decision-making is deemed fully implemented if the following three criteria are met:

1. Teachers systematically examine classroom data (e.g., data of program specific assessments) and progress monitoring data. For progress monitoring data, teachers should

analyze graphs, focusing on students' levels of performance and rates of progress (slope) in relation to the aimline. This standard is set to address what data should be included in the decision-making process.

2. By examining the relative position between the observed scores (the trendline) and the expected scores (the aimline), teachers determine whether to fade, continue, modify, or intensify the students' instructional interventions based on the interpretation of the empirical data gathered. This standard is set to address how the decision should be made.

3. Based on the interpretation of the data gathered, teachers identify and modify specific instructional components to address the targeted areas of skill deficits. This standard addresses what decisions should be made in this evidence-based decision-making process.

Evidence of the implementation of evidence-based decision-making can be obtained by interviewing teachers, reviewing progress monitoring data, and observing school-based meetings in which teachers review progress monitoring data and make decisions based on those data.

Organizational support. In this rubric, organizational support is judged by the evidence of instructional leadership, administrative support, professional development, and professional collaboration. Organizational support is deemed fully implemented if the following four criteria are met:

1. There is strong leadership in school that supports implementing RTI.

2. Teachers receive administrative support in staffing, scheduling, and planning for instruction and assessment.

3. Teachers receive sufficient training so that they can perform their tasks effectively.

4. Staff support and engage in professional collaboration.

Evidence of *organizational support* can be obtained through observation of school-based meetings and inspection of school schedules. Interviews with teachers can also be used to substantiate the evidence of schools implementing the components of RTI.

Ask Content Experts to Review the Rubric

Once test items are developed, Downing and Haladyna (2006) recommend submitting the test to be reviewed by content experts in the targeted domain. The purpose of a content review is to determine whether the included items are representative of the range of the construct. The experts identify and exclude items that might introduce construct-irrelevant variance in the test instrument. They also judge whether the instrument contains all of the key components of the targeted domain.

Three RTI experts reviewed the initial draft of the RTI Assessment Rubric. Drs. Ed Shapiro, Doug Marston, and Teri Wallace were invited to review the rubric because they had published articles on RTI in peer-reviewed journals and they had extensive experience in RTI implementation. All three experts had recently served as principal investigators for Model Demonstration Projects on RTI funded by the Office of Special Education Programs (OSEP). All three could provide insights from the lens of researchers as well as that of practitioners. Information about their qualifications is provided in Appendix C.

These three RTI experts reviewed the rubric using a structured protocol. First, they reviewed the content to determine if there were any components of RTI not included

in the rubric. Second, they indicated whether they agreed with the operational definitions and descriptors provided in the rubric. Finally, they provided reasons for disagreement and/or suggestions for revision.

Preparation of School Reports

After the rubric had been reviewed and revised, I conducted a case study to document the RTI implementation process in two elementary schools. Data from the case study were compiled into detailed school reports. The school reports were shared with teachers and administrators at both case study schools to verify accuracy and completeness prior to being shared with the independent raters, who applied the RTI Assessment Rubric to evaluate RTI implementation at the case study schools.

Setting and Participants

I conducted my case study at two public elementary schools that had been participating in the OSEP Model Demonstration RTI Project since 2006. At each of the schools, I gathered data at the school, grade-level, classroom, and individual student level. The two schools shared the following characteristics: they were both non-Title One Schools and had no ELL programs. They both housed a regional learning center on campus. The schools differed primarily in the school size and staff experience using progress monitoring systems (see Table 4).

Teachers. All teachers at both schools were certified to teach in their current assignments and all but one of the teachers participating in the case study had Masters degrees. In School 1, years of teaching experience of case study teachers were 1, 6, and 10 years, respectively. In School 2, the teaching experience was significantly higher, with the two teachers reporting 26 and 29 years of experience, respectively (see Table 5).

The plan was to recruit six second-grade "Tier 3 students" and the teachers who provided them instruction; however, because not all students originally invited to participate gave permission to take part in the study, I ended up with five students in all. These students were recommended by their homeroom teachers to participate in this study because they had not made progress with the first Tier 2 interventions or they were already receiving Tier 3 interventions. In specifying the selection criteria, I believed that the severity of these students' performance deficits was likely to trigger all possible intervention options available within the district's RTI model. I hoped to compare and contrast across these five cases to demonstrate the variety of ways students might travel through the RTI system.

Two student-teacher dyads were chosen from each of the three participating classes to provide replication in each teacher's class because the students were of comparable reading skill levels at the beginning of the school year and had received the same core curriculum instruction from the same teacher (see Figure 1). This literal replication multi-case study design is an approach similar to conducting multiple replication experiments within qualitative research approaches (Yin, 2003). To further strengthen the study, the dyads across different classes were designed to be a theoretical replication (because students of comparable skill levels received different instructional support from different teachers in different classes). The theoretical replication multi-case study design is an approach similar to having a control group versus an experimental group in experimental design (Yin, 2003). Evidence from multiple replicating cases is more compelling than evidence from a single case study.
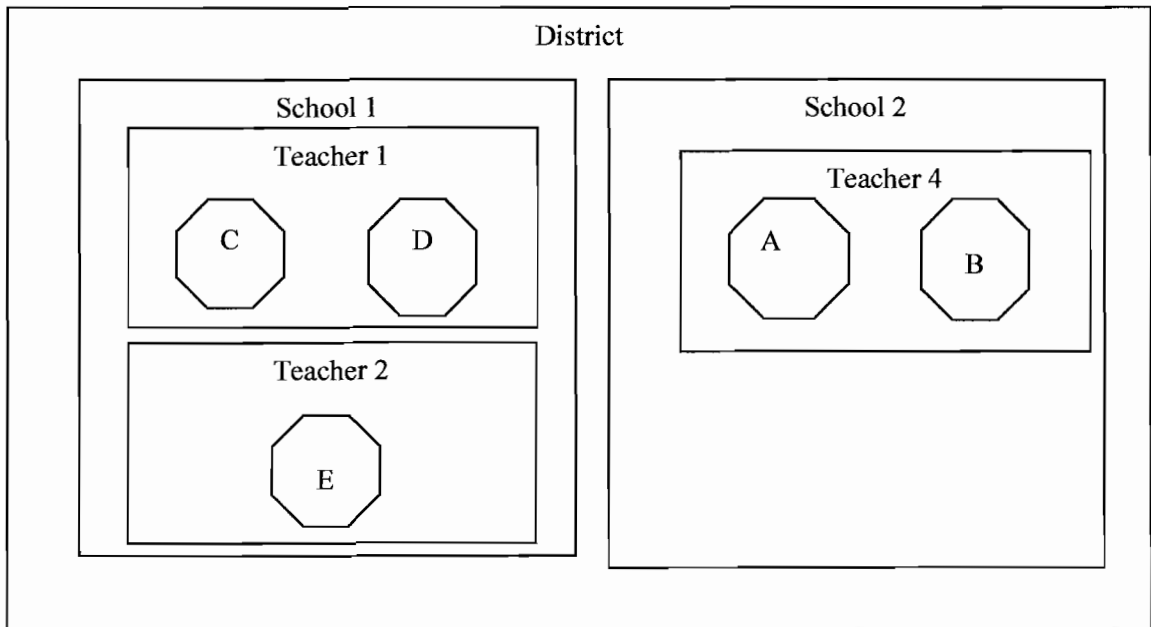
Figure 1

Illustration of Multi-case Design.
Note: C = Carl, D = Diane, E = Eric. These three students were in School 1. Teacher 1 participated in the RTI study, Teacher 2 did not. A = Albert, B = Becky. These two students were in School 2. Teacher 4 participated in the RTI study.

In Figure 1, the comparisons between Students A and B and Students C and D are considered literal replications because the students of similar skill levels receive the same instruction from the same teacher in the same class. The comparison between Teacher 1 and Teacher 4 is considered a theoretical replication because both teachers receive the same training from the district's RTI study but implement their RTI in different schools. The comparison between Teacher 1 and 2 is also considered a theoretical replication because Teacher 1 receives training from the district RTI study and Teacher 2 does not.

Students. After the principals and teachers consented to participate in the study, I asked teachers to identify three appropriate students using the selection criteria described

above. Next, I obtained informed consent from parents and assent from students to observe them in class, to sit in on school-based meetings in which their cases were reviewed, and to gain access to these students' reading assessment data, SST Plans, and IEPs (if applicable).

The five participating students were Caucasian and native English speakers. None of the students had an individualized educational plan (IEP), behavior plan, or had been retained in a grade. Two of the five students were girls, and all were recommended by their teachers to receive Tier 3 interventions because their fall screening scores in reading were at or below the 20th percentile on district norms (see Table 6).

Preparing for Data Collection

I used case study design methodology to organize my data collection and data analysis process. Integrating data from multiple sources, I described what the implementation of the identified components looked like in practice. This approach was modeled after the multi-case study design described in Yin's *Case Study Research* (2003). Yin (2003) suggests preparation for data collection in the case study should include (a) specifying the data to be collected, (b) scheduling the data-collecting tasks, (c) composing interview questions, and (d) describing the instruments for data collection. The preparation of data collection helped organizing multiple strands of data-collection tasks and kept them aligned with the purpose of the study.

Specifying the data to be collected. Yin (2003) listed six types of evidence to be collected in the case study. I used five of the six types of evidence to describe the design and delivery of instructional support that the teachers provided for the target students under the RTI Model. I excluded participant-observation because it was incompatible

with the design of this study; I wanted to document the RTI implementation process as it was at the school, and if I had included myself in the process as a participant, it would have altered the RTI delivery model. The evidence types included in my study were (a) archival records, (b) documents, (c) physical artifacts, (d) direct observation, and (e) interviews. Table 7 provides examples for each type of evidence.

Scheduling the data-collection tasks. All the data-collecting tasks were scheduled within an eight-week timeframe (see Table 8). I observed each student-teacher dyad for one week, during the designated reading time. I followed the students to observe the tiered reading instruction (tier 2 and 3) they received from different teachers. I stayed for the duration of the reading time. I observed grade-level meetings in which these students' cases were discussed. I interviewed the teachers and administrators before and after the classroom observation.

Composing interview questions. I interviewed the participating educators using standardized questions to allow for comparison of answers. Each interview question was linked to one of the six identified components from the literature review (see Table 9 and 10). Each question was asked twice: first, from a general class-wide perspective; second, from the perspective of a specific student. For example, Question #5 of the first interview was, "What instructional support do students with severe skill deficits under the RTI model receive?" and Question #6 was, "How do you deliver the instructional support to the students with severe skill deficits?" In the second interview, I asked Teacher 1, "What instructional support does Carl receive?" and, "How do you deliver instructional support to Carl?" Question #5 was related to the content of instructional support within the multi-tiered intervention model, and Questions #6 was related to the method of

instructional delivery. These two questions were linked to two of the identified

components of RTI implementation: Tier 1 Instruction and Tier 2 and Tier 3

interventions.

In this study, students received instruction from different teachers based on the

different tiers. For example, in School 1, Carl and Diane received Tier 1 Instruction from

Teacher 1 (their homeroom teacher), Tier 2 instruction from Teacher 3 (a special

education teacher), and Tier 3 instruction from Teacher J (an instructional assistant). Eric

also received Tier 1 instruction from his homeroom teacher (Teacher 2) and Tier 3

intervention from Teacher J, but he received Tier 2 instruction from a general education

teacher (Teacher 1) rather than a special education teacher.

Classroom observation measures. I used the Classroom Observation Electronic

System (hereafter, referred as *e-COVE*) in classroom observations. The *e-COVE* is a

computer-based direct classroom observation tool developed by Tenny (2007). I used *e-COVE* to track ten mutually exclusive target behaviors. The first five targeted behaviors

were teacher-led instructional activities, specifically on teaching phonemic awareness,

phonics, fluency, vocabulary, and comprehension. The next three targeted behaviors were

student-led instructional activities, specifically independent seatwork, silent reading, or

reading aloud to others The last two targeted activities were classified as housekeeping,

discipline, transition, and instructional activities unrelated to reading.

The observation was conducted from the perspective of the targeted students. In

other words, if the teacher was working with targeted students on phonics in a small

group setting while the rest of class was doing independent seatwork, that instructional

time would be coded as teacher-led instructional activities on phonics. I used *eCOVE* to

count cumulative minutes of targeted behaviors during instructional sessions. I chose *eCOVE* because its timer allowed direct measure of the duration of target behaviors, in contrast to momentary time sampling or frequency counts of target behaviors, which would provide only estimates based on observed scores.

After observing each instructional session, I filled out the *Classroom Observation Checklist* (COC) (see Appendix D). This checklist was used to document effective teaching behaviors during reading instruction related to the component of high quality instruction; it was adapted from the *English Language Learner-Classroom Observation Instrument Checklist*, developed by Gersten, Baker, Hagger and Graves (2005). The original checklist included six modules. I deleted one module that focused on additional language support for English Language Learners, which was not the focus of my study. The other five modules were adopted verbatim.

The COC is comprised of 24 items; each item rated on a scale between zero and two, with ratings of zero indicating *not effective*, ratings of one indicating *partially effective,* and ratings of two indicating *very effective*. According to Gersten, et al. (2005), observers should supplement the quantitative rating scales with qualitative field notes and should stay for the duration of the reading block to capture the totality of reading instruction. The developers of the checklist do not provide cut scores for determining effective teaching. Instead of calculating the cut score between effective and ineffective schools, the authors provide an average numerical score on the instrument and supplement this information with qualitative descriptions about typical classroom instruction for two effective teachers and one ineffective teacher.

Following Gersten and colleague's example, I calculated the mean and range of scores obtained from the COC during the five-day observation, and I also included qualitative field notes to depict how teachers delivered instruction and intervention and evidence related to the response of targeted students to the provided instruction and intervention.

Data Collection

The duration of the case study was eight weeks (October-November, 2008). During these eight weeks, I was embedded in these classrooms shadowing the targeted students for the duration of their reading time. I also interviewed teachers and administrators, observed school-based meetings, reviewed students' progress monitoring data, and collected documents from various sources (e.g. school schedules and minutes of meetings) to describe how the teachers planned and delivered intervention to these struggling readers.

Classroom observation. I conducted observations of three classrooms to gather information on the nature of students' instructional experiences, including the implementation of core curriculum instruction, differentiated instruction, and use of research-based instructional material. Each class was observed for a week. At the first site visit, I obtained a classroom schedule and school map for the target classes. On the scheduled days of direct observation for each class, I entered the classroom before the class started. I quietly set up my observation station, per the instructions of the classroom teacher, to observe the target student and teacher unobtrusively. The purpose of observing the same classroom for an entire week was to capture the continuity of the instruction and the range of instructional activities that occurred in class.

Observation of school-based meetings. To gather information regarding evidence-based decision making, I observed school-based meetings, including Student Support Team (SST), RTI, and Grade-level meetings, in which the targeted students' cases were discussed. The purpose of observing these school-based meetings was to evaluate evidence-based decision-making processes. The observation focused on: (a) the purpose of the school-based meeting, (b) the case being discussed, (c) the decision-making process, (d) the participants and their level of participation, (e) the conclusion reached in decision-making, and (f) the plan to implement the prescribed instructional support. These foci were selected to highlight three types of best practices identified in the literature review as resulting in improvement of student learning outcomes:

1. Teachers formed a learning community, sharing their expertise with each other to solve a common problem.

2. Teachers focused on the student work through assessment.

3. Teachers changed instructional practice, based on the feedback gained from the assessment, to improve students' learning outcomes (Fullan, 2000).

Reviewing archival records, documents and artifacts. Archival records, documents and artifacts were reviewed to substantiate the presence of the implementation of the six RTI components. Examples of these data sources can be found in Table 11.

Compiling school reports. The case study method employs multiple data collection strands and yields large amounts of data. I followed the data reduction procedures advocated by Miles and Huberman (1994) to create a detailed school report for each participating school. The purpose of synthesizing data was to provide independent reviewers a focused, organized and condensed report documenting the

experiences students received related to their participation in the school's RTI process. (Miles & Huberman, 1994). Because an important expectation of RTI is the improvement of students' learning outcomes, individual students' progress monitoring data were also included in the school reports to provide the data needed to evaluate the effectiveness of instruction for each targeted student.

To ensure the data gathered covered all six identified components, I created a matrix to illustrate the relation between the components and the sources of evidence. Next, I specified the items to be included in school reports. These items were aligned with the data sources specified in the revised RTI Assessment Rubric. Finally, the school reports (see Appendix E) were reviewed by an independent editor to reduce possible ambiguity and value-laden statements.

Use the Rubric to Evaluate the Implementation of RTI in Two Schools

After I made slight revisions to the rubric based on expert reviewers' feedback and completed the detailed case study reports, I recruited two independent raters to use the rubric to evaluate RTI implementation at case study schools. The first independent rater was a project manager for a National Model Demonstration Center on RTI. She holds a Ph.D. in Educational Leadership from the University of Oregon. She was chosen for her expertise in all aspects of RTI implementation. The second independent rater holds a doctoral degree from the Department of Special Education and Clinical Science at the University of Oregon. She was chosen for her extended classroom experience as a special education teacher and her experience in program evaluation and qualitative research. I selected two raters of different backgrounds to minimize the influence of possible shared bias.

I provided the independent raters with two copies of the revised RTI Assessment Rubric, and school reports documenting information gathered at my case study schools (see Appendices C and D). The raters were instructed on how to read and interpret the overall school report for the target students and asked to read the report independently and in its entirety before rating the quality of implementation. I also reviewed the components and use of the RTI assessment rubric with them and answered any questions they had. For each component, raters indicated either 0 for *not implemented*, 1 for *partially implemented*, or 2 for *fully implemented* for each component. After ensuring that the raters understood the purpose of the school reports, steps for reviewing RTI implmentation for each target student, ad how to use the RTI rubric, raters completed the rubric independently.

Validation of the RTI Assessment Rubric

"Validity refers to the degree to which evidence and theory support the interpretation of test scores entailed in the uses of tests" (AERA, APA & NCME, 1999; p. 9). Because the RTI Assessment Rubric is intended to be used as a tool to provide an evaluative judgment about the implementation of the RTI approach in a school—a use with potentially high-stakes results—it is important that the tool itself undergo rigorous evaluation as part of its development. Thus, in this study, I followed the logic model recommended by Kane (1992; 2006) to integrate sources of evidence to substantiate a validity argument. In addition to drawing evidence from the content review, the case study, and the scoring of the rubric, I conducted a focus group to gather evidence about the relevance and ease of using the rubric from the perspective of school teachers and

administrators. The focus group also provided additional information about the accuracy

and adequacy of the school reports written based on the case study.

Focus group study. Principals, general education teachers, special education

teachers, school psychologists, and reading specialists from each of the case study

schools participated in focus groups. Focus groups were conducted in single-school

groupings (see Table 12). Teacher 1 and Teacher 4 from the case study both took part in

the focus groups, which took place the last week of April and first week of May, 2009. In

each focus group, I first asked the teachers, school psychologist, and principal to

independently evaluate their own implementation of RTI, using the RTI Assessment

Rubric. They graded their implementation on each of the six identified components on a

three-point scale: 2 meant fully implemented; 1, partially implemented; and 0, not

implemented at all. General education teachers provided evaluation of the

implementation at their own grade level. Special education teachers, school

psychologists, reading specialists. and principals provided evaluation at the school level.

The participants of the focus group all individually completed the rubric by evaluating

their or their school's implementation of these six identified RTI components and then

discussed as a group.

Next, focus group participants discussed the evaluation of these six RTI

components, one by one. They shared their rating and the reasons for giving that rating.

Next, each school's participants had a brief discussion and reached a consensus on the

composite score for school-wide implementation of that particular component. The

teachers had the option of changing their ratings, but none of them did. For example, in

School 1, Grade 1 and 2 teachers both rated the implementation of RTI at their grade-

level as fully implemented, Grade 5 teachers rated the implementation of RTI at the

Grades 4 and 5 as partially implemented. The principal and the school psychologist gave

an overall school-level of RTI implementation a rating of "partial implementation"

because the variation of implementation across different grade levels. Each focus group

lasted approximately 90 minutes. The focus groups were audio-taped and transcribed for

data analysis.

Data Analysis

Data were analyzed to build a validity argument about the use of the assessment

rubric to judge the quality of RTI implementation in elementary schools. In this study,

validity was defined as "the degree to which evidence support the interpretation of

[rubric] scores entailed by proposed uses of [rubric] scores" (AERA, APA, NCME, 1999,

p. 9). This definition echoes the one given by Messick (1989) that Shephard (1993)

described as the "most cited authoritative reference on the topic" (p. 423) Messick (1989)

defined validity as:

> an integrated evaluative judgment of the degree to which empirical evidence of
>
> theoretical rationale support the *adequacy* and *appropriateness* and *inferences* and
>
> *actions* based on test scores or other modes of assessment. (p. 13, italics in
>
> original)

These two definitions highlight a shift toward the more unified and integrated view of

validity that is prevalent today. Validity is more than demonstrating that test items are

corresponding to the cells of a matrix of test specifications or demonstrating that scores

on a test are correlated to other measures (Linn, 2002). In developing the RTI Assessment

Rubric, I followed Kane's (1992; 2006) advice to integrate different strands of evidence

to develop a plausible and coherent argument to represent the proposed interpretation and use of the scores in the test development process. Although the RTI Assessment Rubric is not a test *per se*, it is intended to be used to provide an assessment of the quality and extent of RTI implementation in elementary schools. Thus, this approach is warranted.

To answer my first research question (To what extent does the RTI Assessment Rubric appropriately measure the construct of RTI implementation?), I gathered experts' judgments as content related validity evidence for the validation of the rubric (Downing & Haladyna, 2006). First, I calculated the number of agreements and disagreements each expert indicated on the rubric regarding the proposed operational definitions and descriptors. Second, I calculated the number of operational definitions and descriptors for which all three experts reached unanimous agreement. Third, I compiled the experts' feedback and made revisions to the RTI Assessment Rubric to bring it in closer alignment to expert content reviewers' judgments about the essential components of RTI implementation. Revisions included clarifying the wording and deleting value-laden terms such as *appropriate* and *sufficient*. In addition, I removed set time limits for instruction in Tier 2 and Tier 3 interventions.

To answer my second research question (What evidence substantiates the reliability of the RTI Assessment Rubric for evaluating the implementation of each identified RTI component?), I followed Miles and Huberman's (1994) recommendations for triangulating data from multiple sources. Sources included: interviews, observations, and document review. The matrix (see Table 7) illustrates the connections among the data, data sources, and the identified RTI components. Following Yin's (2003) recommendation that researchers confirm their findings with key informants, I conducted

member checking in two different ways: during the second interviews with participating teachers and administrators and by having them confirm the accuracy and completeness of the school reports compiled as part of the case study.

To provide evidence for the stability of the assessment rubric, I calculated the inter-rater agreement between the two independent raters who applied the RTI Assessment Rubric to the case study data presented in the detailed school reports. The two raters evaluated the level of implementation of each identified component using the revised rubric. They scored the school reports independently and had no interactions with each other. I also analyzed Teacher 1's and Teacher 4's self-evaluation of implementation by reviewing their completed rubric and compared the teacher's ratings with the independent raters.

To answer my third research question (What evidence substantiates the sensitivity of the RTI Assessment Rubric for measuring qualitative differences in schools' RTI implementation?), I compared the data from my case study with the scores on the RTI Assessment Rubric given by the independent raters, Teacher 1, and Teacher 4. In addition, I evaluated the differences in score on the RTI Assessment Rubric obtained by different grade-levels and different schools, comparing this information to the feedback from participants in the focus groups.

To answer my fourth research question (Do differences in RTI implementation, as measured by the RTI Assessment Rubric, correspond with different student outcomes?), I compared the rating from the RTI Assessment Rubric with the scores from participating students' progress monitoring and benchmark assessment data.

Summary

The purpose of this study was to develop and validate an RTI Assessment Rubric appropriate for use by researchers and school personnel interested in evaluating how well a school has implemented RTI. The instrument development process included creating an initial draft of the rubric based on the synthesis of relevant research literature, having experts conduct a content review, revising the rubric based on this review, conducting a planned case study to gather evidence about the use of the RTI Assessment Rubric, and holding focus groups to gather additional evidence about the reliability and validity of the rubric for its intended use. These data are analyzed to answer the four research questions, representing different facets of reliability and validity evidence, in the process articulating a validity argument to justify the use of the scores of the RTI Assessment Rubric to judge the quality of the implementation of RTI at the elementary school level.

CHAPTER IV

RESULTS

In this chapter, I report findings from the data analysis related to the four research

questions:

1. To what extent does the RTI Assessment Rubric appropriately measure the

construct of RTI implementation?

2. What evidence substantiates the reliability of the RTI Assessment Rubric for

evaluating the implementation of each identified RTI component?

3. What evidence substantiates the sensitivity of the RTI Assessment Rubric for

measuring qualitative differences in schools' RTI implementation?

4. Do differences in RTI implementation, as measured by the RTI Assessment

Rubric, correspond with different student outcomes?

Results for Question 1: Experts' Review of the RTI Assessment Rubric

The RTI Assessment Rubric was comprised of six identified components of RTI.

Within each component, I provided an operational definition and three sets of descriptors,

one for each level of implementation. Thus, each of the three RTI experts had 24

opportunities to state whether he or she agreed with the provided operational definition

and descriptors. These three content experts made their judgments independently and did

not meet to discuss or reconcile their differences. Results across each component are reported in Table 13.

Of the 24 provided definitions and descriptors, Dr. Ed Shapiro agreed with 18 (75%), Dr. Doug Marston agreed with 22 (92%), and Dr. Teri Wallace agreed with 23 (96%). In all, there were 63 agreements and 9 disagreements. The nine disagreements were scattered among 5 of the six components. The percentage of agreement across the components ranged from a low of 75% (9/12 on Tier 2 and Tier 3 interventions) to a high of 100% (12/12 on Progress Monitoring), with a mean agreement of 87.5%.

The three RTI experts provided explanations when they disagreed with the provided definitions or descriptors on the rubric and offered suggestions for revising the instrument. Dr. Shapiro disagreed on the operational definition for evidence-based decision-making because it included teacher judgment. He disagreed on the descriptors for all three levels of implementation on Tier 2 and Tier 3 Interventions because these descriptors specified acceptable ranges of instructional minutes teachers should spend on Tier 2 and Tier 3 Interventions weekly. He disagreed with the inclusion of such criteria, explaining that the appropriate time for different intervention programs varied from grade to grade and program to program. Finally, he disagreed with the wording "sufficient [professional] training and support" on the descriptors for the levels of fully implemented and partially implemented in the component of organizational support because the provided descriptors did not specify what constituted *sufficient.*

Dr. Marston disagreed on the operational definition of *Tier 1 Instruction* because the wording was unclear. He suggested including research-based instruction strategies in the definition and defined "targeted effective teaching behaviors." He asked for

clarification or revision on the following phrases and sentences:

1. On *universal screening*: "Data are consistent with appropriate unit of analysis (teacher roster, grade-level placement)."

2. On Tier 2 and Tier 3 instruction: in place of the words "on-task," he suggested using the term "engaged time."

3. On progress monitoring, he suggested adding the following descriptor: "Teacher administered repeated measures of alternate forms for students receiving additional instructional support weekly or biweekly for at least three times."

Dr. Wallace disagreed on the operational definition for Tier 1 Instruction, because it implied Tier 1 Instruction was a "one-size-fits-all" program. She recommended the following changes:

1. On Tier 1 instruction, she suggested (a) using the term "core curriculum" to describe "Tier 1 Instruction," (b) including writing as part of the core curriculum, (c) considering an alternative time requirement for kindergarten programs because some of them were only half-day, and (d) adding the description "expected behaviors to instructional program after sufficient amount of instruction time was provided" in the rubric.

2. On Tier 2 and Tier 3 interventions, she suggested (a) including the terms "evidence-based intervention" and "implemented with fidelity," and (b) adding, "expected behaviors to instructional program after sufficient amount of instruction time was provided."

3. On progress monitoring, she suggested changing the word "skill" to "performance."

4. On Organizational Support, she posed a question: "How do we measure fidelity of implementation on the Organizational Support?"

Based on the experts' feedback, I made the following revisions:

First, I honored Dr. Shapiro's suggestion by deleting teacher-judgment from the operational definition of evidence-based decision-making. However, I did not explicitly state that teacher judgment should be banned from the evidence-based decision-making process. Dr. Shapiro took a conservative stand in interpreting evidence-based decision-making, and suggested that inclusion of teacher judgment would undermine the tenet of evidence-based decision-making. An equally plausible argument could be made for why teachers' professional judgment should be included in the evidence-based decision-making process. From this perspective, one might argue that just as doctors use their professional judgment to diagnose patients based on empirical evidence, teachers can use their professional judgment to diagnose students' skill deficits based on empirical evidence. Thus, I did not explicitly prohibit the inclusion of teacher judgment.

I deleted the value-laden terms from the rubric or added an expected outcome as the standard of judgment. For example, I revised the operational definition of Tier 1 instruction as

> the comprehensive core reading instruction provided to all students in general education classes for a sufficient amount of time to meet grade-level instructional goals. In Tier 1 Instruction, teacher teaches all "five big ideas" as specified in NRP or relevant ones to specific grade-level, using evidence-based teaching strategies. (RTI Assessment Rubric, 2009)

Whenever possible, I provided measurable criteria in the rubric to help users of the

rubric make objective decisions. In the absence of quantifiable criteria, I provided descriptions containing enough details to assist in differentiating full, partial and non-implementation of specific components. For example, one of the descriptors for full implementation of organizational support was revised as follows:

Teachers receive multiple sessions of professional training and on-going support to implement research-based instruction, assessment and evidence-based decision-making. The professional training is focused and interactive. Teachers collaborate with colleagues to identify students' needs and implement tailored interventions to address the needs.

Its corresponding descriptor for partial implementation stated that

Teachers receive only initial professional training and have no on-going support to implement research-based instruction, assessment and evidence-based decision-making. Professional collaboration occurs in isolated incidences. The extent of collaboration is limited to issues of logistics and role responsibilities.

Its corresponding descriptor for non-implementation stated that

Teachers receive no professional training and rarely collaborate.

Revision following expert review is intended to be more than editing or rephrasing words; it is intended to ensure the validity of the content included on the rubric (Downing & Haladyna, 2006). It should be noted that the experts did not suggest additional components to be added to the rubric nor disagreed with the components provided, indicating support for the overall rubric. I revised the rubric following expert reviewers' feedback prior to its use by the independent raters. The revised rubric can be found in Appendix B.

Results for Question 2: Reliability of the RTI Assessment Rubric

The RTI Assessment Rubric was used by two independent raters to evaluate two

schools' RTI implementation. Each rater gave each school one rating score for the

implementation of each identified RTI component. In all, they had 12 opportunities for

agreeing or disagreeing with each other's appraisal. Both raters were in agreement that

School 1 was fully implementing all of the components of RTI and that School 2 was fully

implementing universal screening and partially implementing Tier 1 instruction, Tier 2 and

Tier 3 interventions, evidence-based decision making and organizational support. The two

raters only disagreed on whether Progress Monitoring in School 2 was fully or partially

implemented (see Table 14). They agreed on 11 of the 12 occasions (see Table 14). The

inter-rater agreement was calculatied by dividing the number of agreements with the

number of possible occasions. The percentage of inter-rater agreement was 92%.

Additional reliability evidence was gathered through the focus group, during

which Teachers 1 and 4 conducted independent self-evaluations on the implementation of

RTI at their grade-level using the rubric. Teacher 1 reported that she fully implemented

RTI across all the six identified components, which is in agreement with the appraisal of

both raters. Teacher 4 reported that the implementation of universal screening at her

school was at the level of full implementation and the other five components were at the

partially implemented level. She marked progress monitoring at the level of partial

implementation because there was often a delay in receiving progress monitoring data

from the school psychologist. Her appraisal is in 100% agreement with one rater and in

92% agreement with the other rater.

Results for Question 3: Sensitivity of the RTI Assessment Rubric

The sensitivity of the rubric can be determined by juxtaposing the school reports obtained from the field study and the scores of the rubric. The purpose of this comparison is to judge whether the scores on the rubric reflect qualitative differences in the implementation between these two schools. I describe the implementation of the six components at these two schools and report their scores from the RTI Assessment Rubric.

Universal screening. *Both schools screened students' reading performance using a battery of research-based, curriculum-based measures, easyCBM (Alonzo, Tindal, Ulmer, & Glasgow, 2006) three times a year, in fall, winter and spring. Two of the screenings occurred outside the timeframe of this study (October-November, 2008), but subsequent visits confirmed that the schools indeed administered the other two benchmark assessments in the winter and spring. Both schools used the district benchmark assessments (easyCBM measures) in fall to create instructional groups for Tier 2 and Tier 3 interventions. In addition to easyCBM, School 1 used the benchmark assessments of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) in their fall screening (see Tables 15 and 16). Teachers used the district recommended cut-score, 20th percentile, to identify students needing Tier 3 interventions. Based on these findings, both schools were graded as fully implementing their universal screening.*

Tier 1 instruction. School 1 provided Grade 2 students 65 minutes of Tier 1 instruction daily. Sometimes Teacher 1 and Teacher 2 (both general education teachers) extended Tier 1 instruction into the story and writing time. On the week in which I conducted my classroom observation, both teachers provided more than 400 minutes of Tier 1 instruction, which was within the district mandate of 60-90 minutes daily Tier 1 instruction. Teacher 1 spent 63% of Tier 1 instruction time teaching the "five big ideas;" Teacher 2 spent 70% of the instructional time directed toward these domains.

As mentioned in the methods section, the Classroom Observation Checklist was comprised of 24 items. For my study, these items were rated using a three-point scale (0 = not effectively at all, 1 = partially effective, and 2 = very effective). The highest possible score on this instrument was 48. In School 1, scores for both classrooms were between 39- 44 (see Table 17), indicating an average score ranging from partially to very effective across all 24 items.

During Tier 1 instruction time at School 1, Carl and Eric were given preferential seating. With their teacher's frequent verbal redirect and positive praise, they maintained on task behavior at least 80% of the time (see Table 17). Carl and Eric could read the grade-level passage from the core reading curriculum when working with teachers one-on- one or in small groups, but they could not match the reading pace when reading in chorus with the entire class. During choral reading, they often lost their place, and their voices faded. Diane was not observed for her independent read-aloud and choral reading because of a three-day absence.

In School 2, Teacher 4 (a general education teacher) provided 60 minutes of daily Tier 1 instruction to her class. However, Albert and Becky did not receive Tier 1

instruction during the allotted Tier 1 instructional time because they had to go to Teacher 5 (a special education teacher) to receive their Tier 3 interventions. Teacher 4 provided some Tier 1 instruction to Albert and Becky outside the scheduled time, which included Teacher 4 reading aloud chapter books and conducting literary discussions with the class and guiding students to generate story starters. In both occasions, Albert and Becky appeared to participate in the instructional activities.

Based on these findings, the raters graded the implementation of Tier 1 instruction at School 1 as full implementation, because it met all the criteria in the rubric, and School 2 as partial implementation, because from Albert and Becky's perspective, Tier 1 instruction was supplanted by Tier 3 interventions.

Tier 2 and tier 3 interventions. The district required all students to receive 30 minutes of daily Tier 2 intervention, and expected students who scored at or below the 20th percentile to receive 60 minutes of Tier 3 interventions weekly in addition to their other instruction and interventions. Each of the two case study schools implemented this expectation with different levels of success.

School 1 created a cascade of seven intervention groups for Tier 2 intervention. Students were assigned to an individual group based on the skill needs of the student. Carl and Diane were assigned to Group 1 (the lowest group), and Eric was first assigned to Group 2 and then moved up to Group 3 because of his progress. To ensure low-performing students did not miss reading and math instruction because of receiving Tier 3 intervention, the Tier 3 interventions in School 1 were scheduled at the time students would have had P.E., music, library or foreign language. However, the teachers in School 1 did not eliminate students' access to specialists or what students perceived as fun activities. Tier 3

students in School 1 missed one of the two opportunities with any specialists, but never missed both opportunities because of receiving Tier 3 interventions.

On the week of the classroom observations at School 1, Teacher 1 provided Carl and Diane 30 minutes of Tier 2 instruction daily, and Teacher 3 provided Eric 30 minutes of Tier 2 intervention daily. Teacher 1 and Teacher 3 spent approximately 81% of Tier 2 instruction time teaching the "5 big ideas of reading." The instructional assistant, Teacher J, provided 38 minutes of Tier 3 intervention to Carl and Diane and 45 minutes of Tier 3 intervention to Eric. Students received Tier 2 instruction in a group of 6-8 students and Tier 3 instruction in a group of 3-4 students. They appeared on task more than 90% of the time during both observations. For Tier 2 and Tier 3 Interventions at the School 1, the average scores on the Classroom Observation Checklist ranged from 40-44 out of 48 possible points (see Table 18).

It should be noted, however, that there were differences in terms of aligning the instructional content between Tier 1 and Tier 2 instruction across teachers at School 1. During Tier 2 instruction, Teacher 1 pre-taught Eric the key vocabulary and phonics patterns and conducted choral reading using the same grade-level materials used in Tier 1 instruction. This linkage appeared to be beneficial to Eric, as he was successful in reading the vocabulary words, identifying their meanings and reading the passages during Tier 1 instruction. For Teacher 3, who worked with Carl and Diane during Tier 2 interventions, this linkage between Tier 1 instruction and Tier 2 intervention was not apparent. Although she used the intervention program designed to supplement the Teacher 1's selected core reading curriculum, the observation revealed that Teacher 3 made no direct connections between the phonics patterns, spelling words, or passages Teacher 1 taught in Tier 1

instruction and she herself covered in Tier 2 intervention. Consequently, it appeared that Carl and Becky could not readily generalize the skills from Tier 2 interventions to Tier 1 instruction.

At School 2, Teacher 4 worked with multiple groups in rotation during Tier 2 intervention time. The rotation of instructional groups appeared to be improvised. While the teacher worked with one group, the other students were assigned to read books silently. But Albert and Becky, instead of reading books at their instructional level, drew, wove, flipped through pages of books, or chatted with peers until redirected by the teacher. It should be noted that Teacher 4 had difficulty keeping track of when students would be removed for additional instruction and voiced a concern about her "fragmented schedule" because she rarely had a time when all of her students were together in her class during the designated reading block. Finding time to teach her entire class new vocabulary words or new comprehension skills was logistically challenging for Teacher 4.

On the week of observation, Teacher 4 of School 2 provided Albert and Becky at least 20 minutes of small group differentiated reading instruction four days of the week. This teaching was accomplished with the assistance of an instructional assistant who came in to read with Albert and Becky twice for 15-20 minutes each. Teacher 4 used first grade basal readers for Albert and Becky's group with an instructional emphasis on phonics and fluency. Albert and Becky appeared on task when they were working with Teacher 4 and the instructional assistant, but they sometimes appeared off task in the assigned silent reading time. As a result, the percentage of on task time and their scores on the Classroom Observation Checklist varied widely. The range of scores on the Classroom Observation Checklist for Teacher 4 and the instructional assistant was between 32 - 42, with an average

score of 37 (see Table 18). These scores suggest that Teacher 4's Tier 2 intervention was inconsistently provided.

At School 2, the special education teacher, Teacher 5, provided Albert and Becky 60 minutes of daily reading instruction in a small group ($n = 4$), using a research-based reading intervention program. Teacher 5 followed the program closely for decoding, oral reading fluency, and spelling. Albert and Becky were given multiple opportunities to respond in chorus or independently. They appeared on task over 90% of the time, answered questions, and finished tasks correctly with minimal teacher correction. The range of scores on the Classroom Observation Checklist for Teacher 5 was 42-44, resulting in an overall classification of "effective" (see Table 18). However, it should be noted that there were no explicit linkages observed between the Tier 2 and Tier 3 interventions received by Albert and Becky and their Tier 1 instruction.

Based on these findings, the raters graded school 1 as fully implementing Tier 2 and Tier 3 interventions because it met all of the criteria; they graded school 2 as partially implementing because of the inconsistency in teaching and student engagement.

Progress monitoring. Both case study schools used the district reading assessment (easyCBM Reading Measures) for progress monitoring. Both schools went beyond the district mandate and progress monitored students who scored at or below the 35th percentile for the district on the fall assessment. For Grade 2, the progress monitoring measure was a one-minute passage reading fluency assessment where the number of words read correctly in a minute was counted. The test was administered individually by a hired instructional assistant who also was responsible for entering scores into the computer and reporting results back to teachers and administrators.

Both schools designated an instructional assistant to collect the school-wide program monitoring data and report back to the teachers. The instructional assistants coordinated with teachers to schedule for progress monitoring; administer the tests, following the standardized protocols and scoring, enter the data in the district website, and deliver a hard copy to classroom teachers. By and large, progress monitoring was conducted following the pre-determined two-week intervals. However, one of the raters graded school 2 with partial implementation because of the occasional delay in testing and data reporting. The other rater gave School 2 a rating of full implementation.

Evidence-based decision-making. All teachers stated that they used the data from easyCBM progress monitoring measures as the primary data source for decision-making. Table 19 documents the implementation of evidence-based decision-making by reporting (a) data being used to make decisions, (b) teacher input, (c) the process of decision-making, and (d) decisions being made.

In School 1, each grade instructional team met every 4-6 weeks to review student assessment data and use those data to adjust instructional groupings in Tier 2 and Tier 3 interventions. At the grade-level meeting observed, two general education teachers, a special education teacher, a facilitating teacher, a school psychologist, and the principal were in attendance. The team compared students' progress against the expected performance (i.e., the aimline) and determined if students were making adequate progress. Based on that determination, the team decided whether or not to modify the student's intervention program or refer students for a full evaluation. In the meeting, the team followed the Grade-level meeting protocol, a document that was generated by teachers in this school and explicitly stated the shared consensus of the procedure and decision rules

for reviewing student data and making instructional decisions. After the instructional team reviewed students' data, they modified the instructional groups in Tier 2 interventions. They decided that Carl and Diane should remain in Group 1 (the lowest group) because both had progress below the aimline. Carl was referred for a special education evaluation because of limited progress and the results of the evaluation were still pending. Diane had not been referred to eligibility evaluation because of her frequent absences and tardiness. Eric was moved from Group 2 to Group 3 in Tier 2 intervention because his progress was on the aimline (See Figures 2-4, p. 86. More information about these figures were provided in page 86.)

In contrast, at School 2 only the two general education teachers attended the grade-level meeting. Although teachers examined student data, they did not look at the graphs comparing student progress against the aimline (i.e., the expected performance). At the grade-level meeting observed, Teacher 4 and her colleague (another general education teacher) admitted that they did not have explicit decision rules that were used uniformly across the school. They both had questions about the scores on the easyCBM progress monitoring measures. They discussed their shared disagreement with the reading specialist, who was not present at the meeting, on the appraisal of student progress.These two general education teachers and the reading specialist disagreed on which students should be included in the Tier 2 intervention group taught by the reading specialist. The reading specialist did not want to exit one of the students in her Tier 2 intervention group because she worried the student might not sustain her progress if the student exited from her Tier 2 intervention group too soon. Nor did the reading specialist want to accept a new student whom Teacher 4 considered in much greater need of support from the reading specialist

because the reading specialist wanted to maintain the group size.

Finally, Teacher 4 and her colleague reviewed all students' progress monitoring data. Neither teacher was satisfied with Albert's and Becky's progress. However, the teachers decided to keep Albert and Becky in the instructional group where they were because they were receiving daily hour-long intensive reading intervention from the most qualified interventionist in the school (Teacher 5, a special education teacher with 29 years of teaching experience). Both teachers agreed that Albert's and Becky's instructional plan would remain unchanged.

Based on these findings, the raters were in agreement in rating School 1 fully implemented and School 2 partially implemented in evidence-based decision-making.

Organizational support. During data collection, I gathered evidence of organizational supports and sorted them into four different categories: (a) instructional leadership, (b) administrative support, (c) professional training and on-going support, and (d) professional collaboration. Examples of instructional leadership included principals allocating resources to purchase research-based programs and to implement small group instruction and progress monitoring. Examples of administrative support included hiring instructional assistants, scheduling progress monitoring, or purchasing equipment for implementing effective instruction. Examples of professional training included the professional training and technical support the school and school district provided to enable teachers to implement research-based instruction and progress monitoring. Examples of professional collaboration included descriptions of how teachers collaborated formally and informally.

Both schools allocated resources to purchase research-based reading programs and to implement small group instruction and progress monitoring. They both designated an instructional assistant to administer the progress monitoring measures for the entire school and to manage the student progress monitoring data. Teachers received initial training on the newly adopted reading curriculum and on progress monitoring measures. A reading specialist was assigned to each school to consult with teachers about implementing new reading curriculum and address issues related to progress monitoring. However, these two schools differed as to how instructional leadership and professional collaboration manifested in organizational support.

In School 1, the principal, two general education teachers, a special education teacher, a facilitating teacher, and a school psychologist all participated in grade-level team meetings to review students' assessment data and discuss how to modify the current instruction program to meet students' instructional needs. They had explicit rules and procedures on how to conduct grade-level data meetings. In School 2, only two general education teachers attended the grade-level team meeting to review student assessment data. The specialists chose when and where they would provide student interventions. As a result, Teacher 4 had a fragmented schedule, and Albert and Becky did not receive Tier 1 Instruction because of the schedule conflict. At School 2, Teacher 4 and the other Grade 2 general education teacher were not familiar with the intervention programs used by the reading specialist and special education teacher. A consensus on the decision rules for using assessment data to make instructional decisions was still emerging.

Based on these findings, once again, the two raters were in agreement in rating School 1 fully implemented and School 2 partially implemented.

By and large, School 1's instructional delivery system was much better coordinated, and its instructional delivery was consistently rated in the effective level. Teachers and administrators in School 1 had better working knowledge of evaluating assessment data and using them to change instruction. The collaboration among staff members was observed in the way they designed and coordinated tiers of interventions. Teachers and administrators at School 1 demonstrated strong buy-in for implementation of RTI.

On the other hand, at School 2, it appeared that RTI there wasn't universal understanding on the point and process of RTI. The buy-in among staff varied, and consensus was still emerging. The scheduling conflicts of Tier 1 and additional support exemplifies a non-coordinated approach to meeting the instructional needs of students. This was further demonstrated in the limited connection of content across tiers of instructional support and the general education teacher's awareness of which programs were being used in these settings. The quality of the instructional delivery was also rated as inconsistent and of lowere quality then in School 1. Teachers' and principal's knowledge of and commitment to using data to guide instruction was still emerging. The principal's leadership was not visible in the evidence-based decision-making process. The collaboration between teachers was not systematic or on-going, as exemplified by how their grade-level team meetings consisted of only the general education teachers. The scores of these two schools on the RTI Assessment Rubric reflected the status of their implementation and clearly indicated a discernable qualitative difference in RTI implementation between these two schools.

Focus group findings. The focus group findings indicated differences in RTI implementation across the grade-levels. For example, in School 1, the team agreed that Grade 2 was at the level of full implementation across six components. In contrast, Grade 4 was judged by the consensus of the team as partially implemented in Tier 1 instruction because the students were not consistently on task 80% of the time; partially implemented in Tier 2 and Tier 3 implementation because of the lack of resources of implementing Tier 3 interventions, and partially implemented in evidence-based decision-making because teachers used data to judge students' progress but did not use that inference to change instructional practice. In School 2, the team confirmed Grade 2's self-evaluation. They also reached consensus on giving all grades full implementation in universal screening and progress monitoring, in spite of the occasional "glitches" in data reporting (Focus Group Study, 2009). They also agreed that the decision-making process was still emerging, some teachers recognized the importance of using data to guide decision-making but not necessarily support collaboration due to the constraint of time. The School 2 team also recognized the unevenness of resource allocation, in responding to the different needs in different grades. While Grade 1 received full support from reading specialists, Grade 3 received support from a veteran special education teacher. But the supports were arranged between teachers, there was no coherent plan. Consequently, schedule conflicts occurred and forced Teacher 4 to re-arrange schedules to ensure at-risk students could receive support from specialists. On the other hand, Grade 3 and 4 did not have established Tier 3 intervention, their implementation was graded by the team as partially implemented. Focus group participants reported that the

scores on the RTI Assessment Rubric reflected the qualitative differences they perceived in RTI implementation between grades.

Results for Question 4: Connection between Scores on Rubric and Student Measures

To answer this question, one must look into the analysis of student progress monitoring data and the school's scores on the RTI Assessment Rubric. Whether students were making good progress was determined by the relative positions between the trend-line of their progress monitoring data (observed performance) and individual aimline (expected performance). In this study, the individual aimline is constructed by connecting the students' oral reading fluency scores to the grade-level target score, which is reading a Grade 2 level passage at the rate of 85 correct words per minute (50th percentile on the District Benchmark Assessment in Reading).

Figures 2, 3, and 4 show the progress monitoring data for the five student case study participants, with red lines indicating aimlines (expected progress) and blue lines indicating trendline (observed progress). The figures provide a way for teachers to examine the relative positions between the trendline and aimline as students progressing toward the year-end goal. Among these five students, Albert of School 1 and Eric of School 2 were both on track in terms of their progress (see Figure 2).
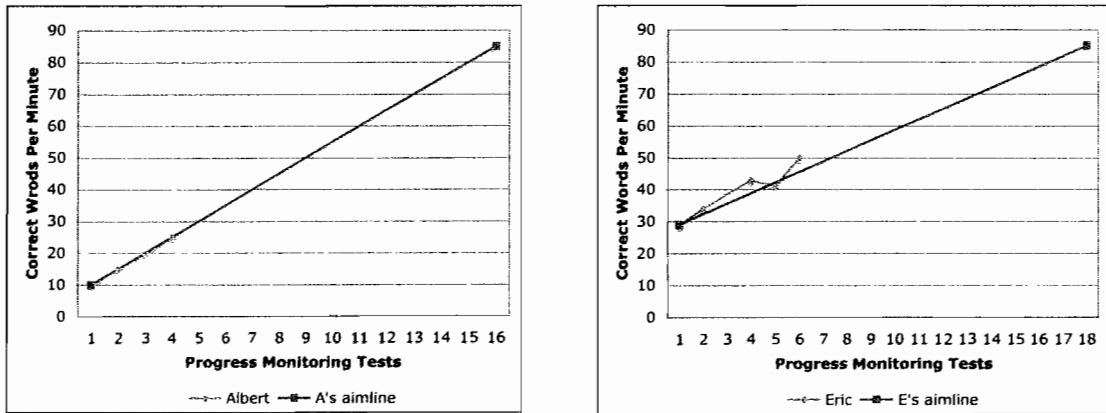
Figure 2

Progress of Two "On-track" Students

Becky of School 1 and Diane of School 2 were both approaching their aimlines (see

Figure 3).
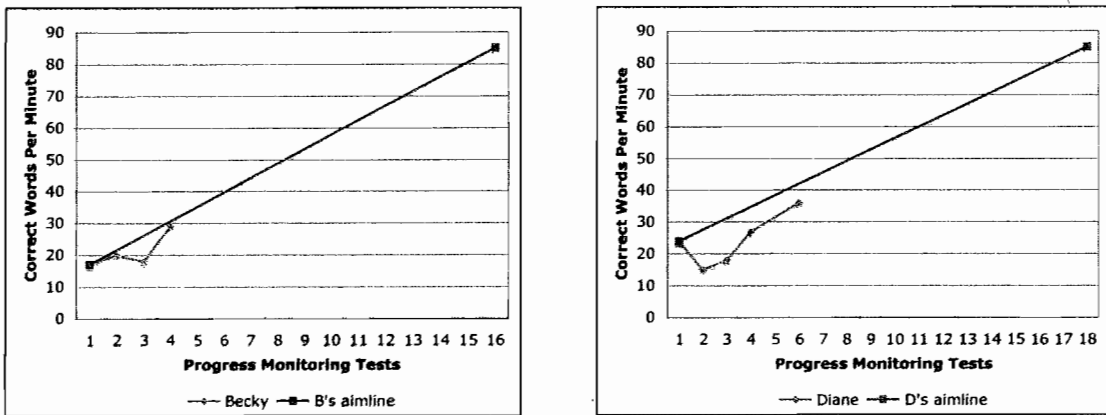
Figure 3

Progress of Two "Approaching" Students

The trajectory of Carl's Progress (School 1) was rather flat in comparison with his

expected progress (see Figure 4). He appeared not to be responding to the provided

instruction and interventions.
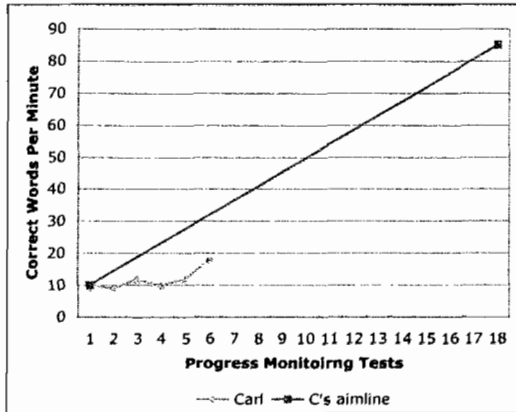
Figure 4

Progress of a "Non-responder"

In School 1, three students had three different outcomes: Eric's progress is on the aimline, Diane's progress is approaching the aimline with an upward trend, and Carl's is below the aimline with little discernable progress. In school 2, two students had two relatively similar outcomes: Albert is on the aimline and Becky has two of the three data points on the aimline and one slightly below the aimline (see Figures 2, 3, and 4). In all, two of the five students were on track, two of them had made some progress and only one student was not making good progress.

Summary

In this chapter, I reported findings for the four research questions about the content-related validity, reliability, sensitivity, and potential use of the RTI Assessment Rubric. These findings collectively represent four different strands of evidence related to the validity and reliability of the rubric. In the next chapter, I integrate these strands of evidence to present a validity argument to justify the proposed use of the rubric, which is to judge the quality of RTI implementation in the elementary school level.

CHAPTER V

DISCUSSION

Evaluation is necessary and instrumental to the implementation of any school-wide effort or change, including RTI (Horner, Todd, Lewis-Palmer, Irvin, Sugai, & Boland, 2004; Kame'enui & Simmons, 2003). My goal in this dissertation was to develop an assessment instrument to assist teachers and administrators to evaluate their implementation of RTI. However, developing a rubric to evaluate RTI is particularly challenging because neither research literature nor federal regulation specifies what constitutes RTI (Burns & Coolong-Chaffin, 2006), and the implementation of RTI is influenced by the individual nuances of the school and the district and community in which it is located (Coyne et al., 2001; Graden et al., 2007).

To overcome these two hurdles, I reviewed relevant RTI literature, identified six essential components of RTI, and developed an initial draft of the rubric based on these components (see Chapter 2). I asked three content experts to review the rubric and used their feedback to revise the rubric (see Chapter 3). I conducted a field study to create school reports, describing how these components worked in practice in two different schools. Finally, I asked two independent raters to evaluate the implementation of RTI in these two elementary schools, using the developed rubric and the school reports (see Chapter 4). Essentially, this dissertation could be viewed as a technical report, designed to

systematically document the process of developing and validating a test instrument, the RTI Assessment Rubric. In this chapter, I will review the major findings of this study, report on the limitations of the instrument development process, and discuss the implications of my findings.

Major Findings of the Study

There are four major findings of this study. Each has significant implications for the validation of the RTI Assessment Rubric. First, the three experts provided content-related validity evidence by agreeing that the six identified components captured the essence of RTI and were sufficient to assess RTI implementation. Although there were some minor differences noted in terms of the operational definitions of the descriptors and how to qualitatively differentiate the level of implementation within each of the components, most of these differences revolved only around the operational definitions related to instruction, both Tier 1 as well as Tier 2 and 3. Second, the high inter-rater agreement between the two raters and between the raters and participating teachers suggests that the description of RTI implementation in these two schools is credible and the rubric is stable. Third, the way in which the scores on the rubric reflected the qualitative differences in RTI implementation observed at the two case study schools provide evidence for the rubric's sensitivity. However, with such a limited sample, findings related to the connection between the scores on the RTI Assessment Rubric and Tier 3 students' learning outcomes were inconclusive.

Limitations of Current Project

As with any study, this project had limitations that impact generalization and interpretations of the findings. These limitations relate to how the rubric was revised, the

process of generating the school reports, and the characteristics of the schools selected for the case study.

Concerns about the Revisions

After the experts had provided feedback on the initial version of the rubric, I made modifications using my professional judgment but did not resubmit the revised version to these same experts for a second review. As a result, I do not know if the revisions completely addressed the concerns of the experts or improved the overall quality of the rubric. Reviewers did not indicate the need for a second review of the rubric; however, additional expert review might have provided a more complete validation process.

Concerns about the Neutrality of the Data

The very complicated process of collecting and creating the school profiles may have been biased because I was the only person to gather the information. Potentially, my relationship and experiences at these school sites may have introduced biases that influenced how the school profile was created. To address this potential concern, I attempted to conduct my case study using the same processes and data sources schools would use to evaluate their own RTI implementation efforts. I interviewed teachers, examined instruction, reviewed school documentation, and observed school-based meetings as I was gathering data to create credible school reports. However, it is important to point out that typically the person rating RTI implementation would be on site to make these observations directly rather than viewing them through the lens of a school report prepared by somebody else. Thus, my own biases might be reflected in the school report and extend from those reports to the independent raters, thereby threatening the validity of the evaluation.

Given the breadth and depth of the RTI Assessment Rubric, one might question whether site-based educators would have the skills, knowledge and professional experience needed to use the rubric to perform a similar evaluation. Specifically, to conduct an evaluation of RTI implementation might require an educator with a great deal of expertise in the areas of reading instruction, assessment, school coordination and documentation, as well as consultation skills in interviewing school personnel. Although the teachers, school psychologists, and principals who participated in the focus groups all appeared quite competent in using the rubric, future studies might examine whether the evaluator's range of experience and relation with the stakeholders influences the reliability and validity of the rubric to evaluate RTI implementation. A follow up study might consider whether the rubric is better completed by a team of specialists (i.e., reading specialists, school psychologist, administrators), rather than one evaluator, to ensure a more accurate and efficient evaluation of the school's RTI implementation.

Concerns about the Representation of the Sample

An additional limitation of the project relates to the nature of the schools used in the data collection process. The schools were from a single school district that had developed its own RTI procedures. These procedures differed in some ways from what the research discusses as best practices in RTI implementation (Fuchs et al., 2003; Grimes & Kurns, 2003; Kamps & Greenwood, 2003). Specifically, the district's model differed in providing Tier 2 intervention to all students. This differs from the typical perception of a three-tiered model that emphasizes providing intervention only to students who are at or below pre-determined standards or to students whose level and slope of progress is significantly less than the average of the class (Fuchs & Fuchs, 2006;

Johnson et al., 2006; Marston, 2005; McMaster et al., 2005; Vaughn et al., 2003).

Because the sample district provided Tier 2 intervention to every student through

differentiated instruction, it is possible that the results obtained in this study might not

generalize to settings using a more typical RTI approach. However, it is worth noting that

the implementation of RTI varies greatly from site to site as it is influenced by the

schools' environment. Thus, the concerns about the degree to which the case study

schools used a unique approach to implement RTI may actually be a strength of this

study. Given the diversity with which RTI is conceptualized in different settings, an RTI

assessment rubric must be able to detect differences in the level of implementation

regardless of the specific details in the model being used.

An additional limitation relates to the specific characteristics of the schools

themselves as well as the students used in the case study. Neither school had a Title One

or ELL program, which does not reflect most high-needs schools in the country. Also, all

of the case study students were Caucasians, native English speakers, and had not been

identified as having behavioral problems. These five students may not be representative

of typical Tier 3 students within the RTI process. Future studies should examine the role

of varying school and student characteristics in connection with the utility of the RTI

Assessment Rubric.

Concerns about the Sample Sizes

Large sample sizes can strengthen the ability to generalize between the observations

derived from the sample and inferences about the target population (Gall, Gall & Borg,

2007). In this study, I used the rubric to evaluate two schools, three classes, and five

students. Although this number of cases is appropriate for a case study (Yin, 2003), it is far

from adequate sampling if my purpose were to generalize the findings related to student

outcomes to the larger population of elementary school students. Future studies should

focus on expanding sample size and diversifying the samples, including schools with ELL

programs, Title One programs, and in urban, suburban, and rural areas. More research is

needed to investigate the influence of students' characteristics on the evaluation of RTI

implementation. It is important to investigate whether this rubric is biased against schools

with a high percentage of English language learners, students with low SES, and students

at-risk of having behavioral problems. In the instrument development process, it is typical

to conduct multiple studies to gather information about the technical adequacy of a

measurement instrument. This study provides initial data that should be supplemented with

findings from additional studies specifically addressing the invariance of the measure

across different student populations (Miles & Huberman, 1994).

Validation of the RTI Assessment Rubric

Validity is the most important consideration in evaluating the quality of the uses

and interpretations of test scores (AERA, APA, NCME, 1999; Haladyna & Downing,

2006; Linn, 2002; Kane, 1992, 2006). Validation of the test instrument for a particular

use is not an afterthought. Rather it is a thoughtful process, integrated in every step of

instrument development. In Chapter 1, I described six steps of instrument development

relevant to the development of the RTI Assessment Rubric (Downing, 2006). The first

three steps (identifying the important components of the targeted domain,

operationalizing these important components and providing descriptors for each

component at different levels of understanding, and creating test items based on the

descriptors listed in the test blueprint) are designed to ensure that the items in a test

instrument assess all of the most important components of the intended construct, or the test instrument preserves "construct integrity" as defined by Messick (1989). The fourth step, gathering content related validity evidence, can be accomplished by relying on dependable judgments of content experts when there are no validated assessments to which to compare (Downing, 2006). The fifth step, conducting a pilot study on the assessment, is intended to gather evidence about the stability of the instrument. The final step of the validation process, creating a technical report and documenting validity evidence, is designed to articulate a validity argument to justify the proposed use of the scores.

In this study, the first three steps (identifying the essential components of RTI through a synthesis of the literature, providing operational definitions and descriptors as observable or measurable criteria for each components, and creating the rubric by which to assess the construct) are designed to ensure the construct integrity of the evaluation instrument. The fourth step (content review) provides content related validity evidence for the rubric. The outcomes of the fifth step (scoring of the rubric by independent raters) yields evidence for the stability or reliability of the rubric. Finally, the raters' and the participating teachers' feedback on the relevance and ease of use of the rubric also provides evidence related to its utility and social validity for use as an instrument to evaluate RTI implementation in an elementary school. In short, this test development procedure lays the groundwork for validation of the RTI Assessment Rubric.

Next, I will discuss the findings related to my four research questions.

Research Question 1: To What Extent Does the RTI Assessment Rubric Appropriately Measure the Construct of RTI Implementation?

This question focuses on the evidence and theory to support the "domain relevance and representativeness of the instrument" (Messick, 1989, p. 17). According to Kane (1992), to substantiate that a proposed test instrument is relevant and representative of the target domain, the content of the instrument must meet two assumptions. First, the test items must be relevant and representative of the target domain. Second, the domain of the instrument must match the domain of the construct. To address the relevance and representativeness of the target domain, I used both a review of the literature and expert feedback.

In conducting the literature review, I explicitly stated the inclusion and exclusion criteria for selecting relevant literature to ensure that all of the articles reviewed focused on the implementation of RTI at the elementary school level in the context of reading, the domain of the RTI Assessment Rubric. The rubric was developed based on the essential RTI components distilled from the results of the literature review. For instance, the purpose, instructional emphasis, explicit and systematic instruction, and evidence for the effectiveness of instructional delivery are all addressed in the operational definitions and descriptors of Tier 1instruction.

When teachers provide all students high quality Tier 1 instruction, they take a proactive approach to prevent students from developing reading skill deficits (Fuchs et al., 2003; Gresham et al., 2005; Grimes & Kurns, 2003; Kovaleski, 2003). Speece, Case, and Molly (2003) referred to students' responsiveness to high quality Tier 1 instruction as the first line of defense in preventing students from developing difficulties in learning to read.

All of the literature reviewed indicated that Tier 1 reading instruction should encompass the "five big ideas" of reading instruction (Foorman & Torgesen, 2001; Grimes & Kurns, 2003; Marston et al., 2003; Mellard et al., 2004; O'Connor, 2000).

Justice (2006), Kamp and Greenwood (2003), and Vaughn, Linan-Thompson, and Hickman (2003) all agreed that Tier 1 instruction should be provided for a sufficient time and in an explicit and systematic manner. All of the literature reviewed yielded a unanimous conclusion on the importance of teaching the five big ideas in Tier 1 instruction, a consensus supported by many reading researchers (Denton et al., 2003; Elmore, 1996; O'Conner et al., 2005a; Simmons & Kame'enui, 2006; Snow et al., 1998; Torgesen et al., 2001).

Many researchers suggest that Tier 1 Instruction alone should meet the instructional needs (i.e., meeting grade-level expectations) of approximately 80% of students without additional instructional support (Foorman & Ciancio, 2005; Foorman & Moats, 2004; Fuchs & Deshler, 2007; Good et al., 2003; Simmons & Kame'enui, 2006; Torgesen et al., 2003; Vaughn & Fuchs, 2003). Based on this consideration, in the rubric the quality of Tier 1 instruction is judged in several ways. It is judged by the alignment between teachers' instructional emphases at each level of instruction / intervention and the grade-level content standards they are expected to address (Foorman & Torgesen, 2001; Grimes & Kurns, 2003; Marston et al., 2003; Mellard et al., 2004; O'Connor, 2000). It is also judged by students' responsiveness to that instruction (Foorman & Ciancio, 2005; Foorman & Moats, 2004; Fuchs & Deshler, 2007; Good et al., 2003; Simmons & Kame'enui, 2006; Torgesen et al., 2003; Vaughn & Fuchs, 2003).

In addition, it is judged by the amount (Bollman et al., 2007; Callender, 2007;

Foorman & Moats, 2004; O'Conner et al., 2005a; O'Conner et al., 2005b; Peterson et al., 2007; Pikulski, 1998; Treptow, 2006; Vaughn et al., 2003) and use (Allington, 1983; Allington, 2002; Hollywood et al., 1995; Treptow, 2006) of instructional time. The presence of effective teaching behaviors (Edmonds & Briggs, 2003; Foorman & Moats, 2004; Foorman & Schatschneider, 2003; Gambrell et al., 1981; Graves et al., 2004; Grek et al., 2003; Hagger et al., 2003; Kim et al., 2003; Klingner et al., 2003; Greenwood, 2003), and teachers' and students' perception of the instruction (Blumenfeld et al., 1991; Carnine, 1997; Kagan, 1992) are also important components.

In the rubric, I emphasized the importance of explicit and systematic instruction by including the use of evidence-based teaching strategies such as Direct Instruction and Peer-Assisted Learning Strategies (Carnine, Silbert, & Kame'enui, 1990; Fuchs, 2003; Rohrbeck, Ginsberg-Block, Fantuzzo, & Miller, 2003; Simmons & Kame'enui, 2006; Vaughn et al., 2003) in the operational definition of Tier 1 instruction on the rubric. The findings from the literature review provide the first piece of evidence for the relevance and representativeness of the domain.

Additional evidence for domain relevance and representativeness relates to the qualifications of the content experts and the procedures and results of the content review. As principal investigators for OSEP-funded Model Demonstration Centers on RTI (see Appendix C), the expert reviewers are qualified to judge whether the rubric is relevant and representative of the domain, RTI implementation. These three experts were unaware of the purpose of this study, and they were not involved in any other aspects of developing the rubric. They conducted their reviews independently. Based on the procedures of the content review, it is plausible to assert that their reviews were unbiased. Downing (2006) states that

it is appropriate to use dependable judgments of content experts in the target domain as the source of content-related evidence when there are no validated test instruments in the same domain.

Prior to my study, there were no validated instruments to evaluate the implementation of RTI. In the absence of validated instruments, I followed Downing's advice and used the feedback of the content experts as a key source of content-related validity evidence. These three experts independently agreed with the provided operational definitions and descriptors for the six identified components, with some minor revisions. Their endorsement provides justification for including these "test items" in the rubric. The experts did not suggest the exclusion of any components or descriptors of the components. They did not suggest that any part of the rubric was irrelevant to an evaluation of the implementation of RTI. In addition, the experts did not indicate any part of the rubric that might bias against certain sub-groups of students. Most importantly, the experts did not identify any additional RTI components, nor did they suggest expanding the scope of the evaluation.

These findings suggest that the rubric encompasses all of the key components of RTI implementation. Messick (1989) defines "content integrity," where the items of a test assess all of the most important components of the targeted domain, as one of the key considerations for the validity of a test instrument. If an important component is excluded from the test, then the validity of the test is undermined because of construct under-representation (Webb, 2006). The fact that none of the content experts suggested any additional RTI components provides evidence to support the claim that this rubric satisfies the requirement of content integrity.

The comprehensiveness of the rubric is further supported by the outcomes of the case study. During the case study, I shadowed five Tier 3 students in three classes for one week each to observe and document how schools provide instruction and assessment to improve these students' reading performance. I delved through various data sources, including reviewing documents, observing classes and school-based meetings, observing tiers of instruction in class, and interviewing teachers and students. During this intense data collection process, I did not find any components that were not already included in the rubric.

Based on evidence gathered from the literature review and content experts, the RTI Assessment Rubric matches the domain of RTI implementation at the elementary school level in the context of reading, and these six identified components are essential and sufficient to evaluate the implementation of RTI.

Research Question 2: What Evidence Substantiates the Reliability of the RTI Assessment Rubric for Evaluating the Implementation of Each Identified RTI Component?

Reliability of the rubric is related to its use, which, different from content-related validity, involves the persons using the rubric and the data being used. Therefore, the evidence of reliability should include confirmation about the credibility of the data being used, the procedure for scoring using the rubric, and that the rubric results in similar evaluations from different raters.

In this dissertation, raters applied the rubric to evaluate the implementation of RTI at two different schools based on detailed school reports. The credibility of the school

reports was substantiated by triangulating multiple data sources gathered during the field study and by confirming the findings with the stakeholders (participating teachers and administrators) during the second interviews (Miles & Huberman, 1994; Yin, 2003). For example, I both used the school schedule to calculate the scheduled instructional time and observed each class for a week to ensure more than 60 minutes of Tier 1 instruction was provided to all students daily. I used the first interview question on Tier 1 instruction (*How do you provide students Tier 1 instruction?*) to guide my classroom observation; I used my observation to verify the statements teachers made in the first interviews, and finally I used the second interview to verify my findings about the provided Tier 1 instruction. I also provided the school reports to Teacher 1 (homeroom teacher of Carl and Diane) and Teacher 4 (home room teacher of Albert and Becky). They both confirmed the findings in the school report. The convergence of evidence suggests that the school report accurately documents the RTI experience that each child received.

Part of evaluating reliability includes ensuring a standardized scoring procedure for an assessment instrument (Downing, 2006). In this dissertation, the scoring procedure was written and provided to raters prior to their scoring. The two raters scored the rubric independently by reading the same school reports and scoring them one component at a time. Their notations on the rubric indicated that they both found justifiable cause to support their ratings. Their annotated rubrics provided evidence to support the claim that their ratings are fair and unbiased.

To evaluate the reliability of the rubric across the two raters, I compared the agreement between their scores and compared these scores to reports from participating teachers during a focus group. The inter-rater agreement between these two raters was

strong with a 92% agreement (as reported in Chapter 4, they agreed on 11 of 12 occasions). In the focus group I asked teachers to first evaluate their own RTI implementation based on their own perception, using the RTI Assessment Rubric, and then to share their ratings and reasons for those ratings with their colleagues. Teacher 1 and Teacher 4 both participated in the focus group study and conducted their own self-evaluation on RTI implementation. Teacher 1 was in 100% agreement with both raters. Teacher 4 was in 92% agreement with one rater and 100% agreement with the other rater. It is noteworthy that the data gathered in the school report about the RTI implementation (written in November, 2008) and the teachers' description of their RTI implementation in the focus group (conducted in April-May, 2009) are very similar. The high inter-rater agreement on the scoring of these two school reports suggests that the rubric provided a reliable estimate of RTI implementation. The agreement between raters and participating teachers further strengthens this claim. Triangulation of data, member checking with participants, and the focus group study, in combination, provide defensible evidence to for the reliability of the rubric.

Research Question 3: What Evidence Substantiates the Sensitivity of the RTI Assessment Rubric for Measuring Qualitative Differences in Schools' RTI Implementation?

As stated before, validity is not a property of a test instrument, it is related to a specific use of the score (Linn, 2002). This rubric is designed to evaluate the implementation of RTI in elementary schools. An important consideration is the sensitivity of the instrument, or the degree to which it is able to differentiate schools at different levels of implementation.

Like other instruments used to evaluate school implementation of academic or behavioral programs, (see, for example, Kame'enui et al.'s Planning and Evaluating Tool, 2003; and Horner et al.'s School-wide Evaluation Tool, 2004), the RTI Assessment Rubric divides the continuum of implementation into three levels and uses multiple data sources including interviews, observation, review of permanent document and work products to render a judgment. Evaluation of the current levels of implementation can guide practitioners as they identify areas for program improvement.

In this study, the RTI implementation in School 1 was judged as fully implemented across all six components by both independent raters and participating teachers. In contrast, both raters agreed that RTI implementation in School 2 was partially implemented in four of six categories and fully implemented in one category (universal screening). They disagreed on whether progress monitoring was fully (Rater 1) or partially implemented (Rater 2). In this case, participating teachers agreed with Rater 2. The consistent difference in ratings across the two schools provides evidence that the rubric is able to detect qualitative differences in implementation.

All sources of data indicate that School 1 is in a more advanced stage of RTI implementation than School 2. School 1 has all the essential components of RTI in place, at least in Grade 2, where this study was based. In contrast, staff at School 2 are still developing consensus about what is meant by evidence-based decision making and solidifying buy-in among staff. Teachers in School 2 need more on-going technical support in interpreting assessment data and using the data to modify instruction.

Research Question 4: Do Differences in RTI Implementation, as Measured by the RTI

Assessment Rubric, Correspond with Different Student Outcomes?

Batsche et al. (2005) and Fuchs et al. (2003) suggest that implementing RTI can improve academic outcomes for all students. Although my study provides some indication that implementation of RTI might be beneficial to students' progress, the data I gathered are insufficient to answer this research question. The developers of the easyCBM assessment system report an expected growth rate of 1.4 words per week on the passage reading fluency measures between September and January for students in second grade performing at the 50th percentile (Alonzo, Tindal, Ulmer, & Glasgow, 2006). On average, students in my case study exceeded this growth rate. The one student (Carl) who failed to make significant progress toward his instructional goal during the eight weeks for which I collected data showed a 9 word improvement, a rate of 1.3 words per week, which is slightly less than the expected rate of growth.

Findings from my study suggest that both partial and full implementation of RTI might lead to improved student learning outcomes, although caution is warranted given the small sample size. The RTI approach at School 1 was considered fully implemented across all six components of the rubric based on all sources of data. Two of the three case study students at School 1 demonstrated above average growth rate with the remaining student demonstrating slightly less than the expected rate of growth. The average rate of growth on the passage reading fluency measures of the three students at School 1 was 2.2 words per week. When the data for the sole non-responding student were removed, the average rate of progress of the two remaining students was 2.7 words per week, almost double the expected growth rate for students in grade 2.

In contrast, School 2 had lower levels of RTI implementation based on the RTI Assessment Rubric yet had an average rate of progress of 2.25 across the two students in my sample. Again, this rate of growth exceeds the expected growth rate for students in this grade level, 1.4 words per minute. Given the short duration of the study, this finding is promising for the RTI approach in general. However, my study does not provide conclusive evidence to suggest that there are differences in student learning outcomes between schools that have partial or full RTI implementation. As in all case studies, the small sample size prevents generalizing beyond the immediate cases. Perhaps a larger sample would reveal more conclusive differences; however, it is equally plausible that the level of RTI implementation is less important to student outcomes than the adoption of the RTI approach in general. Additional research is needed in this area.

Many factors outside the scope of this dissertation might account for the observed learning gains. School effectiveness and teacher effectiveness (Hill & Rowe, 1998) are two potential contributing factors for student achievement. Some students might excel in spite of inadequate instruction. Some students might falter in spite of a well-implemented RTI program and generally effective instruction, in the context of RTI, these students are classified as non-responders (Burns et al., 2005; Fuchs & Deshler, 2007; Gresham et al., 2005; Marston, 2005; Speece et al., 2003; Speece & Hines, 2008). Thus, attempting to use the scores from the RTI Assessment Rubric to predict student learning outcomes would not be a valid use of this instrument because there are other equally plausible explanations for student progress in reading or lack thereof.

Conclusion

The primary purpose of this study was to develop an assessment rubric to evaluate the implementation of RTI. I presented evidence of content validity and reliability to substantiate its intended use. This rubric appears to be appropriate for both researchers and practitioners' use as a tool in the field study of RTI. Given that RTI is an emerging field of study, this rubric fills an important void in the practice of RTI. It can provide relevant, practical, and useful information to guide schools' next steps in RTI implementation.

APPENDIX A

TABLES

Table 1

Summary of the Process of RTI by Selected Articles

Author(s), (Year)	Ways to Describe the Process of RTI	
	Multi-Tiered Intervention	Cycle of Problem-Solving
1. O'Conner (2000)	X	
2. Foorman & Torgesen (2001)	X	
3. Gresham (2002)	X	
4. Fuchs, D., Mock, Morgan, & Young (2003)	X	X
5. Good, Simmons, Kame'enui, & Chard (2003)	X	
6. Grimes & Kurns (2003)	X	X
7. Kovaleski (2003)	X	X
8. Kamps & Greenwood (2003)	X	X
9. Speece, Case & Molly (2003)	X	
10. Tilly (2003)	X	X
11. Vaughn & L.S. Fuchs, (2003)	X	
12. Vaughn, Linan-Thompson & Hickman-Davis (2003)	X	
13. Foorman & Moats (2004)	X	
14. Mellard, Byrd, Johnson, Tollefson & Boesche (2004)	X	
15. Burns & Ysseldyke (2005)	X	X
16. Christro, (2005)	X	

Table 1 (Continued)

Summary of the Process of RTI by Selected Articles

Author(s), (Year)	Multi-Tiered Intervention	Cycle of Problem-Solving
	Ways to Describe the Process of RTI	
17. Foorman & Ciancio (2005)	X	
18. Marston (2005)	X	
19. O'Conner, Fulmer & Harty, (2005)	X	
20. Troia (2005)	X	
21. VanDerHeyden & Jimerson (2005)	X	X
22. Denton, Fletcher, Anthony & Frances (2006)	X	
23. Fuchs & Fuchs (2006)	X	
24. Lau, Sieler, Muysken, Canter, VanKeuren, & Marston (2006)	X	X
25. Justice (2006)	X	
26. Fuchs & Fuchs (2007)	X	

Table 2

Summary of Components of RTI Model by Selected Articles

Selected Articles	Components of RTI					
Author(s), (Year)	US	ET1	ET2	PM	ED	OS
1. O'Conner (2000)	X	X	X	X	X	
2. Foorman & Torgesen (2001)		X	X	X		
3. Gresham (2002)	X	X	X	X	X	
4. Fuchs, D., Mock, Morgan, & Young (2003)		X	X	X	X	X
5. Good, Simmons, Kame'enui, & Chard (2003)	X	X	X	X	X	X
6. Grimes & Kurns (2003)	X	X	X	X	X	X
7. Kovaleski (2003)		X	X	X	X	X
8. Kamps & Greenwood (2003)		X	X	X	X	
9. Speece, Case & Molly (2003)		X	X	X		X
10. Tilly (2003)	X	X	X	X	X	X
11. Vaughn & L.S. Fuchs, (2003)		X	X	X	X	X
12. Vaughn, Linan-Thompson, & Hickman-Davis (2003)		X	X	X	X	
13. Foorman & Moats (2004)		X	X	X	X	X
14. Mellard, Byrd, Johnson, Tollefson & Boesche (2004)		X	X	X	X	X
15. Burns & Ysseldyke (2005)	X	X	X	X	X	X
16. Christro, (2005)		X	X	X	X	X
17. Foorman & Ciancio (2005)	X	X	X	X	X	X

Table 2

Summary of Components of RTI Model by Selected Articles

Selected Articles	Components of RTI					
Author(s), (Year)	US	ET1	ET2/3	PM	ED	OS
18. Marston (2005)		X	X	X	X	
19. O'Conner, Fulmer & Harty, (2005)		X	X	X	X	
20. Troia (2005)		X	X	X	X	X
21. VanDerHeyden & Jimerson (2005)	X	X	X	X	X	X
22. Denton, Fletcher, Anthony & Frances (2006)	X	X	X	X	X	X
23. Fuchs & Fuchs (2006)	X	X	X	X	X	
24. Lau, Sieler, Muysken, Canter, VanKeuren, & Marston (2006)	X	X	X	X	X	X
25. Justice (2006)		X	X	X	X	
26. Fuchs & Fuchs (2007)	X	X	X	X	X	
Number of Articles	12	26	26	26	24	16

US = Universal Screening; ET1= Effective Tier 1 Instruction; ET2/3 = Effective Tier 2 or Tier 3 Intervention; PM = Progress Monitoring; ED = Evidence-based Decision-Making; OS = Organizational Support.

Table 3

Comparison of RTI Components identified in selected articles, NASDSE (2005) and RTI Manual (2006)

Identified Component	NASDSE (2005)	RTI Manual
Universal Screening	√	√√
Effective Tier 1 Instruction	√	√√
Effective Tier 2 and Tier 3 Instruction	√	√√
Progress Monitoring	√	√√
Evidence-Based Decision-Making	√	√√
Organizational Support	√	√√

√: Mentioned, but no specific features.
√√: Mentioned and Emphasized with specific features.

Table 4

School characteristics

	School A	School B
School Structure	K-5, 3 classes per grade	K-5, 2 classes per grade for Grades K-3 and 4 Grades 4/5 combined classes
Enrollment	500+	<300
Special programs (e.g. Title I, ELL programs)	Non -Title 1 school No ELL programs Regional Learning Center on campus	Non-Title 1 school No ELL programs Regional Learning Center on campus
Other instructional supports	Reading Specialist, Student achievement coordinators	Reading specialist
Prior experience in progress monitoring before adopting RTI Model	Yes, they used DIBELS for screening and progress monitoring.	No prior experience in progress monitoring

Data source: *MDP-RTI study: School Profiles*

Table 5

Teacher Characteristics

	School 1				School 2	
Teacher	Teacher 1	Teacher 2	Teacher 3	Teacher J	Teacher 4	Teacher 5
Primary assignment	General education	General education	Special education	Instructional Assistant	General education	Special education
Highest level of education	BA	MA	MA	BA	MA	MA
Teaching certification	Yes, general education	Yes, general education	Yes, general education	Certified as Instructional Assistant	Yes general education	Yes, special education
Years of experience in teaching	6	10	1	8	26	29
Years of teaching students with disabilities	3	10	1	8	26	29
Instructional role(s) in RTI	Tier 1 Tier 2	Tier 1 Tier 2	Tier 2	Tier 2 Tier 3	Tier 1 Tier 2	Tier 2 Tier 3

Data Source: *MDP_RTI Study: Teacher Survey*

Table 6

Student Characteristics

School	School 1			School 2	
Student[1]	Carl	Diane	Eric	Albert	Becky
Grade	2	2	2	2	2
Gen. Ed. teacher	Teacher 1	Teacher 1	Teacher 2	Teacher 4	Teacher 4
Race	White	Unknown	White	White	White
Primary language	English	English	English	English	English
Have IEP?	No	No	No	No	No
Have been retained in grade	No	No	No	No	No
Screening score in cwpm (Percentile)	8 $(1\text{-}10^{th})$	24 $(11\text{-}20^{th})$	29 $(11\text{-}20^{th})$	10 $(1\text{-}10^{th})$	17 $(11\text{-}20^{th})$
Recommended to receive Tier 3 interventions	Yes	Yes	Yes	Yes	Yes

Data Source: *MDP-RTI Student: Characteristic Surveys*; *district assessment data; teacher interviews.*

Grade 2 general education teachers nominated three lowest students in their classes that fit the selection criteria: (a) scoring at or below 20th percentile in the district's fall reading assessment, and (b) being recommended to receive Tier 3 interventions. When three students' parents from a class all gave consent, two with the lowest scores were chosen. In Teacher 2's class, only one student's parents gave consent.

Table 7

Examples of Evidence Categorized by Types

Types of Evidence	Examples
Archival Records	• Students' Student Support Team (SST) Plan or IEPs (if applicable) • Students' scores on the progress monitoring measures • Students' district reading assessment scores
Documents	• Agenda or minutes of professional development meetings. • Agenda or minutes of district-based/school-based RTI meetings. • Agenda or minutes of school-based grade-level meetings. • Students' progress reports retrieved from the easyCBM website that pertain to students' progress monitoring scores, progress measures used, as well as interventions that implemented.
Direct Observation	• Classroom observations: Identify the programs being used in class. • Classroom observations: Use E-COVE to document the number of minutes teachers spend teaching the five big ideas as specified in NRP 2000. • Classroom observation: Evaluate the effectiveness of instruction using Classroom Observation Checklist (Gersten, Baker, Haager, & Grave, 2005). • Classroom observations: Describe the mood of the class, evidence of effectiveness of instruction in the field notes. • School-based meeting: Describe how teachers interpret data and use data to inform decision-making in the field notes.
Interviews	• Two interviews for each participating educator, one prior to the classroom observation, the other after the classroom observation.
Physical Artifacts	• School's Master Schedules. • Students' easyCBM scored protocols. • Students' other work products if appropriate.

Table 8

Data Collection Schedule

Weeks	Description of Tasks
1	• Conduct site visits. Introduce myself to the participating educators. • Obtain the district RTI plan, school master schedules and school-based RTI implementation plan (if available). • Obtain schedules for upcoming SST, Grade-level and RTI meetings. • Identify qualified students; send home the student assent/parent consent letters. • Schedule Interviews for participating educators.
2	• Start initial interviews. Transcribe the interviews and proceed with coding. • Obtain informed consent/assent from parents and students. • Schedule the observations for the SST, Grade-level and RTI meetings.
3	• Conduct direct observation for the first pair of dyads. • Attend scheduled RTI, SST or Grade-level meetings. • Obtain the scores of participating students on the district reading assessment and progress monitoring measures from teachers. • Reconstruct the chronology describing the instructional support students receive over the course of the school year.
4	• Conduct direct observation for the second pair of dyads. • Attend scheduled RTI, SST or Grade-level meetings.
5	• Conduct direct observation for the third pair of dyads. • Attend scheduled RTI, SST or Grade-level meetings.
6	• Update the participating students' progress reports. • Conduct the final interviews. • Transcribe the final interviews and proceed with preliminary coding. • Attend scheduled RTI, SST or Grade-level meetings.
7-8	• Wrap up unfinished tasks, including rescheduled interviews. • Proceed with data analysis.

Table 9

Organization of the Initial Interview Questions and Six Critical Components

	US	T1	T2/3	PM	ED	OS
How do you determine which students receive additional intervention?	X				X	
Tell me about the curriculum you use. How do you know it is a research-based curriculum?		X				X
Tell me about your routine in teaching reading at a typical day.		X	X			
What are your instructional emphases in reading?		X	X			
What kind of instructional supports do the students with most intensive needs receive?			X			
How do you provide instructional support for students who have severe reading skill deficits?			X			
Describe the three-tiered intervention in your school.		X	X			
How do you monitor students' learning outcomes? Tell me about the tool(s) you use.				X		
How do you know if the intervention was effective?					X	
What professional training do you receive for this curricular program?						X
What do you know about RTI? How did you learn about RTI?						X
Describe the instructional support you receive from your principal and other teachers in the school.						X

US = universal screening; T1 = Tier 1 Instruction; T2/3 = Tier 2 & Tier 3 interventions; PM = progress monitoring; ED = evidence-based decision-making; OS = organizational support

Table 10

Organization of the Second Interview Questions and Six Critical Components.

	US	T1	T2/3	PM	ED	OS
How do you determine Student X needs additional intervention?	X				X	
What kind of skill deficits does student have? How does the chosen curriculum address Student X's educational need?		X				X
What were your instructional emphases for Student X?		X	X			
What instructional support does Student X receive?		X	X			
How do you provide Student X additional instructional support?			X			
How does Student X move through the RTI system?		X	X			
How do you determine if Student X makes progress?				X	X	
How does the RTI team determine what support Student X needs?					X	
How do professional training and collaboration with your colleagues help you implement reading programs under RTI?						X
How do professional training and collaboration influence the quality of instruction you provide for Student X?						X
How would you grade the support you receive from your principal and your colleagues?						X

US = universal screening; T1 = Tier 1 instruction ; T2/3 = Tier 2 & Tier 3 interventions; PM = progress monitoring; ED = evidence-based decision-making; OS = organizational support.

Table 11

Data Sources Used to Document Each RTI Component

Data Sources	RTI Critical Components					
	T1	US	T2/T3	PM	ED	OS
Archival Records: school schedules, classroom schedules	√	√	√	√		
Archival Records: student test dates, student screening and progress monitoring data, student intervention profiles		√		√	√	
Document: technical reports of the chosen assessments		√			√	
Document: student screening and progress monitoring data (print out retrieved from test-developer's website)		√		√		
Document: meetings minutes and agenda, grade-level meeting protocols, updates of instructional grouping			√		√	√
Physical Artifacts: Student test booklets		√		√	√	
Classroom Observation	√		√			
Observation of school-based meetings and professional training					√	√
Interviews	√	√	√	√	√	√

T1: Tier 1 instruction; US: Universal Screening; T2/T3: Tier 2 and Tier 3 interventions; PM: Progress Monitoring; ED: Evidence-based Decision-making; OS: Organizational Support

Table 12

Participants of the Four Focus Group Interviews

	School 1	School 2	School 3	School 4
Principal	X	X	X	
Grade 1 teacher	X			X
Grade 2 teacher	X	X		
Grade 3 teacher	X	X	X	
Grade 4 teacher			X	X
Grade 5 teacher	X			
School Psychologist	X	X		X
Special Education teacher		X		X
Reading Specialist/ Facilitating Teacher		X	X	X

School 3 and 4 share the same principal, administrative support, and building.

Table 13

Results of Content Review

		T1	US	T2/3	PM	ED	OS
Dr. Ed Shapiro	Operational Definition	1	1	1	1	0	1
	Full Implementation	1	1	0	1	1	0
	Partial Implementation	1	1	0	1	1	0
	Non-Implementation	1	1	0	1	1	1
Dr. Doug Marston	Operational Definition	0	1	1	1	1	1
	Full Implementation	1	1	1	1	1	1
	Partial Implementation	1	0	1	1	1	1
	Non-Implementation	1	1	1	1	1	1
Dr. Teri Wallace	Operational Definition	0	1	1	1	1	1
	Full Implementation	1	1	1	1	1	1
	Partial Implementation	1	1	1	1	1	1
	Non-Implementation	1	1	1	1	1	1

1: Agree with provided definitions and descriptors; 0: Disagree

Table 14

Evaluation of Two Schools on RTI Implementation

	Critical Components of RTI	Levels of Implementation	
		Evaluator 1	Evaluator 2
School A	Tier 1 Intervention	2	2
	Universal Screening	2	2
	Tier 2 and Tier 3 Interventions	2	2
	Progress Monitoring	2	2
	Evidence Based Decision Making	2	2
	Organizational Support	2	2
School B	Tier 1 Intervention	1	1
	Universal Screening	2	2
	Tier 2 and Tier 3 Interventions	1	1
	Progress Monitoring	2	1
	Evidence Based Decision Making	1	1
	Organizational Support	1	1

2 = Fully Implemented; 1 = Partially Implemented; 0 = Not Implemented

Table 15

Inter-rater Agreement between Two Evaluators

		Evaluator 1	
		Fully Implemented	Partially Implemented
Evaluator 2	Fully Implemented	7	0
	Partially Implemented	1	4

Table 16

Targeted Students' Fall Screening Data: School 1

	Chosen Screening Measures						
	DIBELS[1] CWPM (%ile)		Composite Risk Status	EasyCBM[2] CWPM (range of %ile ranking)			Composite Percentile Ranking
Students	NWF	ORF		WRF	PSF	MCRC	
Carl	30 (8)	6 (4)	Intensive – Needs substantial intervention	8 (1-10)	10 (1-10)	*	9
Diane	32 (11)	16 (13)	Intensive – Needs substantial intervention	20 (11-20)	24 (11-20)	7	25
Eric	41 (21)	28 (26)	Strategic – Additional Intervention	28	29 (11-20)	6	26

1. DIBELS fall benchmark was administered in September 4-10.

2. EasyCBM fall benchmark was administered in September 22-26.

3. Goal: 50 letter sounds per minute.

4. Goal: 44 correct words per minute (cwpm).

5. WRF (Word Reading Fluency). Grade 2 fall: 10%ile =12 cwpm, 20%ile = 20 cwpm, 50%ile = 40 cwpm, 75%ile = 62 cwpm, 90%ile = 80 cwpm.

6. PSF (Passage Reading Fluency). Grade 2 fall: 10%ile =18 cwpm, 20%ile = 30 cwpm, 50%ile = 57 cwpm, 75%ile = 84 cwpm, 90%ile = 127 cwpm.

7. MCRC (Multiple-choice Reading Comprehension. Grade 2 fall: 10%ile =3, 20%ile = 5, 50%ile = 8, 75%ile = 9, 90%ile = 11. Total possible score = 12.

8. Data Sources: teacher interviews, school schedule, School 1 DIBEL Grade 2 fall benchmark report, School 1 EasyCBM Grade 2 fall benchmark report, EasyCBM progress monitoring scores interpretation guidelines.

Table 17:

Targeted Students' Fall Screening Data: School 2

	Chosen Screening Measure			
	easyCBM In Correct Words Per Minute (CWPM) (range of %ile ranking)[1]			Composite Percentile Ranking
Students	WRF[2]	PSF[3]	MCRC[4]	
Albert	14 (11-20)	10 (1-10)	1	6
Becky	8 (1-10)	17 (1-10)	4	11

1. Red = 1-10 percentiles. Yellow = 11- 20 percentiles.

2. WRF (Word Reading Fluency). Grade 2 fall: 10%ile =12 cwpm, 20%ile = 20 cwpm, 50%ile = 40 cwpm, 75%ile = 62 cwpm, 90%ile = 80 cwpm.

3. PSF (Passage Reading Fluency). Grade 2 fall: 10%ile =18 cwpm, 20%ile = 30 cwpm, 50%ile = 57 cwpm, 75%ile = 84 cwpm, 90%ile = 127 cwpm.

4. MCRC (Multiple-choice Reading Comprehension. Grade 2 fall: 10%ile =3, 20%ile = 5, 50%ile = 8, 75%ile = 9, 90%ile = 11. Total possible score = 12.

5. Data Sources: teacher interviews, school schedule, School 2EasyCBM Grade 2 fall benchmark report, easyCBM progress monitoring scores interpretation guidelines.

Table 18

Tier 1 Instruction: Instructional Minutes Spent on Activities during Case Study Week

		School 1		School 2
		Carl & Diane	Eric	Albert & Becky
Teacher-led activities focusing on	Phonemic awareness	0	9	0
	Phonics	31	50	0
	Fluency	23	42	0
	Vocabulary	47	35	0
	Comprehension	63	33	0
Student-led activities	Read aloud[1]	49	31	12
	Silent reading[2]	8	63	17
	Independent seatwork[3]	36	26	0
Total time on the academic activities		257	289	29
Total time spent on other activities such as transitions, discipline or house-keeping		150	119	109
Total time spent on Tier 1 Instruction		407	408	138
% of time spent on the academic activities		63	71	21
% of time students appeared on task		80+	80+	50-80
Average Scores of *Classroom Observation Checklist*[4] (Range of the Scores)		42 (40-44)	41 (39-43)	N/A

1. Students read list of words or connected text with minimal teacher scaffolding.
2: Students read the assigned passages or books of their choice quietly. Teacher sometimes asked vocabulary and comprehension questions afterwards.
3: The worksheets were directly related to phonics, vocabulary, or comprehension.
4: In effective classes, teachers' scored at or above 39 when summed across all 24 items.
Data sources: Classroom observation.

Table 19

Tier 2 and Tier 3 Interventions: Instructional Minutes Spent on Activities during Case Study Week

		School 1				School 2	
		Carl & Diane		Eric		Albert & Becky	
		Tier 2	Tier 3	Tier 2	Tier 3	Tier 2	Tier 3
Teacher-led activities focusing on	Phonemic awareness	0	0	29	0	0	3
	Phonics	1	12	9	11	12	124
	Fluency	54	11	12	12	11	123
	Vocabulary	18	9	34	5	9	10
	Comprehension	5	3	9	12	3	13
Student-led activities	Read Aloud	7	0	15	0	0	27
	Silent Reading	11	0	6	0	0	37
	Independent Seatwork	8	0	12	0	0	0
Total Time on the academic activities		109[1]	35	125	40	35	227
Time on the non-academic activities		25	2	29	5	10	40
Total Time for Reading Instruction		134[1]	38	154	45	45	267
% of time spent on the academic activities		81	92	81	89	78	85
% of time students appeared on task		90+	90+	90+	90+	64	90+
Average Scores of *Classroom Observation Checklist (Range)*		40 (38-42)	42 (41-43)	44 (42-46)	42 (41-44)	37 (32-41)	43 (42-44)

1: Four-day week. Data sources: Classroom observation.

Table 20

Documentation of Evidence-based Decision-making

		School 1		
		Carl	Diane	Eric
Data used: EasyCBM progress monitoring data	3 data points since last data meeting	√	√	√
	Description of Progress	Flatline; below aimline.	Upward trend; below aimline	Upward trend; on track
Teacher input	Attendance	Good	Absence and tardiness	Good
	Classroom Performance	Good, with attending issues	Good	Good, with attending issues
	Additional Assessments	10/29: Test for eligibility	No	No
	Others (Specified)	No	Troubled home life	No
Decision-making process	Follow decision rules	√	√	√
	Multi-disciplinary decision-making[1]	√	√	√
Decisions made	Change of current interventions	NO	NO	Yes. Move from Group 2 to Group 3 in Tier 2 interventions
	Refer to receive special education	Pending the results of the evaluation.	NO	NO

		School 2	
		Albert	Becky
Data used: EasyCBM progress monitoring data	3 data points since last data meeting	√	√
	Summary of data based on the graph	Upward trend; below aimline.	No trend detected; below aimline
Teacher input	Attendance	Some absence	Good
	Classroom Performance	Good, close monitoring is needed	Good, close monitoring is needed
	Additional Assessments	No	No
	Others (Specified)	this is his 7th school	Mom resisted the idea of testing
Decision-making process	Follow decision rules	No explicit decisions rules	No explicit decision rules
	Multi-disciplinary decision-making[1]	Decisions made by two general education teachers	Decisions made by two general education teachers
Decisions made	Change of current interventions	No	No
	Refer to receive special education	No	No

Data Source: Progress monitoring data and observation of school-based meetings.

APPENDIX B

THE RTI ASSESSMENT RUBRIC

RTI ASSESSMENT RUBRIC

Dear Teachers and Principals:

Thank you for choosing RTI Assessment Rubric to evaluate the implementation of RTI at your school. The purpose of this rubric is to assist you evaluating the implementation of RTI at your school or your grade-level. It is comprised of six mini-tests, one for each identified essential component of RTI. It would take approximately 15-20 minutes to conduct a self-evaluation on your own implementation. I recommend general education teachers evaluating the RTI implementation at your own grade-level, and principals, school psychologists, reading specialists and special education teachers evaluate it from a school-wide perspective. After the self-evaluation, you should share your ratings and the reasons to support those ratings with your colleagues. Through this focused and guided discussion, you will have a better understanding about the current status of RTI implementation at your school. You can then develop an action plan based on your conclusion.

In the next page, I include a simple scoring sheet and a template for the school profile. I hope you find this rubric a useful tool to assist you implementing RTI. If you have any questions regarding this rubric or the implementation of RTI, please do not hesitate to contact me (kliu@uoregon.edu). I am looking forward to your feedback. Please send back a copy of your annotated rubric in .pdf or word documents format, via email, briefly describing the reasons/facts to support your rating. Your feedback can expand our collective knowledge on RTI implementation. Thank you in advance for your willingness to help. It is not uncommon that you will have some grades that are more advanced in the RTI implementation than the others. I hope this will help you start your own research and develop a RTI program that can be benefit your students most.

Sincerely,

Kimy Liu

University of Oregon

RTI Assessment Rubric Scoring Sheet

	Universal Screening (US)	Tier 1 Instruction (T1)	Tier 2 and Tier 3 Interventions (T2/T3)	Progress Monitoring (PM)	Evidence-based decision-making (ED)	Organizational Support (OS)
Score						

2 = fully implemented; 1 = partially implemented; 0 = not implemented at all

Action Plan

RTI Components	Area(s) need to be improved	Priority
US		
T1		
T2/T3		
PM		
ED		
OS		

RTI Assessment Rubric

Operational Definition
1. Tier 1 Instruction (in reading) is defined as the core reading instruction provided to all students in general education classes for a sufficient amount of time to meet grade-level instructional goals. In Tier 1 instruction, teacher teaches all "five big ideas" as specified in NRP 2000 or relevant ones to specific grade-level, using evidence-based teaching strategies.
Data sources for documenting the quality of Tier 1 Instruction: □ Classroom observation   □ School Schedule   □ Teacher and student interviews

Level of Implementation	Descriptors/ Features
Fully Implemented	• Teachers explicitly teach all "five big ideas" specified in NRP 2000/or relevant ones to specific grade-level (i.e., teaching phonological awareness might not be expected in Tier 1 after first grade) AND • Teacher teaches all students Tier 1 reading instruction daily for at minimum 60-90 minutes with most of the time spent engaged in grade-level appropriate reading and writing activities. Teacher spends more than 50% of instruction time on explicit instruction of reading and writing. Many effective teaching behaviors identified in a chosen classroom observation checklist are present at the effective level during classroom observation. AND • Student appears on task 80% of time. (On task is defined as sitting in their seats properly, eyes looking either at teachers or at the instructional materials attentively, focusing on assigned reading or writing activities.) • Teacher and student believe instruction is appropriate.
Partially Implemented	Check those that apply (there may be more than one): _____ Teacher explicitly teaches some of the relevant "big ideas" OR _____ Teacher provides instruction less than the allocated time by teaching less each day or for 3-4 days each week. OR _____ Teacher spends 25-50% of instruction time on explicit instruction of reading and writing. Some effective teaching behaviors identified in the chosen classroom observation checklist are present during classroom observation at or above the level of partially effective. OR _____ Students appears on task 50-80% of the time. OR _____ Teacher and student are uncertain whether instruction is appropriate.
Not Implemented	• Teacher does not teach any of the "five big ideas" explicitly OR • Teacher teaches Tier 1 reading instruction for only a small amount of time each day or for 1-2 days each week OR • Teacher spends less than 25% of instruction time on explicit instruction of reading and writing. Few of the targeted effective teaching behaviors are present during classroom observation at or above the level of partially effective. OR • Student appears on task less than 50% of time. OR • Teacher and student believe instruction is inappropriate.

Operational Definition	
2. Universal Screening (in reading) is defined as using a measure (with appropriate evidence of validity and reliability) to assess all students' reading skills and identify students who score below pre-determined benchmarks. Data from universal screening are presented in a manner that reflects an appropriate unit of analysis so that (a) teachers can identify students who need additional instructional support, or (b) schools can configure grade-level interventions.	
Data sources for documenting the quality of universal screening: ☐ Document: Technical report or manual of the chosen screening measure ☐ Archival Record: the test dates, teacher roster, students scores on the screening measures ☐ Document: the test dates, students' scores, cut scores or benchmark for the chosen screening measures	

Level of Implementation	Descriptors/ Features
Fully Implemented	• The chosen screening measure has explicitly documented validity and reliability evidence AND • The screening measure is administered to all students three times a year to identify students who need additional instructional support AND • Data are gathered in a timely manner and displayed in a way that teachers can systematically identify students needing additional instructional support within each grade.
Partially Implemented	Check those that apply (there may be more than one): _____ The measure has face validity and some reliability evidence OR _____ The measure is administered to all students once or twice a year OR only to some students OR _____ Data are not gathered in a timely manner or not displayed in a way that teacher can easily identify students needing additional support.
Not Implemented	• The screening measure has no established validity or reliability evidence OR • The measures are not administered to any students OR • Data are gathered but not used for identifying students needing instructional support.

Operational Definition
3. Tier 2 and Tier 3 Intervention is defined as explicit and differentiated instruction for students who do not make progress with only Tier 1 instruction. It is supplementary to the Tier 1 instruction and is provided for a sufficient amount of time to improve targeted student's learning outcomes in the specific skill deficit areas (e.g., some of the "Five Big Ideas").
Data sources for documenting the quality of Tier 2 and Tier 3 Interventions: □ Observation or record of School-based meetings (SST, RTI or Grade-level Meetings)   □ Classroom observation for Tier 2 and Tier 3 interventions   □ Physical Artifact: School schedule and nature or programs/materials   □ Teacher and student interviews

Level of Implementation	Descriptors/ Features
Fully Implemented	• Tier 2 and Tier 3 interventions are research-based and explicitly targeted to specific "big ideas" matching the identified instructional needs of students needing additional support. AND • The interventions are implemented with high fidelity and for sufficient amount of time as specified by the program recommendations. AND • Teacher spends more than 50% of instruction time on explicit instruction of identified instructional big ideas, using evidence-based strategies. Most targeted effective teaching behaviors emphasizing on differentiating instruction for low-performing students as identified in the chosen checklist are present at the effective level during classroom observation. AND • Student appears on task 80% of time. (On task is defined as sitting in their seats properly, eyes looking either at teachers or at the instructional materials attentively, focusing on assigned reading or writing activities.) • Teacher and student believe instruction is appropriate.
Partially Implemented	Check those that apply (there may be more than one): ____ The content of Tier 2 and Tier 3 interventions is not research-based or it is used to supplant Tier 1 instruction. ____ The interventions don't match the identified instructional needs of students receiving additional support. ____ The interventions are implemented with some fidelity, and/or not for insufficient amount of time. ____ Teacher spends more than 50% of instruction time on explicit instruction of "the 5 big ideas". Some targeted effective teaching behaviors are present during the classroom observation at or above the level of partially effective. ____ Student appears on task 50-80% of time. ____ Teacher and student are uncertain whether instruction is appropriate.
Not Implemented	• The instruction is not explicit or does not have specific instructional emphases. OR • Teacher teaches Tier 2 and Tier 3 interventions for only a small amount of time each day or for only a couple of days each week. OR • Teacher spends less than 50% of instruction time on explicit instruction of the targeted "big ideas". OR • Student appears on task less than 50% of time. OR • Teacher and student believe instruction is inappropriate.

Operational Definition
4. Progress Monitoring is defined as scientifically based practice that is used to assess students' academic performance and evaluate the effectiveness of instruction. Progress monitoring can be implemented with individuals or an entire class. Data from progress measures are presented in a manner that allows change over time to be visible.

Data sources for documenting the quality of Progress Monitoring:
□ Document: Progress monitoring data
□ Review of the technical adequacy of the chosen progress monitoring instrument

Level of Implementation	Descriptors/ Features
Fully Implemented	• Teacher selects progress monitoring measures that have documented validity and reliability evidence for monitoring progress on a specific skill that matches students' instructional needs AND • Teacher uses progress monitoring measures to track student's progress in specific skills. They are administered on weekly or biweekly for at least three times to establish a pattern. AND • Progress monitoring data are collected on schedule and are used to inform instruction. The data are displayed in appropriate graphs so that the change over time can be visible.
Partially Implemented	Check those that apply (there may be more than one): _____ The measures have face validity or some reliability evidence for monitoring progress of reading. _____ Chosen progress monitoring measures is not designed to assess specific skill matching students' instructional needs. _____ Teacher administers the test, but not systematically or less frequently than monthly or only once or twice before changing the interventions. _____ Progress monitoring data are used to inform instruction. The data, however, are not collected on schedule or graphically displayed.
Not Implemented	• The measures have no established validity and reliability evidence for monitoring progress OR • Teacher does not administer progress monitoring measures OR • Progress monitoring data are not collected or used to inform instruction.

Operational Definition	
5. Evidence-based decision-making is defined as using data of the progress monitoring measures and program specific assessments to judge the effectiveness of interventions to determine if intervention modifications are warranted.	
Data sources for documenting the quality of evidence-based decision-making: □ Teacher interviews □ Progress Monitoring Data and program specific assessments □ Observations or records of school-based meetings	
Level of Implementation	Descriptors/ Features
Fully Implemented	• Teacher systematically and regularly examines classroom data (e.g., data of program specific assessments) and data of progress measures. Teacher analyze the graphs of progress monitoring data, focusing on students' levels of performance and rates of progress (slope) AND • Teacher analyzes specific instructional components AND • Teacher determines whether to fade, continue, modify, or intensify the student's current instructional intervention based on the empirical data gathered.
Partially Implemented	Check those that apply (may be more than one): _____ Teacher informally and episodically reviews classroom data and graphs of progress monitoring data. OR _____ Teacher examines the instructional program in general. OR _____ Teacher changes instructional programs based mainly on random classroom observations.
Not Implemented	• Teacher does not review any classroom information or progress measures OR • Teacher makes no reference to current instructional program OR • Teacher provides unsystematic instruction.

Operational Definition
6. Organizational Support is defined as leadership, resources, administrative support, and professional training and collaboration to facilitate implementation of school-wide programs.

Data sources for documenting the quality of organizational support:
- ☐ Teacher and administrator interviews
- ☐ Observations or records of school-based meetings
- ☐ Observations or records of professional training meetings
- ☐ Physical artifacts: school schedules

Level of Implementation	Descriptors/ Features
Fully Implemented	• There is clear and direct evidence of leadership directives, resource allocation, and administrative support AND • Teachers receive multiple sessions of professional training and on-going support to implement research-based instruction, assessment and evidence-based decision-making. The professional training is focused and interactive. Teachers collaborate with colleagues to identify students' needs and implement tailored interventions to address the needs AND • Teachers collaborate to configure the school-wide intervention programs.
Partially Implemented	Check those that apply (there may be more than one): _____ There is indirect evidence of leadership directives, resource allocation, and administrative support. _____ Teachers receive only initial professional training and have no on-going support to implement research-based instruction, assessment and evidence-based decision-making. Professional collaboration occurs in isolated incidences. The extent of collaboration is limited due to issues of logistics and role responsibilities _____Teacher collaborate to configure the interventions only at the grade-level or classroom-level. There is no coherent school-wide configuration of the intervention programs.
Not Implemented	• No evidence is available to show leadership directives, resource allocation, and administrative support AND • Teachers receive no professional training and rarely collaborate AND • Decision-making is made by individual teachers.

APPENDIX C

QUALIFICATION OF RTI EXPERTS

Three content experts on RTI reviewed the proposed *RTI Assessment Rubric*.

Their qualification and experiences were described as followed:

Expert 1: Edward S. Shapiro, Ph.D., currently was Professor of School

Psychology and Director, Center for Promoting Research to Practice in the College of

Education at Lehigh University, Bethlehem, Pennsylvania. He was the author or co-

author of 10 books including his most recently published third edition of Academic Skills

Problems: Direct Assessment and Intervention. Dr. Shapiro was best known for his work

in curriculum-based assessment and non-standardized methods of assessing academic

skills problems. Among his many projects, Dr. Shapiro co-directs a federal

model/demonstration project focused on the development of a multi-tiered, Response-to-

Intervention model in two districts in Pennsylvania. Over the past five years, Dr. Shapiro

has been working as a consultant with the Pennsylvania Department of Education to

facilitate an effort to establish progress monitoring systems for students in special

education and was currently collaborating with the Pennsylvania Department of

Education in developing and facilitating the implementation of the Response-to-

Intervention methodology for the state.

Expert 2: Douglas Marston, Ph.D., currently was a Special Education Administrator in Loring Elementary of Minneapolis Public Schools. He was the co-principal investigator of an OSEP sponsored demonstration grant using progress monitoring to deliver an RTI model. Loring School was one of the elementary schools using this model. Loring Elementary was a K–5 school in the Minneapolis public school district. Loring has an enrollment of 323 students of which 80% live in poverty. Students of color make up 75% of the school population, 21% were English language learners (ELL), and 8% receive special education services. Dr. Marston has published peer-reviewed articles and lectured nationally on RTI.

Expert 3: Teri Wallace, Ph.D. currently was a professor in Institute on Community Integration at University of Minnesota. Her research focuses on special education, paraprofessional and teacher training, general education and special education collaboration, whole school reform, and data-based decision-making. She was currently a Co-principal Investigator of an OSEP Field Initiated Project called "Creating a Progress Monitoring System: Preparing Secondary Students with Disabilities for Success on Standards Tests." This study was focused on identifying measures that teachers can use to determine a student's progress toward success on the Basic Standards Tests in reading or writing. The goal of her study was to encourage teachers use assessment data in their instructional decision-making.

APPENDIX D

CLASSROOM OBSERVATION CHECKLIST (COC)

Date:	School/Grade/Tier:	0	1	2
Cluster 1: Explicit Teaching				
1-1.	Model skills and strategies during lessons			
1-2	Makes relationship among concepts overt			
1-3.	Emphasizes distinctive features of new concepts			
1-4.	Provides prompts and cues in how to use strategies, skills, and concepts			
1-5	Length of literacy activities was appropriate			
1-6	Adjust own use of English during lesson to make concepts comprehensible			
Cluster 2: Instruction Geared Toward Low Performers				
2-1	Achieve high level of response accuracy in context of lesson objectives			
2-2	Ensure quality of independence practice			
2-3	Encourages in on-going monitoring of student understanding and performance during lesson			
2-4	Elicit responses from all students, including students having difficulty with task at hand			
2-5	Modifies instruction for students as needed during the lesson			
2-6	Provides extra instruction, practice or review for students having difficulty with task at hand			
2-7	Checks students comprehension of text by asking questions			
Cluster 3: Interactive Teaching				
3-1	Secure and maintain student attention during lesson.			
3-2	Extent to which students were on task during literacy activities.			
3-3	Select and incorporate students' responses, ideas and examples and experience into lesson			
3-4	Give students wait time to respond to questions			
Cluster 4:Vocabulary Development				
4-1	Teach difficult vocabulary prior and during lesson			
4-2	Structure opportunities to speak English			
4-3	Provides systematic instruction to vocabulary development			
4-4	Encourage students in meaningful interaction about text			
Cluster 5: Phonemic Awareness and Decoding				
5-1	Provide systematic instruction in phonemic awareness			
5-2	Provide systematic instruction in letter-sound correspondence			
5-3	Provide systematic instruction in decoding			

Scores for Cluster 1	Scores for Cluster 2	Scores for Cluster 3	Scores for Cluster 4	Scores for Cluster 5	Total Score

Likert Rating were 0 = not effective; 1= partial effective, 2= very effective. Items adapted from the *English Language Learner Classroom Observation Instrument*. Gersten, Baker, Haager and Graves, (2005). Note: Subscales were clustered were empirically derived by factor analysis with replication.

APPENDIX E

SCHOOL REPORTS

School 1 Report

Response to Interventions (RTI) is comprised of six critical components: (a) universal screening, (b) Tier 1 instruction, (c) Tier 2 and Tier 3 interventions, (d) progress monitoring, (e) evidence-based decision-making and (f) organizational support. A reviewer can read a school report and judge the school's implementation of RTI by its implementation of these six components using the *RTI Assessment Rubric*.

To help reviewers frame school data in a proper context, I first describe the method of data collection, and then report the findings.

Method

I sampled and documented a school's implementation of RTI at one grade-level. This report documents not only the design and execution of the plan, but also the living experience of individual at-risk students receiving reading instruction under the RTI model. The data collection process was guided by Yin's (2003) case study methodology. The guiding principle of Yin's methodology was to gather data from multiple sources so the credibility of findings can be enhanced by the convergence of evidence.

First, I interviewed teachers and administrators before and after the classroom observation. The purpose of these interviews was to understand the school's plan of RTI and to confirm or clarify the findings in direct observation and document review. Second, I reviewed school schedules, class rosters, student assessment data, and teacher-drafted documents such as grade-level meeting protocol, intervention map and assessment grid. The purpose of document review is to depict the school's plan and implementation on instruction and assessment at the grade-level. Third, I shadowed selected at-risk students for a week to observe the Tier 1, Tier 2 and Tier 3 interventions they received within that time frame. I reported the cumulative minutes the students spent on different instructional activities, the percentage of time students appeared on task, and the teacher's scores on the classroom observation instrument. I also supplemented qualitative description of how teachers used instruction to address students' skill deficits. The purpose of classroom observation is to describe the actual implementation of the RTI model, from the perspective of an at-risk student. Next, I observed school meetings to document how teachers used assessment data to make instructional decisions. For additional information about the method of data collection, please refer to the methods section of my dissertation.

Results

This school report documents School 1's implementation of *Response to Intervention (RTI)* using its implementation at Grade 2 as an example. I first describe the characteristics of schools, participating teachers and students (see Tables 1-3), the district's RTI model (Table 4) and the school's design of tiered interventions (Table 5). Next, it provides documentation of the implementation of the six identified RTI critical components: (a) *universal screening*, (b) *Tier 1 instruction*, (c) *Tier 2 and Tier 3 interventions*, (d) *progress monitoring*, (e) *evidence-based decision-making* and (f) *organizational support*. For each identified component, I provide a brief narrative and a table or tables to summarize the findings (see Tables 6-13).

Table 1

School Characteristics

	School 1	School B
School Structure	K-5, straight grades, 3 classes per grade	K-5, 2 classes per grade for Grades K-3 and 4 Grades 4/5 combined classes
Enrollment	500+	<300
Special programs (e.g. Title I, ELL programs)	Non -Title 1 school No ELL programs Regional Learning Center on campus	Non-Title 1 school No ELL programs Regional Learning Center on campus
Other instructional supports	Reading Specialist, Student achievement coordinators	Reading specialist
Prior experience in progress monitoring before adopting RTI Model	Yes, they used DIBELS for screening and progress monitoring.	No prior experience in progress monitoring

Data source: MDP-School Profiles.

Table 2

Teacher Characteristics

School	School 1				School 2	
Teacher	1	2	3	J	4	5
Primary assignment	General education	General education	Special education	Instructional Assistant	General education	Special education
Highest level of education	BA	MA	MA	BA	MA	MA
Teaching certification	Yes, general education	Yes, general education	Yes, general education	certified as Instructional Assistant	Yes general education	Yes, special education
Years of teaching experience (with students with disabilities)	6 (3)	10 (10)	1 (11)	8 (8)	26 (26)	29 (29)
Instructional role(s) in RTI	Tier 1 Tier 2	Tier 1 Tier 2	Tier 2	Tier 2 Tier 3	Tier 1 Tier 2	Tier 2 Tier 3

Data Source: *MDP_RTI Study: Teacher Survey.*

Table 3

Student Characteristics

School	School 1			School 2	
Student[1]	Carl	Diane	Eric	Albert	Becky
Grade	2	2	2	2	2
Gen. ed. teacher	Teacher 1	Teacher 1	Teacher 2	Teacher 4	Teacher 4
Race	White	Unknown	White	White	White
Primary language	English	English	English	English	English
Have IEP?	No	No	No	No	No
Have been retained in grade	No	No	No	No	No
Screening score in cwpm (Percentile)	8 $(1\text{-}10^{th})$	24 $(11\text{-}20^{th})$	29 $(11\text{-}20^{th})$	10 $(1\text{-}10^{th})$	17 $(11\text{-}20^{th})$
Recommended to receive Tier 3 interventions	Yes	Yes	Yes	Yes	Yes

Data Source: *MDP_Student Characteristic Surveys*; district assessment data; teacher interviews.

1: Grade 2 general education teachers nominated three lowest students in their classes that fit the selection criteria: (a) scoring at or below 20th percentile in the district's fall reading assessment, and (b) being recommended to receive Tier 3 interventions. When three students' parents from a class all gave consent, two with the lowest scores were chosen. In Teacher 2's class, only one student's parents gave consent.

Table 4

Description of District Model

Instruction					
Tiers of instruction	Focus	Targeted students	Delivered by whom	How long?	How often?
Tier 1	Core reading instruction, focusing on the "big 5" as defined by NRP (2000).	All students	General education teacher	40-90 minutes	daily
Tier 2	More differentiated and skill oriented than Tier 1 instruction.	All students	General education teachers, special education teachers, reading specialists, IAs, and other qualified personnel	30 minutes	daily
Tier 3	Small group, targeted, explicit and direct instruction that matches students' instructional needs	Students below 20th percentile in screening, and their progress below the aimline (3+ data points)		60 minutes	weekly

Assessment				
Types of assessments	Purpose of assessment	What test?	Target population	How often?
Screening	Identify at-risk students	Easy	All students	2-3 times a year
Progress Monitoring	Determine if students respond to interventions	easyCBM	students below 20th percentile	Biweekly

Data Source: District's RTI model: Instructional Intervention Progress Monitoring Model (IIPM model).

Table 5: *Summary of Instructional Plan: School 1, Grade 2*

Level of Instructional Support	Who? (What Skill-level of Students)	With What? (Which Materials & Activities)	What More? (Supplemental or Additional Support Materials and Activities)	How Are We Doing? (Determining Instructional Effectiveness with Progress Monitoring)
Tier 1 / Primary:	Which Students: **All Students**	Name of Program / Materials: **McGraw Hill Treasures** When: **8:35-9:40** Activities: **All activities in the core** Group Size: **Large and some small group (3-8)**	Name of Program / Materials: **Not Applicable**	Who to Collect: **IA in charge of progress monitoring** How Often: **Beginning, Middle, End of Year** Criteria: **See Easy test result guidelines** Determining Fidelity of Implementation (Who, With What, How often?): **To be determined**
Tier 2 / Secondary:	Which Students: **All students**	Name of Program / Materials: **McGraw Hill Treasures and Triumphs, Phonics in Reading** When: **8:05-8:35** Activities: **All activities in the core** Group Size: **5-8**	Name of Program / Materials: **Not applicable** Who to Deliver: **General Ed teacher** When: **_X_ w/in typical instructional time** Specify Time (minutes, days of week): **8:05-8:35 M-F** Group Size: **3-8**	Who to Collect: **IA in charge of progress monitoring** How Often: **some students biweekly progress monitoring without Tier 3 intervention.** Criteria: **teacher nomination** Determining Fidelity of Implementation (Who, With What, How often?): **to be determined**
Tier 3 / Tertiary:	Which Students: **Students scored at or below 20th percentile of the district norm** As measured by: **easyCBM screening measures**	Name of Program / Materials: **Sidewalks** When: **8:30-9:15** Activities: **Vocabulary and oral/comprehension activities** Group Size: **Small Group (4 or less)**	Name of Program / Materials: **Sidewalks** Who to Deliver: **Instructional assistant** When: **_ _ w/in typical instructional time** **_X_ in addition to typical time** **Twice a week for 25 minutes each.** Specify Time (minutes, days of week): **Mon., Thurs. or Tues. and Fri., 1:05-1:30 or 1:35-2:00** Group Size: **4 or less**	Who to Collect: **IA in charge of progress monitoring** How Often: **Biweekly** Criteria: **in or above the aimline** Determining Fidelity of Implementation (Who, With What, How often?): **Not yet decided**

Universal Screening

    Interviews with teachers. School 1 scheduled to screen all students in reading three times a year (fall, winter and spring), using the benchmark assessments in Dynamic Indicators of Basic Early Literary Skills (DIBELS) and district reading assessment (easyCBM Measures). Teachers used the fall screening data to create the homogenous groups for Tier 2 and Tier 3 interventions. The principal and Grade 2 teachers believed DIBELS was appropriate for screening and progress monitoring because they had used it for years; they believed easyCBM Measures was appropriate because it was approved by the district.

    Document review. Technical adequacy of both measures can be found in the test developer's technical report (Alonzo & Tindal, 2007; Good & Kaminski, 2002). DIBELS have been reviewed by peer-reviewed journals and Buros' Mental Measurement (6th ed.); the newly developed easyCBM Measures have not received any independent reviews.

    The fall benchmark assessment of DIBELS and easyCBM were administered to all students in September 2008. In Grade 2, DIBELS tested students on nonsense word reading fluency (NWF) and oral reading fluency (ORF); easyCBM tested students on word reading fluency (WRF), passage reading fluency (PSF) and multiple-choice reading comprehension (MCRC). The DIBELS and easyCBM screening data reported students' correct words per minute (CWPM) and their percentile ranking in parenthesis. Grade 2 data were presented in a class roster by the order of students' ORF scores.

    Table 6 provided an example of documentation of implementation of screening, using the participating students' fall screening data. The table reported (a) screening measures used, (b) test dates, (c) students' scores, (d) the percentile ranking and cut scores that helped teachers interpreting test results and (e) the data sources.

Table 6

Targeted Students' Fall Screening Data

	Chosen Screening Measures						
	DIBELS[1] CWPM (%ile)		Composite Risk Status	easyCBM[2] CWPM (range of %ile ranking)			Composite Percentile Ranking
Students	NWF	ORF		WRF	PSF	MCRC	
Carl	30 (8)	6 (4)	Intensive – Needs substantial intervention	8 (1-10)	10 (1-10)	*	9
Diane	32 (11)	16 (13)	Intensive – Needs substantial intervention	20 (11-20)	24 (11-20)	7	25
Eric	41 (21)	28 (26)	Strategic – Additional Intervention	28	29 (11-20)	6	26

9. DIBELS fall benchmark was administered in September 4-10.
10. EasyCBM fall benchmark was administered in September 22-26.
11. Goal: 50 letter sounds per minute.
12. Goal: 44 correct words per minute (cwpm).
13. WRF (Word Reading Fluency). Grade 2 fall: 10%ile =12 cwpm, 20%ile = 20 cwpm, 50%ile = 40 cwpm, 75%ile = 62 cwpm, 90%ile = 80 cwpm.
14. PSF (Passage Reading Fluency). Grade 2 fall: 10%ile =18 cwpm, 20%ile = 30 cwpm, 50%ile = 57 cwpm, 75%ile = 84 cwpm, 90%ile = 127 cwpm.
15. MCRC (Multiple-choice Reading Comprehension. Grade 2 fall: 10%ile =3, 20%ile = 5, 50%ile = 8, 75%ile = 9, 90%ile = 11. Total possible score = 12.
16. Data Sources: teacher interviews, school schedule, School 1 DIBEL Grade 2 fall benchmark report, School 1 EasyCBM Grade 2 fall benchmark report, EasyCBM progress monitoring scores interpretation guidelines.

Tier 1 Intervention

Tier 1 instruction is the core curriculum instruction for all students. The findings were derived from interviews with teachers, review documents and weeklong of classroom observation in which I shadow targeted at-risk students.

Interview with teachers. School 1 scheduled Grade 2 Tier 1 instruction at 8:35-9:40 daily. The core reading curriculum was McGraw-Hill's *Treasures*, one of the two research-based core curriculum approved by the school district. Teachers. G, H and principal stated that the emphasis for Tier 1 instruction was to "teach all of 'the big five ideas': phonemic awareness, phonics, fluency, vocabulary and comprehension" and to "help students meet and exceed grade-level benchmarks (in reading)". The teachers reported that they "implemented it [the curriculum] with fidelity", which was aligned with the principal's request. The teachers believed their *Tier 1 instruction* was meaningful to students because most students' had improved in reading.

Document review. School master schedule designated 8:35-9:40 for the Grade 2 reading time; 10-10:30 as open block, 10:30-11:30 for writing time. In class schedule, the open block was changed into snack and story time (teacher read-aloud time).

Classroom observation. During the week of classroom observation, Teachers 1 and 2 delivered Tier 1 instruction *daily* at the designated reading time; sometimes the instruction was extended into the story and writing time. Teachers 1 and 2 both provided more than 400 minutes of Tier 1 instruction in a week, which exceeded the district's mandate of 60 minutes daily Tier 1 instruction. Teacher 1 spent 63% of Tier 1 instruction time teaching the "five big ideas"; Teacher 2, 70%. In both classes, teachers demonstrated most targeted teaching behaviors at effective levels as identified in *Classroom Observation Instruments* (Table 7). Classroom observation Instrument required observers included field notes to address how teachers use instruction to address students' skill deficits (Gersten, Baker, Hagger, Graves, 2005). Following is the field note:

Decoding and fluency were teacher's primary concern for Carl, Diane and Eric. During Tier 1 instruction, Carl and Eric were both given preferential seating. With teachers' frequent verbal redirects and positive praises, they maintained on task 80+% of the time (On task is defined as "sitting in their seats properly, eyes looking either at teachers or at the instructional materials attentively, focusing on assigned reading or writing activities"). When being called to answer questions about vocabulary and comprehension of the text, they answered correctly. They both could read the grade-level passages from the core reading curriculum when working with teachers one on one or in small group. However, Carl and Eric had troubles reading a previously practiced passage with the entire class, because they could not keep up with the pace of their peers. They often lost their places and their voices faded. Diane was not observed for her independent read-aloud and chorus reading, because of her three-day absence.

Interviews with students. When asked how their homeroom teacher teaching them to read, Diane said, "She [Teacher 1] made me read lots and lots of words and books". Eric said, "She [Teacher 2] asked me to pay attention to the letters, taking time to sound out words and paying attention to the 'the bossy e' (the silent e)". Carl said, "I figure out how to read by myself, because I am good at reading". They all liked reading at school.

Table 7

Tier 1 Instruction: Instructional Activities by Cumulative Minutes in a Week

Instructional activities	Instructional emphasis	Cumulative Instructional Minutes	
		Teacher 1	Teacher 2
Teacher-lead activities	Phonemic awareness	0	8.5
	Phonics	30.8	50.4
	Fluency	22.8	42.2
	Vocabulary	47.3	34.6
	Comprehension	63.4	33.4
Student-lead activities	Read aloud[1]	49.3	31.1
	Silent reading[2]	7.9	62.9
	Independent seatwork[3]	36.4	25.7
Total time on the academic activities		257.9	288.8
Total time spent on other activities such as transitions, discipline or house-keeping		149.9	119.4
Total time spent on Tier 1 Instruction		407.8	408.2
% of time spent on the academic activities		63.24	70.75
% of time students appear on task		80+%	80+%
Average Scores of *Classroom Observation Instrument*[5] (Range of the Scores)		42 (40-44)	41 (39-43)

1: Students read list of words or connected text independently or with minimal teacher scaffolding.
2: Students read the assigned passages or books of their choice quietly. After silent reading , teachers asked them comprehension questions.
3: Students work on assigned worksheet directly related to the phonics spelling patterns, vocabulary, and comprehension strategies, skills.
4: *Classroom Observation Instrument* consisted of 24 items. Likert Rating were 0 = not effective, 1=partially effective and 2= effective. In high-performing classes, teachers on average scored 40 out of the 48 possible scores.
Data sources: Classroom observation.

Tier 2 and Tier 3 interventions

School 1 provided all students daily 30 minutes of Tier 2 interventions (at 8:05-8:35) and some students Tier 3 interventions twice a week for 25 minutes (at 1:05-1:30 or 1:35-2:00). Table 6 provides an overview of Tier 2 and Tier 3 interventions.

Teacher interviews. The school district required all students received 30 minutes of Tier 2 intervention daily. "Students who scored at or below 20th percentile (of district norm in the screening) would receive Tier 3 intervention and progress monitoring every other week." (Teachers 1, 2, 3 and Principal). Teachers provided students who scored between 20-35th percentiles of district norm Tier 3 instruction if slots were available (Teacher 1 and principal). Teachers covered all of the "big 5" in reading during the Tier 2 instruction, with emphasis on phonics and fluency. Tier 3 interventions focused only on phonics and fluency.

Review and observation of school-based meetings. At School 1, general education teachers, special education teacher, facilitating teachers, school psychologist and administrator reviewed screening and progress monitoring data at grade-level data meetings. The purpose of the meetings was to use data to make informed instructional decisions. The meeting of Oct. 12 yielded an updated Tier 2 class roster and a new Tier 3 class roster (Tables 8 and 9). On Nov. 15, Grade 2 team reviewed progress-monitoring data and updated the Tier 2 class roster. The team decided that Carl and Diane should remain in Group 1 (the lowest group) and Eric was promoted from Group 2 to Group 3.

Document review. The Tier 2 and Tier 3 class rosters specified the instructor, curriculum, instructional emphasis, designated time slot and students of each instruction group in Tier 2 and Tier 3 interventions (Tables 8 and 9). According to the roster, Carl and Diane received Tier 2 intervention from the special education teacher, Teacher 3. Eric received Tier 2 intervention first from Teacher 2; later from Teacher 1. They all received Tier 3 intervention from the instructional assistant, Teacher J.

Classroom observations. During the week of classroom observation, Teachers 2 and 3 provided 30 minutes of Tier 2 instruction daily. They spent approximately 81% of Tier 2 instruction tire teaching the "big 5". Teacher J. provided 38 minutes of weekly Tier 3 interventions to Carl and Diane and 45 minutes to Eric while their peers were with specialists. She spent approximately 90% of Tier 3 instruction time teaching the "big 5". In all, teachers demonstrated effective teaching behaviors identified in *Classroom Observation Checklist*; and students appeared on task 90+% of the time (Table 10).

Following is the summary of field note: Decoding and fluency are teacher's primary concern. During Tier 2 and 3 interventions, Carl, Diane and Eric had multiple opportunities practicing decoding and reading connected text. They read with accuracy, but at a slower pace than the one in Tier 1 chorus reading. Because of her absence, Diane received one day of Tier 2 intervention and no Tier 3 intervention in the week of 11/10.

Interviews with students. Carl, Diane and Eric did not notice differences in the ways teachers teaching reading in different groups. When asked if they liked to go to different reading groups, Carl and Diane said, "it's ok"; and Eric said, "it's fun because I can see what other classes were doing." Eric and Diane liked reading and did not think reading was hard. Eric noted, "comprehension was easy, but reading words was hard."

Table 8:

School 1, Grade 2, Class Roster for Tier 2 intervention

Time	Group	Teacher	Curriculum[1] (focus)	Class roster[2] (recommended group size)
8:05-8:35, daily	1	Special education teacher (Teacher 3)	Triumphs (phonics)	Carl, Diane, Greg, Ida, Patty, Rachael (4-6 students)
	2	Class A teacher (Teacher 2)	Triumphs (phonics)	Adam, Jay, Max, Opal, Eric, (4-6 students)
	3	Class B teacher (Teacher 1)	Pre-teaching core curriculum (phonics)	Betsy, Daisy, Jessie, Nancy, Kelly, Steve (6-8 students)
	4	Facilitating teacher	*Phonics for Reading I* (phonics and fluency)	Albert, Bo, Candy, Danny, Ellen, Fred, Ginny, Holly, Iris, Kathy, Jack, Larry, Martha, Ned (12-20 students)
	5	Instructional assistant	*Phonics for Reading II* (phonics and fluency)	Oliver, Paul, Quincy, Robert, Sammie, Tracie, Ursula, Victoria, Walter, Xavier, Yolanda, Zach, Amanda, Beatrice, Colby, Dillon, Emma, Fredrick (12-20 students)
	6	Other Grade 2 General Education Teacher (Teacher 6)	*Treasures* (comprehension skill and strategy)	Gina, Hunter, Jacob, Toni, Maria, Laurie, Emily, Kevin, Nathan, Patricia, Alex, Ryan, Andrew, Cameron, Andy, Harriet (12-20 students)
	7	Instructional Assistant (Teacher J)	Literature discussion using chapter books (comprehension strategies)	Natalie, Owen, Madison, Bianca, Joy, Hope, Charity, Jessica, Nicole, Zoe, Shannon, Connor, Wesley, Irene, Carson, William, Yvonne, Bernie, George (12-20 students)

[1]: *Treasures, Triumphs* are district-approved core and intervention programs because they are research-based. *Phonics for Reading* is also research-based. (Teacher interviews)

[2]: All students' names were changed to protect their privacy.

Table 9 :

School 1, Grade 2, Class Roster for Tier 3 intervention

Grade 2, Tier 3 Instructor: Teacher J			
Time	Monday	Tuesday	Thursday	Friday
1:05 -1:35	Group 3	Group 1	Group 4	Group 2
	Adam	Carl	Jessie	Max
	Betsy	Diane	Kelly	Nancy
	Jay	Greg	Eric	Opal
	Daisy	Ida		
Curriculum	Sidewalks_Level A_Changes	Sidewalks_Level A_Animals	Sidewalks_Level A_Treasures	Sidewalks_Level A_Community

Time	Monday	Tuesday	Thursday	Friday
1:35 -2:05	Group 4	Group 3	Group 2	Group 1
	Jessie	Adam	Max	Carl
	Kelly	Betsy	Nancy	Diane
	Eric	Jay	Opal	Greg
		Daisy		Ida
Curriculum	Sidewalks_Level A_Treasures	Sidewalks_Level A_Changes	Sidewalks_Level A_Community	Sidewalks_Level A_Animals

Table 10

Time spent on "the Big 5" during Tier 2 and Tier 3 interventions

		Minutes Spent on Selected Activities			
Teacher vs. student lead activities	Instructional emphasis	Carl and Diane's group		Eric's Group	
		Tier 2 (n=5)	Tier 3 (n=3)	Tier 2 (n=6)	Tier 3 (n=3)
Teacher-lead activities	Phonemic awareness	0	0	29.3	0
	Phonics	1.4	11.5	8.6	11
	Fluency	54.2	11.4	12.1	11.7
	Vocabulary	17.9	9.1	33.5	5
	Comprehension	4.7	3.2	8.7	12.3
Student-lead activities	Read Aloud	7.1	0	14.8	0
	Silent Reading	10.7	0	5.8	0
	Independent Seatwork	7.7	0	12.2	0
Total Time on the academic activities		108.5	35.6	125	40
Total Time on the non-academic activities		24.5	2.4	28.5	5.2
Total Time for Reading Instruction		133[1]	38	153.5	45.2
% of time spent on the academic activities		81.58	91.99	81.38	88.50
% of time students appear on task		90+	90+	90+	90+
Average Scores of *Classroom Observation Instrument*		40	42	44	42

1: It is a four-day week. 11/12 was a no-school day.

2 *Classroom Observation Checkli*st consisted of 24 items. Likert Rating were 0 = not effective, 1=partially effective and 2= effective. In high-performing classes, teachers on average scored 40 out of the 48 possible scores.
Data sources: Classroom observation.

Progress Monitoring

School 1 used district reading assessment (easyCBM) for progress monitoring; it progress monitored students who scored at or below 35th percentile of district norm in fall screening in reading every other week. For Grade 2, the progress monitoring measures was one-minute passage reading fluency. The test was administered individually.

Interviews with teachers. Teachers and principal stated that an instructional assistant (IA) was assigned to administer the tests to targeted students every other week. The IA also were responsible for entering test scores into a password-protected district website and plotted the data on the graph, at the front page of individual students' test booklets. General education teachers were responsible to keep the booklets of students in their classes.

Document review. Table 11 and Figure 1 showed the progress monitoring data for the targeted students. It indicated test dates, students scores and scores that helped teachers interpreted the test data. Technical adequacy of EasyCBM word reading fluency and passage reading fluency can be found in the test developer's technical report (Alonzo & Tindal, 2007). It provided strong evidences for internal reliability and some evidences for validity. Easy CBM has not been reviewed independently.

Table 11:

Student Progress Monitoring Data: in Correct Words per Minute

	Screening	10/16	10/20	10/28	11/7	11/20	Annual Goal
Carl	10	9	12	10	12	18	110
Diane	24	15	18	27	NA	36	110
Eric	29	34	NA	43	41	50	110

Passage Reading Fluency Second Grade Fall Data Benchmarks: 10th percentile = 18 cwpm, 20th percentile = 30 cwpm, 50th percentile = 57 cwpm. A = Absent.

Carl's Progress Monitoring Data

Diane's Progress Monitoring Data

Eric's Progress Monitoring Data
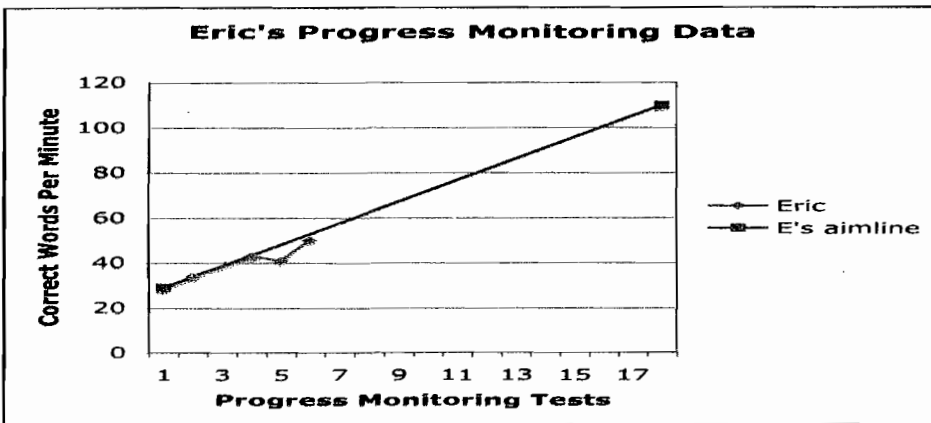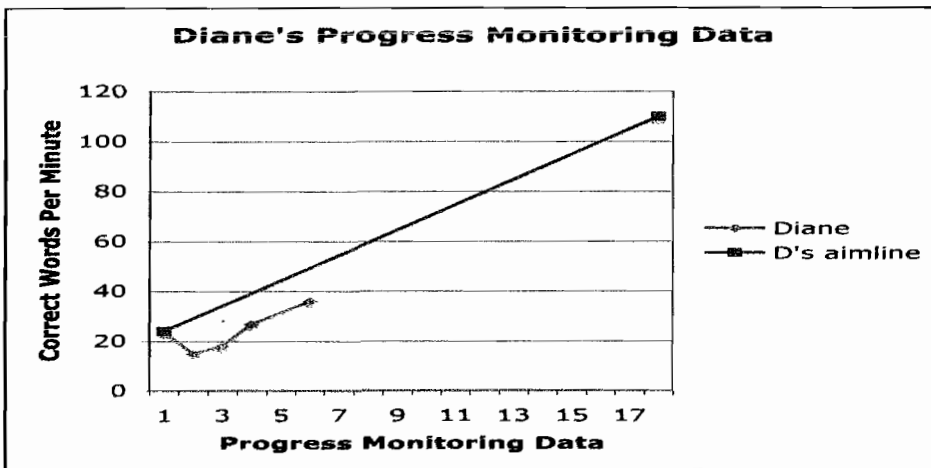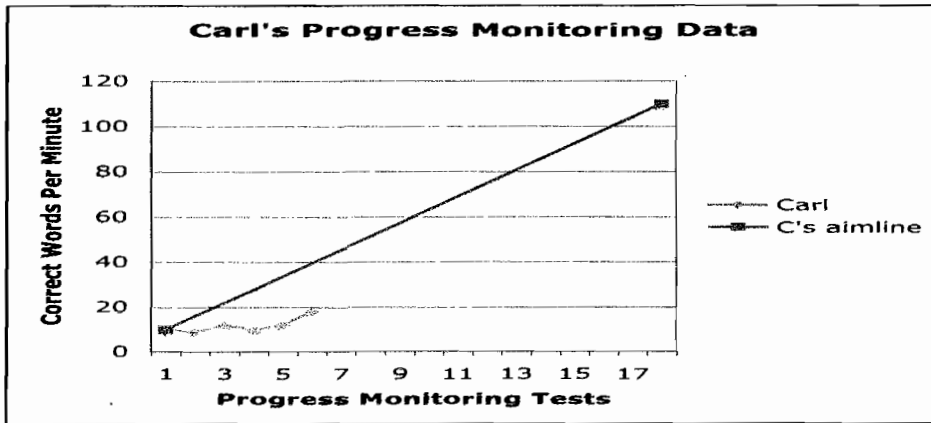
Evidence-based Decision-making

 Teacher interviews. Teachers stated that they used EasyCBM progress monitoring data as the primary data source for decision-making. Each grade met every 4-6 weeks at grade-level data meeting to review students' progress monitoring data and corresponding graph. The principal, school psychologist, general education teacher, special education teacher and facilitating teacher all attended the data meeting. The team would compare student's progress against the expected performance (i.e. the aimline); and determine if student is making good progress. Based on that the determination, the team would decide whether or not to modify student's intervention program or refer students for a full evaluation for eligibility for special education.

 Observation of school-based meetings. I observed the school-based meetings on Oct. 3 and Nov 15. On Oct 3, the staffs were grouped according to grade-levels and collaborated to create a reading assessment grid that identified different types of screening, diagnostic and progress monitoring measures students would receive at grade-level. On Nov. 15, the principal, school psychologist, general education teachers, special education teacher and facilitating teacher attended Grade 2 data meeting and reviewed the progress monitoring data of all Grade 2 student. The team followed the grade-level meeting protocol in conducting the data review process. Teachers looked at individual student's progress monitoring data. Each student data were presented in the booklet, with the front page showed the data points and aimline that indicated expected progress. The team reviewed the EasyCBM progress monitoring data. The teachers who provided Tier 1, Tier 2 and Tier 3 interventions provided additional information on classroom performance, attendances and other relevant information. The school psychologist reported upcoming or pending evaluations. Based on all information available, the team first decided if the student made progress. Next, the team decided whether to modify student's placement in Tier 2 or Tier 3 interventions groups. Finally, teachers also decided whether to refer students to be tested for special education. Table 12 documented the data and information used in decision-making, as well as its process and conclusion. Base on the conclusions, the team updated the students' individual intervention profiles and the configuration of instructional groups for Tier 2 interventions.

Table 12

Documentation of Evidence-based Decision-making

		Students		
		Carl	Diane	Eric
Data used: EasyCBM progress monitoring data	3 data points since last data meeting	✓	✓	✓
	Summary of data based on the graph	Flatline; below aimline.	Upward trend; below aimline	Upward trend; on track
Teacher input	Attendance	Good	Absence and tardiness	Good
	Classroom Performance	Good, with attending issues	Good	Good, with attending issues
	Additional Assessments	10/29: Test for eligibility	No	No
	Others (Specified)	No	Troubled home life	No
Decision-making process	Follow decision rules	✓	✓	✓
	Multi-disciplinary decision-making[1]	✓	✓	✓
Decisions made	Change of current interventions	NO	NO	Yes. Move from Group 2 to Group 3 in Tier 2 interventions
	Refer to receive special education	Not yet, pending the results of the evaluation.	NO	NO

Data Source: Progress monitoring data and observation of school-based meetings.

Organizational Support

Organizational support can manifest in instructional leadership, administrative support, availability of professional training and on-going support, and professional collaboration. Table 13 provided examples of organizational support in School 1.

Table 13

Documentation of Organizational Support

Types of Support	Examples
Instructional leadership	1. Principal and leadership team allocated resources for purchasing research-based intervention programs. 2. Principal allocate resources for implementing small group instruction and progress monitoring. 3. Principal facilitated and participated in grade-level data meetings. 4. Principal facilitated the consensus building among the staff on implementing effective research-based instruction and refining the evidence-based decision-process.
Administrative support	1. School 1dministrators provided the materials and equipments for implementing effective instruction and progress monitoring. 2. Principal hired an IA to administer progress monitoring and manage student data for the entire school. 3. School 1dministrators supported teachers in dealing with attendance or behavioral issues.
Professional training and on-going support	School district provided teachers and instructional assistants training and on-going technical support in implementing research-based reading instruction and progress monitoring.
Professional collaboration	1. Teachers created schedules to support tiered interventions so that at-risk students would not react negatively for receiving interventions and missing fun activities. 2. Teachers collaborated to create "cascade of intervention groups" that matched students' ability level and skill deficits with assigned intervention. 3. Teachers coordinated the content and pacing of instruction so that students could benefit from pre-teaching and tiered interventions. 4. Teachers worked as a multidisciplinary team, sharing resources and expertise in teaching, assessing students, and problem-solving.

Data sources: teacher interviews, document review (of school schedules and student progress monitoring data), classroom observation and observation of school-based meetings.

<u>School 2 Report</u>

Response to Interventions (RTI) is comprised of six critical components: (a) universal screening, (b) Tier 1 instruction, (c) Tier 2 and Tier 3 interventions, (d) progress monitoring, (e) evidence-based decision-making and (f) organizational support.
A reviewer can read a school report and judge the school's implementation of RTI by its implementation of these six components using the *RTI Assessment Rubric*.

To help reviewers frame school data in a proper context, I first describe the method of data collection, and then report the findings.

Method

I sampled and documented a school's implementation of RTI at one grade-level. This report documents not only the design and execution of the plan, but also the living experience of individual at-risk students receiving reading instruction under the RTI model. The data collection process was guided by Yin's (2003) case study methodology. The guiding principle of Yin's methodology was to gather data from multiple sources so the credibility of findings can be enhanced by the convergence of evidence.

First, I interviewed teachers and administrators before and after the classroom observation. The purpose of these interviews was to understand the school's plan of RTI and to confirm or clarify the findings in direct observation and document review. Second, I reviewed school schedules, class rosters, student assessment data, and teacher-drafted documents such as grade-level meeting protocol, intervention map and assessment grid. The purpose of document review is to depict the school's plan and implementation on instruction and assessment at the grade-level. Third, I shadowed selected at-risk students for a week to observe the Tier 1, Tier 2 and Tier 3 interventions they received within that time frame. I reported the cumulative minutes the students spent on different instructional activities, the percentage of time students appeared on task, and the teacher's scores on the classroom observation instrument. I also supplemented qualitative description of how teachers used instruction to address students' skill deficits. The purpose of classroom observation is to describe the actual implementation of the RTI model, from the perspective of an at-risk student. Next, I observed school meetings to document how teachers used assessment data to make instructional decisions. For additional information about the method of data collection, please refer to the methods section of my dissertation.

Results

This school report documents School 2's implementation of *Response to Intervention (RTI)* using its implementation at Grade 2 as an example. The report first described the characteristics of school, participating teachers and students (see Tables 1-3), the district's RTI model (Table 4), the school's design of tiered interventions (Table 5). Next, it provided documentation of the implementation of the six identified RTI critical components: (a) *universal screening*, (b) *Tier 1 instruction*, (c) *Tier 2 and Tier 3 interventions*, (d) *progress monitoring*, (e) *evidence-based decision-making* and (f) *organizational support*. For each identified component, I provide a brief narrative and a table to summarize the findings (see Tables 6-10).

Table 1
School characteristics

	School 1	School 2
School Structure	K-5, straight grades, 3 classes per grade	K-5, 2 classes per grade for Grades K-3 and 4 Grades 4/5 combined classes
Enrollment	500+	<300
Special programs (e.g. Title I, ELL programs)	Non -Title 1 school No ELL programs Regional Learning Center on campus	Non-Title 1 school No ELL programs Regional Learning Center on campus
Other instructional supports	Reading Specialist, Student achievement coordinators	Reading specialist
Prior experience in progress monitoring before adopting RTI Model	Yes, they used DIBELS for screening and progress monitoring.	No prior experience in progress monitoring

Data source: *MDP-RTI Study: School Profiles.*

Table 2

Teacher Characteristics

School	School 1				School 2	
Teacher	1	2	3	J	4	5
Primary assignment	General education	General education	Special education	Instructional Assistant	General education	Special education
Highest level of education	BA	MA	MA	BA	MA	MA
Teaching certification	Yes, general education	Yes, general education	Yes, general education	certified as Instructional Assistant	Yes gneral education	Yes, special education
Years of teaching experience (with students with disabilities)	6 (3)	10 (10)	1 (11)	8 (8)	26 (26)	29 (29)
Instructional role(s) in RTI	Tier 1 Tier 2	Tier 1 Tier 2	Tier 2	Tier 2 Tier 3	Tier 1 Tier 2	Tier 2 Tier 3

Data Source: *MDP_RTI Study: Teacher Survey.*

Table 3

Student Characteristics

School	School 1			School 2	
Student[1]	Carl	Diane	Eric	Albert	Becky
Grade	2	2	2	2	2
Gen. ed. teacher	Teacher 1	Teacher 1	Teacher 2	Teacher 4	Teacher 4
Race	White	Unknown	White	White	White
Primary language	English	English	English	English	English
Have IEP?	No	No	No	No	No
Have been retained in grade	No	No	No	No	No
Screening score in cwpm (Percentile)	8 $(1\text{-}10^{th})$	24 $(11\text{-}20^{th})$	29 $(11\text{-}20^{th})$	10 $(1\text{-}10^{th})$	17 $(11\text{-}20^{th})$
Recommended to receive Tier 3 interventions	Yes	Yes	Yes	Yes	Yes

Data Source: *MDP_Student Characteristic Surveys*; district assessment data; teacher interviews.

1: Grade 2 general education teachers nominated three lowest students in their classes that fit the selection criteria: (a) scoring at or below 20th percentile in the district's fall reading assessment, and (b) being recommended to receive Tier 3 interventions. When three students' parents from a class all gave consent, two with the lowest scores were chosen. In Teacher 2's class, only one student's parents gave consent.

Table 4

Description of District Model

Instruction					
Tiers of instruction	Focus	Targeted students	Delivered by whom	How long?	How often?
Tier 1	Core reading instruction, focusing on the "big 5" as defined by NRP (2000).	All students	General education teacher	40-90 minutes	daily
Tier 2	More differentiated and skill oriented than Tier 1 instruction.	All students	General education teachers, special education teachers, reading specialists, IAs, and other qualified personnel	30 minutes	daily
Tier 3	Small group, targeted, explicit and direct instruction that matches students' instructional needs	Students below 20th percentile in screening, and their progress below the aimline (3+ data points)		60 minutes	weekly

Assessment				
Types of assessments	Purpose of assessment	What test?	Target population	How often?
Screening	Identify at-risk students	EasyCBM	All students	2-3 times a year
Progress Monitoring	Determine if students respond to interventions	EasyCBM	students below 20th percentile	Biweekly

Data Source: *District's RTI model: Instructional Intervention Progress Monitoring Model (IIPM model).*

Table 5: *Summary of Instructional Plan: School 2, Grade 2*

Level of Instructional Support	Who? (What Skill-level of Students)	With What? (Which Materials & Activities)	What More? (Supplemental or Additional Support Materials and Activities)	How Are We Doing? (Determining Instructional Effectiveness with Progress Monitoring)
Tier 1 / Primary:	Which Students: **All Students**	Name of Program / Materials: **Houghton Mifflin Reading** When: **8:30 – 9:30** Activities: **All activities in the core** Group Size: **Large and some small group (3-8)**	Name of Program / Materials: **Not Applicable**	Who to Collect: **IA in charge of progress monitoring** How Often: **Beginning, Middle, End of Year** Criteria: **See EasyCBM test result guidelines** Determining Fidelity of Implementation (Who, With What, How often?): **To be determined**
Tier 2 / Secondary:	Which Students: **All students**	Name of Program / Materials: **Houghton Mifflin Reading** When: **10:00-10:30** Activities: **All activities in the core** Group Size: **5-8**	Name of Program / Materials: **Not applicable** Who to Deliver: **General Ed teacher** When: **__X__ w/in typical instructional time** Specify Time (minutes, days of week): **8:05-8:35 M-F** Group Size: **3-8**	Who to Collect: **IA in charge of progress monitoring** How Often: **some students biweekly progress monitoring without Tier 3 intervention.** Criteria: **teacher nomination** Determining Fidelity of Implementation (Who, With What, How often?): **to be determined**
Tier 3 / Tertiary:	Which Students: **Students scored at or below 20th percentile of the district norm** As measured by: **the EasyCBM screening measures**	Name of Program / Materials: **Horizon** When: **8:30-9:30** Activities: **Vocabulary and oral reading /comprehension activities** Group Size: **Small Group (4 or less)**	Name of Program / Materials: **Horizon** Who to Deliver: **Special Ed teacher** When: **__X__ w/in typical instructional time __ __ in addition to typical time Every day for an hour** Specify Time (minutes, days of week): **Monday- Friday 8:30-9:30** Group Size: **4 or less**	Who to Collect: **IA in charge of progress monitoring** How Often: **Biweekly** Criteria: **in or above the aimline** Determining Fidelity of Implementation (Who, With What, How often?): **Not yet decided**

Universal Screening

   *Interviews with teachers. School 2*scheduled to screen all students in reading three times a year (fall, winter and spring), using district reading assessment (EasyCBM Measures). Teachers used the fall screening data to identify students who need Tier 3 interventions. They also used the scores to determine the instructional grouping for Tier 2 interventions. The principal and Grade 2 teachers believed EasyCBM Measures was appropriate because the district approved it.

   Document review. Technical adequacy of both measures can be found in the test developer's technical report (Alonzo & Tindal, 2007). The newly developed EasyCBM Measures have not received any independent reviews.

   The fall benchmark assessment of EasyCBM was administered to all students in September 2008. In Grade 2, EasyCBM tested students on word reading fluency (WRF), passage reading fluency (PSF) and multiple-choice reading comprehension (MCRC). The EasyCBM screening data reported students' correct words per minute (CWPM) and their percentile ranking in parenthesis. Grade 2 data were presented in a class roster by the order of students' ORF scores.

   Table 7 provided an example of documentation of implementation of screening, using the participating students' fall screening data. The table reported (a) screening measures used, (b) test dates, (c) students' scores, (d) the percentile ranking and cut scores that helped teachers interpreting test results and (e) the data sources.

Alonzo, J. & Tindal, G. (2007). Examining the Technical Adequacy of Word and Passage Reading Fluency Measures in a Progress Monitoring Assessment System (Technical Report #40). Eugene, OR: Behavioral Research Teaching.

Good, R. H. & Kaminiski, R. A. (2002). DIBELS Oral Reading Fluency Passage for First Through Third Grade (Technical Report No. 10). Eugene, OR: University of Oregon.

Table 7

School 2 Targeted Students' Fall Screening Data

	Chosen Screening Measure			
	EasyCBM In Correct Words Per Minute (CWPM) (range of %ile ranking)[1]			Composite Percentile Ranking
Students	WRF[2]	PSF[3]	MCRC[4]	
Albert	14 (11-20)	10 (1-10)	1	6
Becky	8 (1-10)	17 (1-10)	4	11

6. Red = 1-10 percentiles. Yellow = 11- 20 percentiles.
7. WRF (Word Reading Fluency). Grade 2 fall: 10%ile =12 cwpm, 20%ile = 20 cwpm, 50%ile = 40 cwpm, 75%ile = 62 cwpm, 90%ile = 80 cwpm.
8. PSF (Passage Reading Fluency). Grade 2 fall: 10%ile =18 cwpm, 20%ile = 30 cwpm, 50%ile = 57 cwpm, 75%ile = 84 cwpm, 90%ile = 127 cwpm.
9. MCRC (Multiple-choice Reading Comprehension. Grade 2 fall: 10%ile =3, 20%ile = 5, 50%ile = 8, 75%ile = 9, 90%ile = 11. Total possible score = 12.
10. Data Sources: teacher interviews, school schedule, School 2EasyCBM Grade 2 fall benchmark report, EasyCBM progress monitoring scores interpretation guidelines.

Tier 1 Intervention

Tier 1instruction is the core curriculum instruction for all students. The findings were derived from interviews with teachers, review documents and weeklong of classroom observation in which I shadow targeted at-risk students.

Document review. School master schedule designated 8:30-9:30 for the Grade 2 reading time, 10:30-11:30 for open block. In class schedule, the open block was further specified. On Mondays, Tuesdays and Thursdays, 10:30-11:00 was designated for language arts. The story time was scheduled at 11:00-11:30 daily (see Table 6). It is noted in the class schedule that Albert and Becky were to leave to receive reading instruction from special education teacher between 8:30-9:30.

Interview with teachers. School 2scheduled Grade 2 Tier 1 instruction at 8:30-9:30 daily. The core reading curriculum was Houghton Mifflin's Reading Program, one of the two research-based core curriculum approved by the school district. Teacher 4 and principal stated that the emphasis for Tier 1 instruction was to teach all of 'the big five ideas': phonemic awareness, phonics, fluency, vocabulary and comprehension and to "help students meet and exceed grade-level benchmarks (in reading). Albert and Becky received daily one hour of small group reading instruction from the special education teacher while   peers received Tier 1 instruction from Teacher N. Principal counted the time as Tier 1 plus Tier 3 intervention time. Teacher 4 counted the time with special education teacher as Tier 2 plus Tier 3 interventions. Teacher 4 acknowledged that it is challenging to teach whole class reading under the current schedule (see Table 6) because students left for different intervention groups at different times.

Classroom observation. During the week of classroom observation, Teacher 4 delivered Tier 1 instruction to the whole class *daily* at the designated reading time. However, Albert and Becky were not in the room to receive it. The Tier 1 instruction Albert and Becky received from their homeroom was mostly made up by story time and writing time. In the story time, Teacher 4 read aloud the chapter books to the whole class and conducted literary discussion with students. Albert and Becky both appeared listening attentively and sometimes they raised hands answering teacher's comprehension questions. In writing time, students were brainstorming story starters together. Teacher wrote the key vocabulary on the board and students proceeded to write the passage independently. While the students were writing, Teacher 4 either circulated the room or writing words on the sticker notes per students' requests. In the writing time, some students spent significant amount time in illustrating and standing in line to ask the spelling of words. In a 20-minute of writing session, average students in this class wrote 2-5 complete sentences with appropriate capitalization, punctuation and some inventive spelling. Albert wrote two complete sentences by copying the dictation Teacher wrote for him. Becky wrote the following sentence independently: "Mi bruthir uent to the pukin" (My brother went to the pumpkin patch.).

Interviews with students. When asked how their homeroom teacher teaching them to read, Albert said, "She [Mrs. G] told me the words I don't know. She made me sound out words". Becky said, "She read the best stories. Her books are awesome." They both liked reading at school, even though reading made them "tired and confusing" at time.

Table 8

Instructional Activities by Time during Tier 1, Tier 2 and Tier 3 interventions

Instructional Activities		Tiers of Interventions		
Teacher vs. student lead activities	Instructional emphasis	Tier 1	Tier 2	Tier 3
Teacher-lead activities	Phonemic Awareness	0	0	2.7
	Phonics	0	11.5	124.2
	Fluency	0	11.4	12.7
	Vocabulary	0	9.1	10
	Comprehension	0	3.2	12.8
Student-lead activities	Read aloud[1]	12.2	0	27
	Silent reading[2]	16.5	0	36.6
	Independent seatwork[3]	0	0	0
Total Time on the targeted reading activities		28.7	104.4	226.7
Total Time on the other activities, such as creative writing, arts and crafts and transition, discipline and house-keeping		108.9	45.6	39.8
Total Time for Reading Instruction		137.6	150	266.5
% of time spent on the academic activities		20.9	69.6	85
% of time students appear on task		50-80	50-80	90+
Average Scores of *Classroom Observation Instrument*[4]		N/A	37 (32-42)	432 (42-44)

1: Students read list of words or connected text independently or with minimal teacher scaffolding.

2: Students read the assigned passages or books of their choice quietly. Some silent reading was followed by teacher asking vocabulary or comprehension questions.

3: Students work on assigned worksheet directly related to the phonics spelling patterns, vocabulary, and comprehension strategies, skills.

4: *Classroom Observation Instrument* consisted of 24 items. Likert Rating were 0 = not effective, 1=partially effective and 2= effective. In high-performing classes, teachers on average scored 40 out of the 48 possible scores.

Data sources: Classroom observation.

Tier 2 and Tier 3 interventions

Table 5 provides an overview of school's tiered intervention plan. Table 6 showed the class schedule in which schedules for different intervention groups was embedded.

Document review. The class schedule (see Table 6) showed reading specialist took one of the Tier 2 groups during the story time; and two instructional assistants took some students some time during the workshop time. Special education teacher took Tier 3 group during the designated Tier 1 reading time.

Teacher interviews. Albert and Becky received Tier 2 instruction from Teacher 4 and sometimes from an instructional assistant; and they received Tier 3 instruction from special education teacher. Teacher 4 used her own basal program in Tier 2 interventions. She first taught

students the targeted vocabulary and then asked students to read the sentences. Special education teacher used a research-based intervention program *Horizons* for Becky and Albert's Tier 3 group. Both teachers focused on decoding and sight word reading, emphasizing read with accuracy, appropriate pacing and prosody.

Review and observation of school-based meetings. Two second-grade teachers met to discuss their students' progress in a grade-level meeting. In the meeting, the teachers reviewed student progress monitoring data by class roster; and they agreed that keeping Albert and Becky in the current instructional group was the best available choice.

Classroom observations. Teacher 4 worked with multiple small groups in rotation during Tier 2 intervention time. It is noted that the rotation of instructional group was often improvised. When the teacher was working with one group of students, the other students were assigned to read book of their choices silently. But they drew, weaved, flipped through pages or chatted with peers till being redirected by teacher. Teacher 4 and an instructional assistant provided Albert and Becky at least 20 minutes of small group ($n = 2$ or 6) differentiated reading instruction four days of the week. Wednesday's lesson was cancelled due to dance practice and music lesson. Teacher used first grade basal readers and the instructional focus was on decoding and fluency. In classroom observation, student "on task" is defined as "sitting in their seats properly, eyes looking either at teachers or at the instructional materials attentively, focusing on assigned reading or writing activities". Albert and Becky were "on task" 80-90% of time when reading with Teacher or IA; but they appeared sometimes off task in assigned silent reading time. As a result, the percentage on task time varied widely (Table 8).

Special education teacher, Teacher 5, provided Albert and Becky 60 minutes of daily reading instruction in a small group ($n=4$), using a research-based reading intervention program, *Horizons*. The teacher followed the teacher's manual closely for the decoding, oral reading fluency and spelling. Albert and Becky were given multiple opportunities to response in chorus or independently. Most of the time they finished the task correctly the first time. Teacher 5 demonstrated most targeted teaching behaviors in effective level. Albert and Becky appeared on task 90+% of the time (Table 8).

Interviews with students. Albert and Becky stated both Teacher 4 and Teacher 5 showed them how to sound out words. Albert and Becky both liked to do reading in small groups. They found comprehension was easier than reading words. They found Teacher 4's books were more interesting.

Progress Monitoring

School 2 used district reading assessment (EasyCBM) for progress monitoring; it scheduled to monitor students who scored at or below 20th percentile of district norm in fall screening in reading every other week. For Grade 2, the progress monitoring measures was individually administered one-minute passage reading fluency.

Interviews with teachers. Teachers and principal stated that an instructional assistant (IA) was assigned to administer the tests to targeted students in Grade K-5 (except Grade 2, which was administered by school psychologist) every other week. The IA also were responsible for entering test scores into a password-protected district website, and reporting data to general education teachers and principal.

Document review. Table 9 and Figure 1 showed the progress monitoring data for the targeted students. It indicated test dates, students scores and reference scores that helped teachers interpreted the test data.

Technical adequacy of EasyCBM word reading fluency and passage reading fluency can be found in the test developer's technical report (Alonzo & Tindal, 2007). It provided strong evidences for internal reliability and some evidences for validity. Easy CBM has not been reviewed independently.

Table 9:

Student Progress Monitoring Data: in Correct Words per Minute

	Screening	10/16/2008	10/28/2008	11/12/2008	Annual Goal
Albert	10	15	20	25	110
Becky	17	20	18	29	110

Passage Reading Fluency Second Grade Fall Data Benchmarks: 10th percentile = 18 cwpm, 20th percentile = 30 cwpm,50th percentile = 57 cwpm. A = Absent.

Albert's Progress Monitoring Data

Becky's Progress Monitoring Data
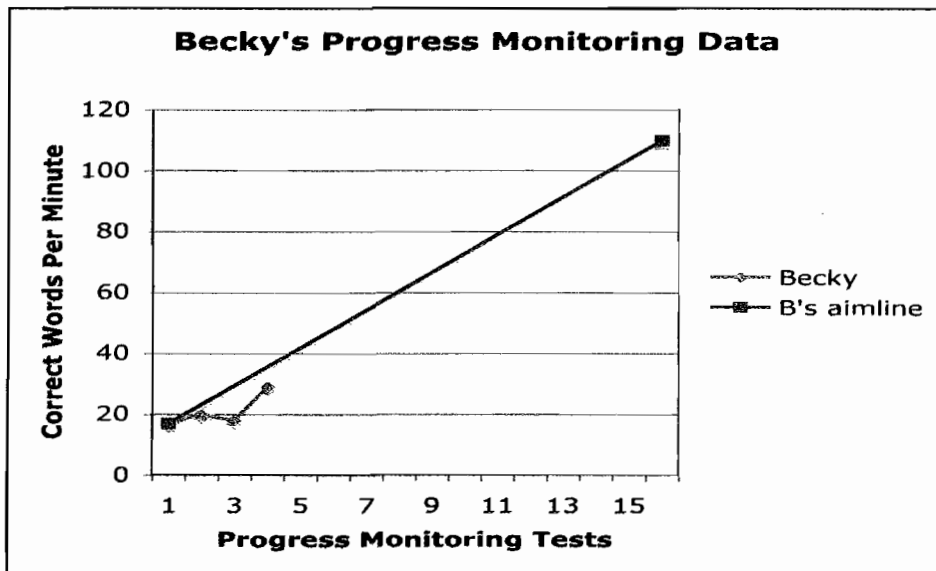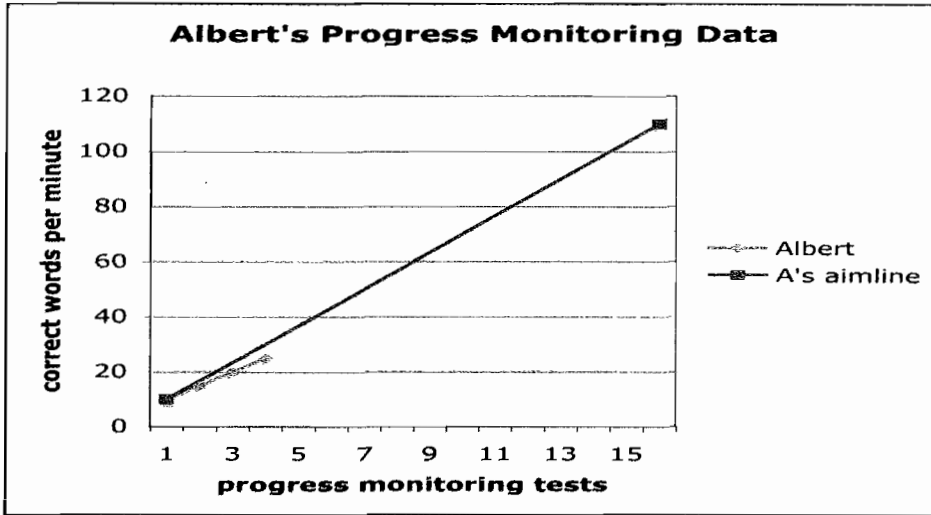
Evidence-based Decision-making

Teacher interviews. Teachers stated that they used EasyCBM progress monitoring data as the primary data source for decision-making. The principal and teacher both stated that they were still learning about how to interpret the progress monitoring data and use them to improve instruction. They do not have protocols to conduct the data meetings or explicit decision rules.

Observation of school-based meetings. I observed a grade level meeting on November 18. Two-second grade teachers attended Grade 2 data meeting and reviewed the progress monitoring data of all Grade 2 students. In the data meeting, two teachers comments on impacts of fragmented instruction schedule and curriculum programs. They discussed their disagreement with the reading specialist on the assessment with student's progress. Both teachers felt hesitated to mix two classes of students during the Tier 2 interventions and at the same time felt challenging to provide cascade of interventions to students with wide range of skill differences. The teachers had questions about the reference scores of the EasyCBM progress monitoring measures. Both teachers reviewed Albert and Becky's data. Teacher 4 provided additional information about students' classroom performance. She commented on the students' commitment and stress in learning to be an independent reader as well as Albert's and Becky's inability to retain previously taught sight words. Neither of the second grade teachers was satisfied with these students' progress. However, they decided to keep the students in the current intervention group because it would be the similar arrangement, had students being identified to be eligible to receive special education.

Table 10

Documentation of Evidence-based Decision-making

		Albert	Becky
Data used: EasyCBM progress monitoring data	3 data points since last data meeting	✓	✓
	Summary of data based on the graph	Upward trend; below aimline.	No trend detected; below aimline
Teacher input	Attendance	Some absence	Good
	Classroom Performance	Good, close monitoring is needed	Good, close monitoring is needed
	Additional Assessments	No	No
	Others (Specified)	School 2 is his 7th school	Mom resisted the idea of testing
Decision-making process	Follow decision rules	No explicit decisions rules	No explicit decision rules
	Multi-disciplinary decision-making[1]	Decisions made by two general education teachers	Decisions made by two general education teachers
Decisions made	Change of current interventions	NO	NO
	Refer to receive special education	NO	NO

Data Source: Progress monitoring data and observation of school-based meetings.

Organizational Support

Organizational support can manifest in instructional leadership, administrative support, availability of professional training and on-going support, and professional collaboration. Table 11 provided examples of organizational support in School 2.

Table 11

Documentation of Organizational Support

Types of Support	Examples
Instructional leadership	5. Principal and leadership team allocated resources for purchasing research-based intervention programs. 6. Principal allocate resources for implementing small group instruction and progress monitoring. 7. Principal facilitated the consensus building among the staff on implementing effective research-based instruction and refining the evidence-based decision-process.
Administrative support	4. School administrators provided the materials and equipments for implementing effective instruction and progress monitoring. 5. Principal hired an IA to administer progress monitoring and manage student data for the entire school.
Professional training and on-going support	School district provided teachers and instructional assistants training and on-going technical support in implementing research-based reading instruction and progress monitoring.
Professional collaboration	5. Two Grade 2 teachers had frequent informal discussions about their students' progress. 6. Teacher 4 had high percentage of at-risk students and went to principal to request additional support. Reading specialist and special education teacher volunteered to provide some of Teacher 4's students Tier 2 and Tier 3 interventions. The interventionists chose students, intervention curriculum and time for the groups to meet. Teacher 4 appreciated colleagues' support. 7. Teacher 4 asked the interventionists about her students' performance in the intervention groups. The check-in was informal and episodic. 8. Teacher 4 was not familiar with the intervention programs used in other teacher's classes. She and the reading specialists disagreed on which students had higher needs and should receive more intensive interventions.

Data sources: teacher interviews, document review (of school schedules and student progress monitoring data), classroom observation and observation of school-based meetings.

REFERENCES

AERA, APA, & NCME. (1999). *Standards for educational and psychological testing.* Washington, D.C.: American Educational Research Association.

Allington, R. (1983). The reading instruction provided readers of differing reading abilities. *The School Psychologist, 83*(5), 548-559.

Allington, R. (2002). What I've learned about effective reading instruction from a decade of studying exemplary elementary classroom teachers. *Phi Delta Kappan, 83*(10), 740-744.

Alonzo, J., Tindal, G., Ulmer, K., & Glasgow, A. (2006). easyCBM online progress monitoring assessment system. . Eugene, Oregon: Behavior Research and Teaching, University of Oregon.

Batsche, G. M., Elliott, J., Graden, J., Grimes, J., Kovaleski, J. F., & Prasse, D., et al. (2005). *Response to Intervention: Policy considerations and implementation.* Alexandria, VA: National Association of State Directors of Special Education, Inc.

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: sustaining the doing, supporting the learning. *Educational Psychologist, 26(3/4),* 369-398.

Bollman, K. A., Silberglitt, B., & Gibbons, K. A. (2007). The St. Croix River education district model: Incorporating system-level organization and a multi-tiered problem solving Process for intervention delivery. In S. R. Jimerson, M. K. Burns & A. M. VanDerHeyden (Eds.), *The handbook of Response to Intervention: The Science and Practice of Assessment and Intervention*, 319-330. New York: Springer.

Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: a meta-analysis. *Review of Educational Research, 73*, 125-231.

Burns, M. K., Appleton, J. J., & Stehouwer, J. D. (2005). Meta-analysis review of responsiveness-to-intervention research: Examining field-based and research implemented models. *Journal of Psychoeducational Assessments*, 381-394.

Burns, M. K., & Coolong-Chaffin, M. (2006). Response to Intervention: The role of and effect on school psychology. *School Psychology Forum: Research in Practice, 1*(1), 3-15.

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.

Callender, W. A. (2007). The Idaho result-based model: implementing Response to Intervention statewide. In S. R. Jimerson, M. K. Burns & A. M. VanDerHeyden (Eds.), *The Handbook of Response to Intervention: The science and practice of assessment and intervention*, 331-342. New York: Springer.

Carnine, D. (1997). Bridging the research-to-practice gap. *Exceptional Children, 63*(4), 513-521.

Carnine, D., Silbert, J., & Kame'enui, E. J. (1990). *Direct instruction reading*. Columbus, OH: Merrill.

Chard, D. J. (2004). Toward a science of professional development in early reading instruction. *Exceptionality, 12*(3), 175-191.

Chard, D. J., Stoolmiller, M., Harn, B., Wanzek, J., Vaughn, S., Linan-Thompson, S., et al. (2008). Predicting reading success in a multilevel schoolwide reading model: a retrospective analysis. *Journal of Learnign Disabilities, 41*(2), 174-189.

Christ, T. J., Burn, M. K., & Ysseldyke, J. E. (2005). Conceptual confusion within Response-to-Intervention vernacular: clarifying meaningful differences. *NASP Communique, 34*(3).

Christ, T. J., & Hintze, J. M. (2007). Psychometric consideration when evaluating Response to Intervention. In S. R. Jimerson, M. K. Burns & A. VanDerHeyden (Eds.), *Handbook of Response to Intervention*, 93-105. New York, NY: Springer.

Coyne, M. D., Kame'enui, E. J., & Simmons, D. C. (2001). Prevention and intervention in beginning reading: two complex systems. *Learning Disabilities Research & Practice, 16*(2), 62-73.

Coyne, M. D., Kame'enui, E. J., Simmons, D. C., & Harn, B. A. (2004). Beginning reading intervention as inoculation or insulin: first-grade reading performance of strong responders to kindergarten intervention. *37*(2), 90.

Crawford, L., Tindal, G., & Stieber, S. (2001). Using oral Reading Rate to Predict Student Performance on Statewide Achievement Tests. *Eudcational Assessment, 7*(4), 303-323.

Darling-Hammond, L., & Mclaughlin, M. W. (1995). Policies that support professional development in an era of reform. *76*(8).

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

Denton, C. A., Fletcher, J. M., Anthony, J. L., & Frances, D. J. (2006). An evaluation of intensive intervention for students with persistent reading difficulties. *Journal of Learning Disabilities, 39*(5), 447-466.

Denton, C. A., Vaughn, S., & Fletcher, J. M. (2003). Bringing research-based practice in reading intervention to scale. *Learning Disabilities Research & Practice, 18*(3), 201-211.

Desimone, L. (2002). How can comprehensive school reform models be successfully implemented? *Review of Educational Research, 77*, 433-480.

Dickson, S. V., & Bursuck, W. D. (1998). Implementing a model for preventing reading failure: a report from the field. *Learning Disabilities Research & Practice, 14*, 191-195.

Downing, S. M. (2006). Twelve steps for effective test development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 3-26). Mahwah, N. J.: Lawrence Erlbaum.

Downing, S. T., & Haladyna, T. M. (Eds.). (2006). *Handbook of test development*. Mahwah, NJ: Lawrence Erlbaum.

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Education Research, 18*(2), 237-256.

Edmonds, M., & Briggs, K. L. (2003). The instructional content emphasis instrument. In S. Vaughn & K. L. Briggs (Eds.), *Reading in the classroom*. Baltimore: Paul H. Brooks.

Elmore, R. F. (1996). Getting to scale with good educational practice. *Harvard Educational Review, 66*(1), 1-26.

Engelmann, S., & Carnine, D. (1991). Theory of instruction: principles and applications.

Ervin, R. A., Schaugency, E., Goodman, S. D., & McGlinchey, M. T. (2007). Moving from a model demonstration project to a statewide initiative in Michigan: Lessons learned from merging research-practice agenda to address reading and behavior. In S. R. Jimerson, M. K. Burns & A. M. VanDerHeyden (Eds.), *The handbook of Response to Intervention: the science and practice of assessment and intervention* (pp. 354-377). New York: Springer.

Fletcher, J. M., Francis, D. J., Morris, R. D., & Lyon, G. R. (2005). Evidence-based assessment of learning disabilities in children and adolescents. *Journal of Clinical Child & Adolescent Psychology, 34*(3), 506.

Foorman, B. R., & Ciancio, D. J. (2005). Screening for secondary intervenitons: concept and context. *Journal of Learning Disabilities, 38*, 494-499.

Foorman, B. R., & Moats, L. C. (2004). Conditions for sustaining research-based practices in early reading instruction. *Remedial & Special Education, 25*(1), 51-60.

Foorman, B. R., & Schatschneider, C. (2003). Measurement of teaching practice during reading/ language arts instruction and its relationship to student achievement. In S. Vaughn & K. L. Briggs (Eds.), *Reading in the classroom.* Baltimore: Brooks.

Foorman, B. R., & Torgesen, J. (2001). Critical elements of classroom and small-group instruction promote reading success in all Children. *Learning Disabilities Research & Practice, 16*(4), 203-212.

Friend, M. (2000). Myths and misunderstandings about professional collaborations. *Remedial & Special Education, 21*(3), 130-132.

Fuchs, D., & Deshler, D. D. (2007). What we need to know about Responsiveness To Intervention (and shouldn't be afraid to ask). *Learning Disabilities Research & Practice, 22*(2), 129-136.

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, L. S. (1995). *Curriculum-based measurement and eligibility decision-making: An Emphasis on treatment validity and growth.* Washington, DC: National Research Council.

Fuchs, L. S. (1998). Treatment fidelity : A unifying concept for reconceptualizing the identification process of learning disabilities. *Journal of Learning Disabilities.*

Fuchs, L. S. (2003). Assessing intervention responsiveness: conceptual and technical issues. *Learning Disabilities Research & Practice, 18*(3), 172-186.

Fuchs, L. S., & Fuchs, D. (2006). A framework for building capacity for Responsiveness to Intervention. *School Psychology Review, 35*(4), 621.

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

Fullan, F. (2000). The three stories of education reform. *Phi Delta Kappan, 81*(8), 581-584.

Gambrell, L. B., Wilson, R. M., & Gantt, W. (1981). Classroom observation of task-attending behaviors of good and poor readers. *Journal of Educational Research, 74*(6), 400-404.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Kwang, S. Y. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915-945.

Gerber, M. M. (2005). Teachers are still the test: limitations of response to instruction strategies for identifying children with learning disabilities. *Journal of Learning Disabilities, 38*(6), 516.

Gersten, R., Chard, D., & Baker, S. (2000). Factors enhancing sustained use of research-based instructional practices. *Journal of Learning Disabilities, 33*(5), 445-457.

Gickling, E. E., & Armstrong, D. I. (1978). Levels of instructional activities as related to on task behavior, task completion, and comprehension. *Journal of Learning Disability, 11*, 559-566.

Glickman, C. D., Gordon, S. P., & Ross-Gordan, J. M. (2001). *Supervision and instructional leadership: a developmental approach* (6th ed.). Nedham Height, MA: Ally & Bacon.

Good, R. H., Gruba, J., & Kaminski, R. A. (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 IV*, 679-700.

Good, R. H., & Kaminski, R. A. (2003). *Dynamic Indicators of Basic Literary Skills (DIBELS): Administration and scoring guide (6th Edition)*. Eugene, OR: Sopris West.

Good, R. H., Simmons, D., & Kame'enui, E. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third grade high stakes outcomes. *Scientific Studies of Reading, 5*, 257-288.

Good, R. H., Simmons, D., Kame'enui, E., & Chard, D. (2003). *Operationalizing Response to Intervention in eligibility decisions.* Paper presented at the National Research Center on Learning Disabilities Response to Intervention Symposium, Kansas City, MO.

Good, R. H., Simmons, D. C., Kame'enui, E. J., Kaminski, R. A., & Wallin, J. (2002). *Summary of decision rules recommendations in kindergarten through third grade* (Technical Report No. 11). Eugene, OR: University of Oregon.

Graden, J. L., Stoller, S. A., & Poth, R. L. (2007). The Ohio Integrity System Model: an overview and lessons learned. In S. R. Jimerson, M. K. Burns & A. M. VanDerHeyden (Eds.), *Handbook of Response to Intervention: The Science and Practice of Assessment and Intervention*, 288-299. New York: Springer.

Graves, A. W., Gersten, R., & Hagger, D. (2004). Literacy instruction in multiple-language first-grade classrooms: linking student outcomes to observed instructional practice. *Learning Disabilities Research & Practice, 19*(4), 262-272.

Greenwood, C. R., Abott, M., & Tapia, Y. (2003). Ecobehavioral strategies: observing, measuring, and analyzing behaviors and reading interventions. In S. Vaughn & K. L. Briggs (Eds.), *Reading in the classroom.* Baltimore, NY: Brooks.

Grek, M. L., Mathes, P. G., & Torgesen, J. (2003). Similarities and differences between experienced teachers and trained paraprofessional: an observational analysis. In S. Vaughn & K. L. Briggs (Eds.), *Reading in the classroom.* Baltimore, N.Y.: Brooks.

Gresham, F. M. (2002). Responsiveness to intervention: An alternative approach to the identification of learning disability. *Identification of Learning Disabilities: Research to Practice,.*

Gresham, F. M., MacMillan, D. L., Beebe-Frankenberger, M. E., & Bocian, K. M. (2000). Treatment integrity in learning disabilities intervention research: Do we really know how treatments are implemented? *Learning Disabilities Research & Practice, 15*, 198-205.

Gresham, F. M., Reschly, D. J., Tilly, W. D., Fletcher, J., Burns, M., & Prasse, D. (2005). A response to intervention perspective. *The School Psychologist, 59*(1), 26-33.

Grimes, J., & Kurns, S. (2003). *An intervention-based system for addressing NCLB and IDEA expectations: A multiple tiered model to ensure every child learns.* Paper presented at the National Research Center on Learning Disabilities Responsiveness to Intervention Symposium Kansas, MO.

Hagger, D., Gersten, R., Baker, S., & Graves, A. W. (2003). The English language learner classroom observation instrument for beginning readers. In S. Vaughn & K. L. Briggs (Eds.), *Reading in the classroom*. Baltimore: Paul H. Brooks.

Harn, B., Kame'enui, E. J., & Simmons, D. C. (2007). The role of the third tier. In D. Hagger, J. Klinger & S. Vaughn (Eds.), *Evidence-based reading instruction for Response to Intervention* (pp. 161-184). Baltimore: Paul H. Brooks.

Heller, K. A., Holtzman, W. H., & Messick, S. (Eds.). (1982). *Placing children in special education: A strategy in equity*. Washington, DC: National Academy Press.

Hiebert, J. (1999). Relationship between research and NCTM standards. *Journal for Research in Mathematics Education, 30*(1), 3-19.

Hill, P. W., & Rowe, K. J. (1998). Modeling student progress in studies of educational Effectiveness. *School Effectiveness & School Improvement, 9*(3), 310-333.

Hollywood, T. M., Mott, S. E., & Dodson, C. I. (1995). Use of instructional time serving students with and without severe disability. *Exceptional Children, 61*, 242-252.

Horner, R. H., Todd, A. W., Lewis-Palmer, T., Irvin, L., K., Sugai, G., & Boland, J. B. (2004). The School-wide evaluation tool (SET): A research instrument for assessing school positive behavior support. *Journal of Positive Behavioral Support, 6*(1), 3-12.

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.

Jenkins, J. R. (2003). *Candidate measures for screening at-risk students.* Paper presented at the National Research Center on Learning Disabilities (NRCLD): "Response to Treatment as Learning Disabilities Identification", Kansas, MO.

Johnson, E., Mellard, D. F., Fuchs, D., & McKnight, M. A. (2006). *Responsiveness to intervention: How to do it.* Lawrence, KS: National Research Center on Learning Disability.

Justice, L. M. (2006). Evidence-based practice, Response to Intervention, and the prevention of reading difficulties. *Language Speech Hearing Serving School, 37*(4), 284-297.

Kagan, D. M. (1992). Implication of research on teacher belief. *Educational Psychologist, 27*(1), 65-90.

Kame'enui, E. J., & Simmons, D. C. (2003). Planning and evaluating tools for effective school-wide reading programs-revised. Retrieved Jan. 23, 2009, from http://oregonreadingfirst. uoregon.edu/download/PET-R.doc.

Kamps, D. M., & Greenwood, C. R. (2003). *Formulating secondary level reading interventions.* Paper presented at the National Research Center on Learning Disabilities Response to Intervention Symposium, Kansas, MO.

Kavale, K. A., & Forness, S. R. (2000). What definitions of learning disability say and don't say. *Journal of Learning Disabilities, 33*(3), 239-256.

Kim, A., Briggs, K. L., & Vaughn, S. (2003). The classroom climate scale: observing during reading instruction. In S. Vaughn & K. L. Briggs (Eds.), *Reading in the classroom.* Baltimore: Brooks.

Klingner, J., Sturger, K., & Harry, B. (2003). Conducting ethnographic observation of reading in elementary schools. In S. Vaughn & K. L. Briggs (Eds.), *Reading in the classroom* Baltimore, N.Y. : Brooks.

Kovaleski, J. F. (2003, December 2003). *The three-tier model for identifying learning disability: critical program features and system issues.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-intervention Symposium, Kansas City, MO.

Kraochwill, T. R., Clements, M. A., & Kalymon, K. M. (2007). Response to intervention: conceptual and methodological issues in implementation. In S. R. Jimerson, M. K. Burns & A. M. VanDerHeyden (Eds.), *The Handbook of Response to Intervention: The science and practice of assessment and intervention.* New York: Springer.

Larson, J. (1994). Violence prevention in the schools: A review of selected programs and procedures. *School Psychology Review, 23*(2), 151.

Lau, M. Y., Sieler, J. D., Muyskens, P., Canter, A., VanKeuren, B., & Marston, D. (2006). Perspectives on the use of the problem-solving model from the viewpoint of a school psychologist, administrator, and teacher from a large midwest urban school district. *Psychology in the Schools, 43*(1), 117-127.

Little, J. W. (1993). Teachers' professional development in a climate of educational reform. *Educational Evaluation and Policy Analysis, 15*, 129-151.

Lyons, G. R., Fletcher, J. M., & Barnes, M. C. (2002). Learning disability. In E. Mash, A. Schulte & R. Olson (Eds.), *Child psychology, 2*, 520-586. New York: 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.

Marston, D., Lau, M., & Muskens, P. (2007). Implementation of Problem-Solving Model in the Minneapolis Public Schools. In S. R. Jimerson, M. K. Burns & A. M. VanDerHeyden (Eds.), *The Handbook of Response to Intervention: The science and practice of assessment and intervention,* 279-287. New York Springer.

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.

Marston, D., Pickart, M., Reschly, A., Heistad, D., Muyskens, P., & Tindal, G. (2007). Early literacy measures for improving student reading achievement: translating research into practice. *Exceptionality, 15*(2), 97.

McGrew, J. H., Bond, G. R., Dietzen, L., & Salyers, M. (1994). Measuring the fidelity of implementation of a mental health program model. *Journal of Consulting and Clinical Psychology, 62*, 670-678.

McMaster, K. L., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2005). Responding to non-responders: an experimental field of identification and intervention methods. *Exceptional Children, 71*, 445-463.

McMaster, K. L., & Wagner, D. (2007). Monitoring response to general education instruction. In S. R. Jimerson, M. K. Burns & A. M. VanDerHeyden (Eds.), *The Handbook of Response to Intervention: The science and practice of assessment and intervention,* 223-233. New York: Springer.

Mellard, D. (2004). *Understanding Responsiveness to Intervention in learning disability determination*: National Research Center for Learning Disability

Mellard, D. F., Byrd, S. E., Johson, E., Tollefson, J., M., & Boesche, L. (2004). Foundations and research on identifying model responsiveness-to-intervention sites. *Learning Disability Quarterly, 27*(4), 243-256.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.), 13-103. New York: American Council on Education, MacMillan.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded source book.* (2nd ed.). Thousand Oaks, CA: Sage.

National Center on Education and the Economy. (1999). *Reading and writing grade by grade*. Pittsburgh, PA: author

National Reading Panel. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH pub. 00-4754)*: National Institute of Health.

O'Conner, R., Fulmer, D., Harty, K. R., & Bell, K. M. (2005a). Layers of reading intervention in kindergarten through third grade: change in teaching ad student outcomes. *Journal of Learning Disabilities, 38*, 440-445.

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

O'Connor, R. (2000). Increasing the intensity of intervention in kindergarten and first grade. *Learning Disabilities Research & Practice 15*(1), 43.

Peterson, D. W., Prasse, D. P., & Shinn, M. R. (2007). The Illinois flexible service delivery model: a problem-solving model initiative. In S. R. Jimerson, M. K. Burns & A. M. VanDerHeyden (Eds.), *The handbook of Response to Intervention: the science and practice of assessment and intervention*, 300-318. New York: Springer.

Pikulski, J. J. (1998). Preventing reading failure: Review effective programs. In R. Allington (Ed.), *Teaching struggling readers*, 35-45. Newark, DE: International Reading Association.

Reschly, A., Coolong-Chaffin, M., Christenson, S. L., & Gutkin, T. (2007). Contextual influences and Response to Intervention: critical issues and strategies. In S. R. Jimerson, M. Burns & A. VanDerHeyden (Eds.), *The handbook of Response to Intervention*. New York: Springer.

Rhoades, L., & Eisenberger, R. (2002). Perceived organizational support: a review of the literature. *Journal of applied psychology, 87*(4), 698-714.

Rohrbeck, C. A., Ginsberg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning intervention for elementary school students: A meta-analysis review. *Journal of Educational and Psychology, 95*(2), 240-257.

Setton, R. P., Bennett, N., & Linden, R. C. (1996). Social exchange in organizations: Perceived organizational support, leader-member exchange, and employee reciprocity. *Journal of Applied Psychology, 81*(3), 219-227.

Shinn, M. (2002). Best practice in using curriculum-based measurement in problem-solving model. In A. Thomas & J. Grimes (Eds.), *Best practice in school . psychology IV*. Bethesda, MD: National Association of School Psychologists.

Showers, B., Joyce, B., & Bennett, B. (1987). Synthesis of research on staff development: A framework for future study and state-of-the-art analysis. *Educational Leadership, 45*(3), 77-87.

Simeonsson, R. J. (1994). Promoting children's health, education, and well-Being. In R. J. Simeonsson (Ed.), *Risk, resilience and prevention: Promoting the well-being of all children.* Baltimore: Paul Brooks Publishers.

Simmons, D., & Kame'enui, E. (2006). *A Consumer's Guide to Analyzing a Core Reading Program Grades K-3: A Critical Elements Analysis.* Eugene, OR: Oregon Reading First Center, College of Education, University of Oregon.

Simmons, D., Kame'enui, E., Beck, C. T., Brewer, N. S., & Fien, H. (2003). *A Consumer's Guide to Evaluating Supplemental and Intervention Reading Programs Grades K-3: A Critical Elements Analysis.* Eugene, OR: Oregon Reading First Center, College of Education, University of Oregon.

Smith, W. F., & Andrew, R. L. (1989). *Instructional leadership: how principals make a difference* Alexandria, VA: Association of supervision and curriculum development.

Snow, C., Burns, M., & Griffin, P. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academies Press.

Speece, D. L., Case, L. P., & Molloy, D. E. (2003). Responsiveness to general education instruction as the first gate to learning disabilities identification. *Learning Disabilities Research & Practice, 18*(3), 147-156.

Speece, D. L., & Hines, S. (2008). Identifying children who require different instruction in a Response to Instruction framework. *Perspectives on Language Learning and Education 15 34-40 March 2008, 15*, 34-40.

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

Stoner, G., & Green, S. K. (1992). Reconsidering the science-practitioner model for school psychology practice. *School Psychology Review, 21*, 155-166.

Tilly, W. D. (2003, December 2003). *How many tiers are needed for successful prevention and early intervention? : Heartland Area education agency's evaluation From four to three Tiers.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas, MO.

Tindal, G., Yovanoff, P., & Alonzo, J. (2006). DIBELS in the context of reform and the measurement of reading (Vol. 1). Newport, OR: The Advocate, a publication of Oregon Reading Association.

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

Torgesen, J. K., Rashotte, C. A., Alexander, A., Alexander, J., & MacPhee, K. (2003). Progress towards understanding the instructional conditions necessary for remediating reading difficulties in older children. In B. R. Foorman (Ed.), *Preventing and Remediating Reading Difficulties: Bringing Science to Scale.* 275-298. Parkton, MD: New York Press.

Treptow, M. A. (2006). *Reading at student' frustrational, instructional and independent levels: effects on comprehension and time on task.* Minneapolis, MN: University of Minnesota Center for Reading Research.

Troia, G. A. (2005). Responsiveness to intervention roles for speech--language pathologists in the prevention and identification of learning disabilities. *Topics in Language Disorders, 25*(2), 106.

VanDerHeyden, A., Witt, J. C., & Naguin, C. (2003). The development and validation of a process for screening referrals to special education. *School Psychology Review, 32*, 204-207.

VanDerHeyden, A. M., & Jimerson, S. R. (2005). Using Response-to-Intervention to enhance outcomes for children. *The California School Psychologist, 10*, 21-32.

VanDerHeyden, A. M., Witt, J. C., & Gilberton, D. A. (2007). 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., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: to 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 of identifying students with reading/learning Disabilities. *Exceptional Children, 69*(4), 391-409.

Vaughn, S., Wanzek, J., Linan-Thompson, S., & Murray, C. S. (2007). Monitoring response to supplemental service for students at risk for reading difficulty: high and low responders. In S. R. Jimerson, M. K. Burns & A. M. VanDerHeyden

(Eds.), *The Handbook of Response to Intervention: the science and practice of assessment and instruction*. Baltimore: Springer.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., & Chen, R. S. (1996). Cognitive profiles of difficult to remediate and readily remediated poor readers: early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific learning disability. *Journal of Educational Psychology, 88*, 601-638.

Vellutino, F. R., Scanlon, D. M., Small, S. G., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between reading disabled and non-reading disabled children: evidence for the role of kindergarten and first grade intervention. *Journal of Learning Disabilities, 38*, 157-169.

Vellutino, F. R., Scanlon, D. M., & Tanzman, M. S. (1998). The case for early intervention in diagnosing specific learning disability. *Journal of School Psychology, 36*, 367-397.

Walker, H. M., Horner, R. H., Sugai, G., Bullis, M., Sprague, J. R., Bricker, D., et al. (1996). Integrated approaches to preventing antisocial behavior patterns among school-age children and youth. *Journal of Emotional and Behavioral Disorders, 4*(4), 194-209.

Webb, N. R. (2006). Content for achievement tests. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 155-180). Mahwah, N. J.: Lawrence Erlbaum.

Yin, R. K. (2003). *Case study research: design and methods* (3 ed.). Thousand Oaks, CA.