1-30-2015

Recommended Practices for Determining Rate of Improvement: A Review of State-Level Guidance

Caitlin S. Flinn
Indiana University of Pennsylvania

RECOMMENDED PRACTICES FOR DETERMINING RATE OF IMPROVEMENT: A REVIEW OF STATE-LEVEL GUIDANCE

A Dissertation

Submitted to the School of Graduate Studies and Research

in Partial Fulfillment of the

Requirements for the Degree

Doctor of Education

Caitlin S. Flinn

Indiana University of Pennsylvania

December 2014

Indiana University of Pennsylvania
School of Graduate Studies and Research
Department of Educational and School Psychology

We hereby approve the dissertation of

Caitlin S. Flinn

Candidate for the degree of Doctor of Education

October 28, 2014 — Signature on File
Joseph F. Kovaleski, D.Ed.
Professor of Educational and School Psychology,
Advisor

October 28, 2014 — Signature on File
Mark J. Staszkiewicz, Ed.D.
Professor of Educational and School Psychology

October 28, 2014 — Signature on File
Courtney L. McLaughlin, Ph.D.
Assistant Professor of Educational and School
Psychology

October 28, 2014 — Signature on File
David J. Lillenstein, D.Ed.
School Psychologist
Derry Township School District

ACCEPTED

Signature on File
Timothy P. Mack, Ph.D.
Dean
School of Graduate Studies and Research

Title: Recommended Practices for Determining Rate of Improvement: A Review of State-Level Guidance

Author: Caitlin S. Flinn

Dissertation Chair: Dr. Joseph F. Kovaleski

Dissertation Committee Members: Dr. Mark J. Staszkiewicz
Dr. Courtney L. McLaughlin
Dr. David J. Lillenstein

The reauthorization of the Individuals with Disabilities Education Act (IDEA; 2004) fundamentally permitted a response to intervention (RTI) process for identifying students with specific learning disabilities (SLD) after decades of problematic criteria with the ability-achievement discrepancy model. Monitoring student progress to determine rate of improvement (ROI) is an important component of RTI (Fuchs & Fuchs, 1998; Gresham, 2001; Kovaleski et al., 2013) and while abundant in the literature, recommended practices continue to surface that require inclusion in state regulation and guidance documents.

This exploratory, qualitative study built on previous studies (Hauerwas, Brown, & Smith, 2013; Zirkel & Thomas, 2010b) that systematically reviewed state regulation and guidance documents for determining student eligibility for SLD. Documents representing all 51 states (including the District of Columbia) were used to gather information about the inclusion of a set of recommended practices that were gleaned from the literature and formed into a checklist. Because a checklist of recommended practices for determining ROI did not already exist, inter-rater reliability was calculated between two researchers for a subset of documents and was found to be between fair and perfect for checklist items.

Evidence for recommended practices were inconsistent across states, with some being included more frequently than others. Monitoring student progress over time was found most often, but specific practices related to the quality of data used to determine

special education eligibility were scant (i.e., using linear regression to generate a trend line representing student progress, considering outlier data). Few states included specific procedures for requiring or recommended a specific number of data points for making instructional decisions, quantifying a magnitude of deficiency for determining when a student is significantly below expectations, or specifying procedures by SLD category or grade level.

Implications for future research include refining the research tool used to gather information and gathering updated information about state requirements and recommendations as states continue to refine regulation and guidance documents. Implications for educators and school psychologists include the need to remain current with ever-changing regulatory and guidance documents for identifying students with SLD and bring empirical research findings to practice for determining student ROI.

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

For several decades, school-based teams have been identifying students as having specific learning disabilities (SLDs) using problematic eligibility criteria. Research has highlighted a framework that can reduce the number of students referred for SLD evaluations by improving academic outcomes for all students (Burns, Appleton, & Stehouwer, 2005; Gibbons, 2008; VanDerHeyden, Witt, & Gilbertson, 2007). This framework includes the measurement of students' response to intervention (RTI) with brief, frequent academic assessments. Legislation has been passed to allow the use of student data gathered through an RTI process as a way of determining SLD eligibility (Individuals with Disabilities Education Act [IDEA], 2004), but research indicates that there are better ways than others (i.e., recommended practices) for measuring and analyzing student academic progress. In order to avoid allowing a new set of problematic criteria for SLD, states need to incorporate recommended practices from research into their guidance documents for schools regarding student progress data. This study will review state regulation and guidance documents that include language relevant to monitoring student progress to determine if recommendations align with important research findings.

Background

Specific Learning Disabilities

The Educate All Handicapped Children Act (EHA) of 1975 was the first legislation to allocate federal funds for students in need of special education (EHA, 1975). Regulations followed in 1977 and included the long-standing definition of SLD and other disability categories. The rationale behind the regulatory definition of SLD is that a student's

academic capability is directly related to intellectual ability, which practitioners interpreted as a measurement of IQ (Kovaleski, VanDerHeyden, & Shapiro, 2013). Therefore, if a student with an average ability performed significantly below average on a standardized assessment of reading, he or she may have an SLD. In other words, the requirement to qualify as having an SLD was a discrepancy between ability and achievement. In this ability-achievement discrepancy model of determining SLD eligibility, students are typically administered standardized assessments of intelligence and achievement, which can usually be accomplished in a couple of school days. This type of data is essentially a "snap shot" of the student's performance and does not usually include multiple data sources to confirm that there is a significant skill deficit that requires special education services. Further defining this ability-achievement discrepancy was largely left to individual states, which lead to an inconsistency of eligibility between states and even between schools (Bocian, Beebe, MacMillan, & Gresham, 1999; Ysseldyke, Gresham, & Bocian, 1983).

Other problems surfaced as schools continued to use the ability-achievement discrepancy approach to identify SLD. One problem, as revealed in meta-analyses, is that schools tend to identify students with SLD when there were low achievement scores, sometimes regardless of ability, as measured by IQ (Gresham & Vellutino, 2010; Stuebing, Barth, Molfese, Weiss, & Fletcher, 2009). A second problem is that researchers are divided as to the utility of measuring ability or cognitive processes for predicting a student's acquisition of basic skills (e.g., reading). Some researchers reported that the measurement of cognitive processes does not correlate with the remediation of student skill deficits (Vellutino, Scanlon, & Lyon, 2000). Others from the neuropsychology perspective agree

that the use of a general IQ score compared to achievement performance is not reliable or helpful, but report that cognitive processing data are necessary to meet the definition of SLD and individualize educational plans for students (Wright, Hale, Backenson, Eusebio, &Dixon, 2013). A third problem is that the ability-achievement discrepancy relies heavily on data that has to be created (i.e., with standardized assessments) versus RTI data, which are collected over time. Making the decision to classify a student with an educational disability should not be taken lightly, yet evaluations for SLD have often times been boiled down to the comparison of two scores: IQ and academic achievement (Bocian et al., 1999; Meyer, 2000; Restori, Katz, & Lee, 2009; Stuebing, Fletcher, LeDoux, Lyon, Shaywitz, & Shaywitz, 2002). A fourth issue is that it is difficult to demonstrate a severe discrepancy in the primary grades; therefore students are usually identified as SLD in the third or fourth grade after potentially missing several years of necessary special education services (Lyon, Fletcher, Fuchs, & Chhabra, 2006).

Because of the issues surrounding the ability-achievement discrepancy, researchers in education have proposed alternative methods of identifying SLD that focus on achievement levels and student achievement progress over time while receiving interventions in the regular education settings, dating back to the 1980s (Heller, Holtzman, Messick, 1982, p. 69). One alternative method that is prevalent in the literature is the dual-discrepancy approach for determining significant skill deficits (Fuchs & Fuchs, 1998). In this approach students are considered dually discrepant if they are significantly below level and not responsive to intensive instruction. For example, a fifth grade student who is reading on a second grade level and is also not demonstrating progress to close the gap

between his performance and that of his typically performing peers could be considered dually-discrepant.

The dual-discrepancy approach requires data to be collected over time to determine if the student is responding to instruction and intervention. Unlike the ability-achievement discrepancy that compares a student's achievement to that same student's IQ, the dual-discrepancy approach compares a student's level and rate of improvement (ROI) to his or her typically performing peers (Fuchs, 2003). This approach relies on a set of basic academic skill measures that can be administered more frequently than the traditional standardized assessments of achievement. One type of measure that can be collected over time and provide information about a student's level and ROI is called curriculum-based measurement (CBM).

Curriculum-Based Measurement

While working at the University of Minnesota, Stanley Deno and colleagues were charged with developing a set of brief and inexpensive assessment procedures for measuring basic skills that could inform teachers about their students' progress. In order to report on special education goals and to rely less on lengthier assessments of achievement, CBM was designed to help teachers know when and how to intervene with a change of instruction more effectively and efficiently (Deno, 1985). Probes were originally developed from a teacher's curriculum (e.g., students' reading material) to assess the basic skills of reading, writing, and mathematics, and were aptly named CBM. A common example of CBM is having a student read aloud from a grade-level text for one minute and counting the number of words read correctly and incorrectly. Math CBM might consist of a student completing single-digit by single-digit addition problems for a minute and writing

could be a three-minute timed assessment of the student's total number of words produced.

One of the easiest ways to view data collected over time is to plot data points (student scores) on a chart that has an x- and y-axis. The x-axis could be the number of school weeks and the y-axis would be the score, such as the number of words read correctly per minute. The teacher could graph the student's data and quickly determine if the student is progressing toward a pre-set goal. Although the teacher or student could graph data by hand on standard graph paper, technology allows for more efficiency and can provide additional information for consideration. Microsoft Excel offers a chart feature that not only graphs student data but can create a trend line, or the average skill acquisition over time (Kovaleski & Flinn, 2011; Flinn & McCrea, 2012; Kovaleski et al., 2013; VanDerHeyden & Burns, 2010).

There are benefits to analyzing data across time. One example is measuring a person's weight. The number on the scale is accepted as a general indicator of a person's health when other factors are considered such as height, sex, or family medical history. However, as anyone who has tried to lose weight has found, the number on the scale can vary considerably depending on many other factors such as time of day the person was weighed, their recent eating and exercising habits, their water intake, and their level of stress. To determine if people are making progress toward their goal weight, they might weigh themselves once a week. Based on those data, the person will be able to decide if they need to make additional changes to better ensure attainment of their goal weight. Similarly, by reviewing student achievement across time, the inevitable "bad days" will factor less into the trend of overall progress and more information will be available for

when school teams need to make high-stakes decisions such as special education eligibility. This process of intervening and measuring a student's academic growth over time can be described as measuring and individual student's response to instruction and intervention.

Standardized achievement assessments typically used for state reporting or by school psychologists are time-consuming and are generally not sensitive to small increments of growth, but they can provide information about a student's level of performance. Data from CBMs indicate a student's level, but can also be used frequently enough to illustrate a student's ROI. Because CBM is sensitive to change over time and can accurate determine a student's level, it is most often the assessment of choice for accurately determining dual discrepancy which requires consideration of both level and rate of performance (Elliott & Fuchs, 1997; Marston & Magnusson, 1985; McMaster, Fuchs, Fuchs, & Compton, 2005) but is not the only option. For instance, with advancements in technology, assessments have been developed that can be taken by students on a computer, such as STAR Reading (Renaissance Learning, 2013).

Over time, the technical adequacy of CBM has improved; CBM measures what it intends to measure (validity) and repeated administrations of CBM consistently yields similar results (reliability; Fuchs, Fuchs, & Speece, 2002; Wayman, Wallace, Wiley, Tichá, & Espin, 2007). Following the initial research project at the University of Minnesota, CBMs have been refined into commercially available products that include national norms and standard rates of academic growth. CBM tools continue to be well represented in research for the fields of education and school psychology for making instructional decisions based on student performance (Fuchs, 2004).

In addition to the original purposes of CBM set forth by Deno (1985) paired with enhanced technical adequacy, the functional utility of CBM has evolved. Educators can analyze student achievement with CBM in numerous ways including determining the level at which a student is performing, the rate at which a student is progressing, and what the student's data means in comparison to other student data. Using CBM has been validated in the research literature for these purposes; however, CBM can also be used to help with high stakes decision such as special education eligibility (Kovaleski et al., 2013). Recent research has reviewed the technical adequacy of using CBM data for high-stakes decisions and has not found empirical support for such practices (Ardoin et al., 2013; Thornblad & Christ, 2014). Because of the popularity of CBM for making instructional decisions (Ardoin et al., 2013; Fuchs, 2004) it is important to have the most sound data possible as psychometric properties of CBM continue to be refined.

Theoretical Framework

The collection of individual student data, graphed over time, is only one definition of RTI. The other definition falls within the broader context of a multi-tiered system of supports (MTSS; Batsche et al., 2005; Brown-Chidsey & Steege, 2010; Kovaleski et al., 2013) for which the focus is on the improvement of academic and behavioral outcomes for all students (VanDerHeyden & Burns, 2010) rather than on individual student RTI. Because of the long-standing history, ease of use, and multiple utilities, CBM is often used to help educators make instructional decisions for students who meet or do not meet grade-level expectations (Fuchs, 2004). Most MTSS models utilize an RTI framework in which students can move between tiers. For instance, a student who meets the grade-level expectations (i.e., benchmark) would continue to receive Tier 1 supports, whereas a

student who does not meet the benchmark may be provided with Tier 2 supports, which are in addition to the instruction provided at Tier 1. Students who do not make progress in Tier 2 may benefit from more strategic instruction at Tier 3. Data from CBMs can be used to determine which students are at benchmark and which students require strategic or intensive instruction in order to make progress in a district's curriculum. Student progress can be monitored more frequently for students who require more intensive instruction (e.g., one CBM probe per week) to ensure that interventions are working.

As educators continue to use CBM as a means for measuring student growth, policies have been set in place in various states (Zirkel & Thomas, 2010b) without necessarily having sufficient research supporting practice (Ardoin et al., 2013; Christ, Zopluoglu, Long, & Monaghen, 2012; Thornblad & Christ, 2014). For instance, Oklahoma currently requires four data points collected in four weeks before moving a student from Tier 1 to Tier 2 and nine data points collected in nine weeks to move from Tier 2 to Tier 3. A current research study by Christ et al. (2012) suggests that many more data points are required to establish a stable trend line than the four or nine data points recommended by Oklahoma state guidance documents. Christ et al. (2012) studied oral reading fluency progress monitoring data from three different commercial providers of CBM (e.g., AIMSweb) to determine the minimum number of data points required to produce a statistically stable trend line. For some measures, the minimum number of required data points was equal to one oral reading fluency probe administration per week for 18 weeks. This is an example of policy driving practice without having sufficient research to support either policy or practice.

More recently, the quality of trend lines created from CBM have come into question (Christ et al., 2012), particularly as policy-makers and practitioners consider RTI data to make high stakes decisions, such as special education eligibility. While there are no current studies that compare various methods for calculating the trend of CBM data over time and the resulting quality of educational outcomes for students, a review of CBM literature reveals that the majority of studies employ an ordinary least squares method of simple linear regression to determine a student's ROI (Ardoin, Christ, Morena, Cormier, & Klingbeil, 2013; Christ et al., 2012).

Statement of the Problem

In response to decades of documented problems associated with the ability-achievement discrepancy model of identifying SLD (Bocian et al., 1999; Gresham & Vellutino, 2010; Vaughn & Klingner, 2007; Vellutino et al., 2000), the most recent reauthorization of IDEA (2004) does not allow states to mandate the use of an ability-achievement discrepancy. Instead, the IDEA (2004) regulations (§300.309) support the use of an RTI process when considering SLD for any student. A component of an RTI process is the monitoring of student progress, which is quantifying and describing "data-based documentation of repeated assessments of achievement at reasonable intervals, reflecting formal assessment of student progress during instruction" (IDEA, 2004, § 300.309(b)(2)). In order to avoid the psychometric problems of the past, it is important to investigate how states have incorporated this information that recommends monitoring of student progress into their state policies or guidelines. Furthermore, it is important to determine if state policies or procedures are aligned with recommended practices in the literature regarding the monitoring of student progress.

Following the reauthorization of IDEA (2004), several studies have reviewed state level changes related to the implementation of RTI as a framework and using an RTI process for individual students when determining eligibility for SLD. Some studies included surveys of state directors of special education to determine if states have simply adopted the federal regulations or if they have further defined criteria for SLD (Ahearn, 2008; Hoover, Baca, Wexler-Love, & Saenz, 2008; Zirkel & Krohn, 2008). Other studies focused directly on reviewing state documents for evidence of change (Hauerwas, Brown, & Scott, 2013; Zirkel, & Thomas 2010a, Zirkel, & Thomas, 2010b). The study by Hauerwas and colleagues provided one of the first studies of state documents that explicitly set out to identify best practices related to how states are defining responsiveness. Their study used a qualitative content analysis approach to describe themes in state guidance documents (i.e., policies or recommendations) and determine if a consistent use of an RTI process for SLD has emerged. The study concluded that consensus has not been reached.

None of the studies reviewed included a comparison of requirements or recommendations found in state guidance documents to a set of recommended practices that have surfaced in research. This study intends to fill the gap in the literature of recommended practices and state guidance regarding student ROI. Because policy has been put in place via IDEA for progress monitoring, at least when considering students for special education eligibility for SLD, the purpose of this study would be to explore and report evidence of requirements and recommendations aligned with the current literature related to measuring student achievement over time (i.e., ROI) in state regulations or guidance documents.

Research Questions and Hypotheses

Research Question 1

Do state regulation and guidance documents include recommended practices for monitoring student progress? This question will be answered through exploratory research for which results can be used to further define the problem and/or develop hypotheses for future research.

Research Question 2

What are the differences, if any, for including recommended practices of monitoring student progress for states that mandate RTI versus states that do not mandate an RTI process for identifying SLD eligibility? This question will be answered through exploratory research for which results can be used to further define the problem and/or develop hypotheses for future research.

Research Question 3

What are states requiring or recommending in terms of the number of data points needed to make instructional changes? Hixson et al. (2008) reported that instructional change can be made confidently with very few data points. Zirkel and Thomas (2010b) reported various suggestions or requirements for the number of data points used for instructional decision-making. Hauerwas et al. (2013) had reported that state guidance documents most frequently suggested three to four data points be used before making instructional change. This question will be answered through exploratory research for which results can be used to further define the problem and/or develop hypotheses for future research.

Research Question 4

Do states have specific, quantifiable requirements or recommendations for defining the magnitude of deficiency in level when a student is performing significantly below specified expectations? This question will be answered through exploratory research for which results can be used to further define the problem and/or develop hypotheses for future research.

Research Question 5

Do states specify requirements for identifying students with SLD for any or all of the eight categories of SLD? This question will be answered through exploratory research for which results can be used to further define the problem and/or develop hypotheses for future research.

Research Question 6

Do states specify requirements for identifying students with SLD for specific grade levels? This question will be answered through exploratory research for which results can be used to further define the problem and/or develop hypotheses for future research.

Significance of the Study

In the school setting, SLD is the most prevalent disability category (Technical Assistance Coordination Center, 2012; Zirkel, 2013); therefore, studies related to SLD can positively impact a large population of students with special needs. The ability-achievement discrepancy model for identifying SLDs has largely been contested as problematic and even harmful for students because criteria can vary between schools and states and because students may not qualify for special education for several years despite having academic difficulties (Reschly & Hosp, 2004). Noting that the way that students

were identified prior to the reauthorization of IDEA (2004) was problematic and inconsistent, it is important that the adoption of new eligibility criteria using an RTI process results in better outcomes for students and is based on research. Because monitoring of student progress is an important component of an RTI framework, and because repeated achievement assessments is written into the 2006 IDEA regulations, the focus of this study will be a review of whether states have written requirements or suggestions that align with recommended practices found in research for monitoring student progress. This study is also significant because it will contribute to a growing body of research that has reviewed changes in state documents as a result of the IDEA (2004) reauthorization. Reviewing state guidance documents for recommended practices will highlight where states need to improve support for schools and where training institutions need to update pre-service programs for educators.

Assumptions

There are a few assumptions inherent in this study. The first assumption is that the state documents reviewed for this study are accurate and in the most current version. A second assumption is that the research tool developed for this study is technically adequate for reporting results. A third assumption is that practitioners in schools follow state guidelines and regulations.

Limitations

There are limitations inherent in conducting qualitative research such as the role and biases of the researcher completing the study. Previous research that has included reviews of state documents indicated that said documents seem to be fluid and ever changing (Hauerwas et al., 2013; Zirkel & Thomas, 2010b). Therefore, documents may be

in the process of being changed and would not be captured in this study. However, this study will still be able to document the progress of a few years since the documents that were reviewed in previous research. Because there is no one method for interpreting data and making instructional decisions from CBM results (VanDerHeyden & Burns, 2010), this study will use a set of recommended practices developed by the researcher and reviewed by a panel of experts. Once the set of recommended practices was reviewed and revised as needed, the researcher used the list of recommended practices as a checklist when reviewing state documents; therefore, some judgment is involved on behalf of the researcher and there is potential for researcher bias.

Delimitations

Delimitations are frequently reported in qualitative research to describe what the study will not include. Informational text reviewed in this study did not include training materials set forth by the SEA since they may be interpretations of policy or guidance documents. The focus of this study is to review macro-level changes in state requirements or recommendations for determining student ROI; therefore, this study will not review what individual local education agencies (LEAs) within states have required or recommended for determining student ROI.

Definition of Terms

Curriculum-Based Measurement

CBM is a brief measure of basic academic skills such as reading, writing, or math that serves as a general indicator of achievement (Deno, 1985). The results of CBM can help educators determine when to intervene with an instructional change (Fuchs, 2004). When given approximately three times per school year to all students in a grade level or

school building, CBM can serve as a universal screening tool. When given monthly, every other week, weekly, or more frequently, CBM can help educators determine if progress is being made in the respective skill area being measured.

Guidance Documents

For the purpose of this study, guidance documents are documents that are considered to be a set of recommendations, as opposed to requirements, that are set forth by the state education agency. If evidence of a recommended practice is found in a guidance document, it is considered to be a suggestion rather than a requirement.

Local Education Agencies

An LEA is the term used to describe the decision-making authority of a school district (U.S. Department of Education, n.d.).

Progress Monitoring

In the fields of education and school psychology, monitoring student academic progress is usually completed with CBM (Shinn, 2008) or other technically adequate measures that can be repeatedly administered, are sensitive to small increments of growth, and are observed over time to determine the effectiveness of instruction and intervention (Hixson, Christ, & Bradley-Johnson, 2008).

Rate of Improvement

ROI is a term used to quantify the results of frequent monitoring of a student's progress over time (Deno, Fuchs, Marston, & Shin, 2001; Flinn, 2011; Kovaleski et al., 2013). For instance, if a student is participating in a reading intervention, the teacher may administer a curriculum-based measurement (CBM) for reading each week. The teacher can then calculate how many words the student is gaining per week on average. ROI can

also be a statistic calculated from the line of best fit, or trend line, of student data points that can be compared to the progress of typical students, or other students performing at the same grade level (Fien et al., 2010; Silberglitt & Hintze, 2007).

Regulations

For the purpose of this study, regulations are defined as documents that are legally-binding. Recommended practices found within regulation documents are considered to be required as opposed to recommended since they are contained within a legally binding document.

Response to Intervention

RTI on a broader scale falls within the definition of a MTSS to improve outcomes for all students both academically and behaviorally (Averill & Rinaldi, 2011; Buffman, Mattos, & Weber, 2009; Kovaleski et al., 2013). Individual student RTI involves data that demonstrate the level and rate of student achievement given multiple tiers of instruction and intervention.

Specific Learning Disability

The definition for an SLD has remained consistent for several decades and is written as follows in the IDEA (2004) legislation:

> Specific learning disability means a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations, including conditions such as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. (34 CFR 300.8(c)(10))

Standardized Assessments

When administered by school psychologists, standardized assessments can refer to individually administered, norm-referenced tests of intelligence or academic achievement. In general, a standardized assessment is any kind of assessment that has a specific set of procedures or protocols that are administered to a student or groups of students in a uniform manner to allow for comparison.

State Education Agencies

State education agencies (SEAs) are the people with decision-making power within the state such as state departments of education or state board of education (US Legal, 2013). For the purpose of this study, the term state and SEA will be used interchangeably.

Trend Line

In statistics, a trend line is the line of best fit for a set of data points that allows one to determine the average increase or decrease in said data set (Gall, Gall, & Borg, 2007). In education, the trend line can represent average performance over time on an academic skill as measured by CBM or other measures (Hixson et al., 2008).

Universal Screening

In order to determine which students may require changes to the instructional environment to improve their academic success and to estimate students' performance on state assessments, schools can administer short assessments (such as CBM) three to four times per year to all students (National Center on Response to Intervention [NCRTI], n.d.).

Summary

This chapter reviewed the issues surrounding the ability-achievement discrepancy model for SLD identification and established a need for updated eligibility criteria that

provides meaningful information for implementing interventions. The dual-discrepancy method was reviewed, including the need for monitoring student progress, as a method for determining a student's response to instruction and intervention. The benefits of using CBM to monitor student progress were discussed especially in relation to requirements from the reauthorization of IDEA for repeated measures and the allowance for states to use an RTI process for SLD eligibility. The need for states to develop guidance documents congruent with recommended practices for monitoring and interpreting student progress was discussed. In addition, the need to gather updated information about state regulation and guidance documents and the need to determine whether state policy or recommendations align with recommended practices as found in the research was outlined as the focus of this study. Assumptions, limitations, delimitations, and a description of important terms were provided.

CHAPTER II

REVIEW OF THE RELATED LITERATURE

This chapter will review relevant literature describing events that have led to macro-level changes in education and the implementation of response to intervention (RTI) as a framework for improving outcomes for all students. The use of an RTI process for identifying students with specific learning disabilities (SLDs) is discussed including research results that illustrate problems with the ability-achievement discrepancy model. Because an RTI process for special education eligibility includes the collection of student response data, several key features of monitoring student progress are described, especially research-based suggestions for determining student rate of improvement (ROI). Studies pertaining to the benefits and cautions of determining student ROI are discussed and summarized into a set of recommended practices. Finally, an outline is provided of former studies that have captured changes over time with regards to the implementation of RTI and will be used to support the methodology proposed for the current study.

School Reform

There have been several events that have occurred since the year 2000 that have had a major impact on education. Information has come forth regarding key reading and math skills for students and schools have been set to task to improve accountability for teaching and student learning. High stakes student assessments have become a popular and required method of measuring school accountability and schools have a renewed need to determine student progress. Updated legislation for students with disabilities has also been passed and has highlighted components of promising models of assessment,

instruction, and intervention to meet students' learning needs. A review of literature reveals that there is still much work to be done to improve outcomes for all students.

The National Reading Panel disseminated one of the first momentous reports regarding student learning in 2000. This report was the result of several years of the panel members gathering evidence of scientific support for the skills and instructional methods students need to become effective readers. The report concluded with a recommended five big ideas in reading, which consists of phonemic awareness, phonics, fluency, vocabulary, and comprehension. A National Mathematics Advisory Panel (2008) was later formed to outline similar information for the instruction and sequence of mathematical skills.

Soon after the National Reading Panel Report, the president signed into action the No Child Left Behind Act of 2001 (NCLB; enacted in 2002). The purpose behind this act was to raise standards of learning and require accountability via annual yearly progress (AYP) on state assessments. The act also required the use of scientifically based instructional strategies to improve student achievement.

In line with the NCLB requirement for scientifically based instruction, the reauthorization of Individuals with Disabilities Education Act (IDEA, 2004) maintained a requirement set forth in 1997 that for students evaluated for special education, the school teams had to demonstrate appropriate instruction in reading and math. In other words, schools could not identify students with educational disabilities if there was no documentation of evidence-based reading and math instruction. When IDEA enacted the regulations in 2006, schools had been hearing for several years that education in the general education setting needed to be improved. A promising framework (i.e., RTI) was emerging in the literature as a way of attaining the goal of improved student outcomes.

Although RTI models were and continue to be a regular education initiative, the IDEA demonstrated confidence in using said model for individual students as a method for determining eligibility for SLD. In order for schools to use this new option of documenting a student's response to instruction and intervention, schools needed to consider screening, intervention, and monitoring of student progress at a larger scale. With special education law allowing for individual student RTI to be used for evaluations, coinciding efforts were developing for large scale RTI implementation.

Fuchs and Fuchs (2001) described a blueprint for putting into practice an RTI framework. Specific components of a multi-tiered framework were outlined as well as recommended responsibilities for both general education and special education. Tier 1of this framework was described as being the high-quality research-based general education curriculum used for educating all students. Also within Tier 1, general education would complete universal screenings to determine which students required additional, more focused instruction beyond the Tier 1core curriculum. Students performing below a specific score could receive additional instruction in the general education setting (i.e., Tier 2). While receiving instruction at the Tier 2 level, student progress would be monitored more frequently than the universal screenings. This monthly or bi-monthly (data gathered every other week) would help educators determine if students needed additional learning opportunities at the Tier 3 level, should remain at the Tier 2 level, or could fade the intervention provided at the Tier 2 level. Fuchs and Fuchs (2001) proposed that general education and special education work together to instruct students at the Tier 2 and Tier 3 level, including parents in the decision-making process. If students do not make adequate progress in this system, it was recommended that they be evaluated (with parent consent)

for a possible learning disability, following the required procedures and considerations of special education evaluations.

The National Association of State Directors of Special Education (NASDSE) drafted a significant document regarding the implementation of RTI models that further defined the components and benefits (Batsche et al., 2005). Components included screening for all students, interventions for students not meeting benchmarks (i.e., Tier 2), the use of standard protocol interventions (interventions that have been shown to be effective for a large number of students with similar skill deficits), the monitoring of student progress, intensive interventions (i.e., Tier 3) when students have not progressed, and possible referral for an evaluation if necessary. Emphasis was placed on using a problem-solving model of defining the problem, gathering data to determine the cause of the problem, intervening, and determining effectiveness of the intervention. The importance of determining a student's level of performance and rate of learning were described.

Following the reauthorization of IDEA (2004), the Office of Special Education Programs (OSEP) was asked to clarify a number of aspects related to the change in evaluation procedures for students suspected as having an SLD. In 2007, OSEP created a questions and answers document that described the Department of Education's understanding of common components for RTI frameworks including:

> (1) students receive high quality research-based instruction in their general education setting; (2) continuous monitoring of student performance; (3) all students are screened for academic and behavioral problems; and (4) multiple levels (tiers) of instruction that are progressively more intense, based on the student's response to instruction. (p. 14)

Other literature has outlined notable components of RTI frameworks that corroborate the information provided by OSEP (2007; Glover & DiPerna, 2007; Kovaleski, 2007) including studies that have captured changes in state policies and recommendations for the use of an RTI process for evaluating SLD (Hauerwas, Brown, & Scott, 2013; Zirkel & Thomas, 2010b).

In recent years, many initiatives have taken place to improve student achievement in the United States. Information about best practices for teaching reading and mathematics paired with new accountability systems for educators to instruct students to proficiency on common standards have led to a movement of school reform. Promising educational frameworks have been noted from influential organizations (i.e., NASDSE) and regulations regarding the evaluation of students with for SLD (i.e., IDEA) has allowed for a new method of identification that could serve as a framework for improving the learning outcomes of all students. Although RTI has been proposed as a promising method of SLD identification, research needs to be conducted and reviewed for large scale RTI beyond the evaluation of individual students.

Student Outcomes From Large Scale RTI

As RTI became a common term in education, studies surfaced evaluating the impact of implementation of RTI models on a large scale. In several large-scale evaluations of RTI, benefits were described as reduced numbers of students referred for special education, shrinking special education rates, and improved achievement for the student population.

A pertinent meta-analysis (Burns, Appleton, & Stehouwer, 2005) compared outcomes for sites that had been implementing an RTI framework for several years to sites that received university funds to implement RTI. Outcomes were measured by a reduced

number of students found eligible for special education. The researchers had hypothesized that sites receiving training from the university would have better outcomes, assuming that specific training and fidelity of implementation would surpass a district-run initiative. Although the university-supported sites and the district-run sites both experienced positive results, the researchers concluded that the district-run sites experienced better outcomes, possibly because of better buy-in from faculty since implementation of RTI was over a longer period of time.

Another large scale RTI implementation effort has been taking place through the Screening to Enhance Educational Performance project (STEEP; VanDerHeyden, Witt, & Gilbertson, 2007). Research results in their 2007 study found that implementation of RTI led to a reduction in the number of students requiring special education. Similarly, the St. Croix River Education District in Minnesota has adopted common components of RTI on large scale, including research-based instruction, continuous monitoring of student progress, and evaluation of intervention effectiveness by determining level and rate of student academic growth (Gibbons, 2008). After years of implementation, a significant increase in student achievement occurred as measured by district screening data and state performance assessments.

The Reading First program was an initiative created in response to the NCLB Act (2001) to improve student achievement in reading in the primary grades through the use of a multi-tiered framework (Moss, Jacob, Boulay, Horst, & Poulos, 2006). Although NCLB and Reading First did not specifically mention the use of an RTI framework, the screening and intervention process of the program is similar to the multiple tiers described for RTI models. A study by Torgesen (2009) reviewed the benefits of using an RTI framework

through the Reading First program and found a decrease in the number of students referred for evaluations for learning disabilities and number of students requiring intensive instruction.

Although specific initiatives and research studies have demonstrated benefits of scaling up an RTI framework, continued research is necessary to provide professional development and refine the process of determining eligibility for SLD (Griffiths, Parson, Burns, VanDerHeyden, & Tilly, 2007; Kratochwill, Volpiansky, Clements, & Ball, 2007).

Individual Student RTI

Problems with the Ability-Achievement Discrepancy Model

The ability-achievement discrepancy model can be interpreted to be an unexpected low performance on a standardized assessment of achievement compared to an average performance on IQ measures. This method of determining the presence of a severe discrepancy is heavily dependent on results from standardized assessments and does not necessarily consider information from other sources such as performance on classroom assessments or rate of skill improvement measured over time. Despite the apparent black-and-white formula of ability-achievement discrepancy, it appears that a plethora of studies exist confirming a very gray set of criteria (Bocian, Beebe, MacMillan, & Gresham, 1999; Meyer, 2000; Restori, Katz, & Lee, 2009; Stuebing, Fletcher, LeDoux, Lyon, Shaywitz, & Shaywitz, 2002).

Soon after the regulations from the EHA (1975) were finalized, studies questioning the use of the ability-achievement discrepancy model for SLD identification came forth. The quantification of severe discrepancy between ability and achievement could take many forms and the lack of consistency with which students were identified was illustrated in a

study by Ysseldyke, Algozzine, and Epps, (1983). Essentially, two studies were completed with descriptions of 17 possible interpretations of severe discrepancy. The first study reviewed data from regular education (i.e., non-identified) students spanning across grade levels. When the student data were run through the 17 possible variations of ability-achievement discrepancy criteria, between 2 and 65% of students met one or more of the 17 criteria for learning disabled. Results from a second group of fourth grade students, half of whom were regular education and low achieving and half of whom were identified as learning disabled, indicated that between 1 and 78% of students met one or more of the 17 possible criteria for learning disabled. This study suggested that more specific criteria are necessary for the identification of students with SLD. A similar study was completed where data from 150 students in California schools who were previously referred for SLD were reviewed. According the researchers, approximately half of the students who were found eligible for services did not meet the state criteria (MacMillan, Gresham, & Bocian, 1998).

Because of the inconsistency of interpretations for determining SLD, Bocian et al. (1999) reviewed data from a sample of 76 students in an attempt to uncover deciding factors for whether students were found eligible or not. Information from classroom-based reading assessments, individualized assessments of ability and achievement, and the eligibility determination was reviewed. Overall, they found agreement between classroom-based assessment results and special education placement and between a severe ability-achievement discrepancy and special education placement, but there was little agreement between the classroom-based assessment levels and ability-achievement discrepancy results. The study also found that students who were found eligible who did not demonstrate a severe ability-achievement discrepancy were male students with behaviors

that teachers found to be problematic. Results indicated that there are instances where other variables are considered beyond the ability-achievement discrepancy, which can lead to inconsistent placement decisions.

Other studies have illustrated that students are found eligible as having an SLD regardless of whether students met severe ability-achievement discrepancy criteria. Stuebing, Barth, Molfese, Weiss, and Fletcher (2009) completed a meta-analysis of studies to determine the relation between IQ tests and reading achievement. After reviewing over 1000 studies, the researchers included 22 studies that met their specific criteria. Studies needed to have a well-defined intervention for reading achievement with reading achievement results reported for both before and after the intervention. The studies needed to report enough data for researchers to determine the effect of IQ on the reading achievement scores. If available, data from other variables were recorded including age, types of IQ tests, and types of reading skills (i.e., phonological awareness, word reading analysis, reading fluency, and reading comprehension). The essential question of the study was to determine if IQ scores predict student response to reading interventions. Results concluded that IQ results are poor predictors of student response and account for only one to three percent of the variance (Stuebing et al., 2009).

To address the questionable relationship between IQ and achievement, Gresham and Vellutino (2010) reviewed results of several meta-analyses and two studies that directly measured student response to reading intervention. The authors concluded that, "(1) IQ is not highly correlated with reading achievement, (2) IQ does not predict response to intervention, and (3) measures of response to intervention are better predictors of long-term risk status than IQ" (p. 194). Another review of existing research included

conclusions that the ability-achievement discrepancy does not help educators determine which students have a reading disability and which students are poor readers (Vellutino, Scanlon, & Lyon, 2000). The article also concluded that IQ does it clearly discriminate between typical and poor reading performance.

Given the evidence that general IQ has very little utility for helping educators make decisions about student need for special education, is there a need for completing cognitive assessments at all? Those from the neuropsychological standpoint have argued that measures of cognitive processing are required to meet the definition of SLD, but agree that a general IQ score itself is purposeless for SLD evaluations (Wright, Hale, Backenson, Eusebio, & Dixon, 2013).

In light of the problems with the ability-achievement discrepancy model, a Learning Disabilities Summit was held in 2001 to discuss the need for a change in eligibility criteria. A variety of experts in the field were consulted and other options for deter mine SLD eligibility were proposed, including the use of an RTI process (Bradley, Danielson, & Hallahan, 2002).

RTI as a Process for Identifying SLD

A landmark study by Fuchs and Fuchs (1998) described the use of a treatment-validity model of identifying students with SLD. Within this model, students received high quality instruction and were screened for academic difficulties using curriculum-based measurement (CBM). Students who scored below benchmark received interventions in their classroom and their progress was monitored over time. This study referred to the dual-discrepancy approach for determining eligibility and proposed that students who were dually-discrepancy in level (e.g., grade level performance) and rate of progress (i.e.,

ROI) at or greater than one standard deviation from the mean of typically performing peers should be considered for special education evaluation.

Gresham (2001) reviewed the weaknesses of the ability-achievement model while proposing that RTI is viable alternative for identifying learning disabilities. In his study, three models of RTI were examined: (a) predictor-criterion, (b) dual discrepancy, and (c) applied behavior analytic. The predictor-criterion model employed potent interventions to improve reading achievement and determined effectiveness. The dual-discrepancy model involved instructional, environmental, and supplemental intervention variables to determine if a student responds. The study determined this model to be reliable and valid in terms of revealing students in need of and eligible for specially designed instruction through special education. The applied behavior analytic model entailed a behaviorist framework, examining antecedents, behavior, and consequences in terms of educational performance. All three models involved multiple tiers of intervention, progress monitoring, and effective instructional strategies. Results yielded support for applying an RTI model to identifying students for learning disabilities.

In support of the dual-discrepancy approach, Speece, Case, and Malloy (2003) found that students who were dually discrepant had more significant academic and behavioral needs than student who met criteria for an ability-achievement discrepancy. Another study examined variations of the dual-discrepancy model (Burns & Sensenac, 2005) by initial performance (percentile rank versus one standard deviation from the mean) and ROI (with benchmark data, not weekly slope). The results provided supportive evidence that the dual-discrepancy model using percentile ranks as initial indicators of level can be a valuable method for determining which students require intensive remediation in reading.

Other definitions of dual discrepancy were analyzed in the Burns, Scholin, Kosciolek, and Livingston study (2010). Two types of decision-making models with student RTI data were examined. The first type of analysis included the use of an aim line and end-of-year goal and the second type was a dual-discrepancy model of low rate of progress and below level performance. Results indicated that instructional decisions would differ depending on the approach used (aim line versus dual discrepancy) which is problematic when educators need reliable criteria for determining instructional needs for students. Specifically, the dual-discrepancy approach was less affected by standard error of measurement than the aim line approach. The researchers sought to determine whether the two methods would yield similar numbers of students requiring more intensive intervention but found that 40% of the students would have been determined to need more intensive interventions in one model (aim line) than in the other (dual-discrepancy).

Overall, an RTI process for identifying students with SLD has indicated the need to monitor student progress over time. Monitoring student progress has also been described as one of the key features of an RTI framework (Stecker, Fuchs, & Fuchs, 2008).

Key Features of Monitoring Student Progress

A review of literature surrounding the monitoring of student progress revealed several key features including the types of assessments used, the popularity of commercial assessment products, how to define progress and a lack of progress, creating goals for student growth, and the frequency for which data should be collected. Other features include the various methods for determining ROI, the use of technology as a tool for graphing and calculating ROI, and the number of data points needed to make sound instructional decisions.

Types of Assessments Used for Monitoring Progress

In the Fuchs (2004) commentary to the field of school psychology, the historical and current use of CBM was discussed and how it differs from mastery measures which are short-term measures for more specific skills. CBM has been described as being the most prevalent in educational and school psychology research (Shinn, 2007) for having strong technical adequacy for determining level and rate of progress (Fuchs, 2004).

Mellard, McKnight, and Woods (2009) reviewed various screening and progress monitoring practices from local schools that had recently started to develop an RTI framework. The researchers outlined and defined various types of assessments that could be used for progress monitoring (reading only) including published reading program assessments, CBM, and district assessments. Computer adaptive tests (CATs) are also becoming prevalent in the literature and come with unique benefits and disadvantages (Stone & Davey, 2011). An example of a CAT is the STAR Reading system by Renaissance Learning (2013). Students complete approximately 36 multiple-choice questions that are pulled from a large pool of possible questions. These assessments are based on item response theory (IRT) that assumes the knowledge of prerequisite skills when an item is answered correctly. When students achieve a correct answer, the system will provide a more challenging question and vice versa to provide information about a student's performance on a continuum ranging from early elementary grades to post-high school performance.

Commercial Products for Monitoring Progress

Of the various types of progress monitoring tools that could be used for progress monitoring, one of most popular is in the form of commercially available CBM.

The National Center on Response to Intervention (NCRTI, 2011) has reviewed numerous CBM products to determine levels of technical adequacy for general outcome measures (GOMs), which appraise the basic academic skills all students should acquire such as reading, writing, and mathematics. The forms of technical adequacy reported on the NCRTI web site (www.rti4success.org) include reliability of the performance level score, reliability of the slope, validity of the performance level score, predictive validity of the slope of improvement, alternate forms, sensitive to student improvement, end-of-year benchmarks, ROI specified, norms disaggregated for diverse populations, and disaggregated reliability and validity. As reported on the NCRTI web site, this type of measurement can be used to determine a trend in student growth in the form of a slope of progress or ROI. In addition to CBM, CATs are also being used to monitor student progress (e.g., STAR Reading) and have varied technical adequacies (NCRTI, 2011).

Defining Progress and Lack of Progress

The National Joint Committee on Learning Disabilities (2005) reported that there is no one way for quantifying and defining progress or RTI. Hixson, Christ, and Bradley-Johnson (2008) refer to multiple ways of interpreting student academic growth data including ordinary least squares regression and visual analysis of data graphed over time. In their reviews of state policies and recommendations, neither Zirkel and Thomas (2010 b) or Hauerwas et al. (2013) found consistent definitions of lack of responsiveness across states. It is important to understand how, if at all, states are defining lack of progress because it may contribute to high stakes decisions (i.e., special education eligibility).

Designating Goals for Student Progress

When Deno and colleagues (1985) proposed the use of CBM to track student progress, the intention was to help educators evaluate whether students were on target for reaching a preset goal. Educators often ask how much growth can be expected from students who appear to be several grade levels behind in basic academic skills. The purpose behind the Fuchs, Fuchs, Hamlett, Walz, and Germann (1993) study was to answer that question and measured oral reading fluency and digits correct for a sample of students. Results indicated that most students made an average gain of one word or digit correct per minute per week and concluded that goals could be set as realistic or ambitious by multiplying that ROI by two or three when intervention was provided. Shapiro (2008) built on this work and suggested a model of goal-setting based on student ROI.

Frequency of Monitoring Student Progress

One component of an RTI framework is ongoing progress monitoring (VanDerHeyden & Burns, 2010), but the key is defining "ongoing." Brief, reliable, and valid assessments of basic skills are administered at specific intervals in order to determine a student's ROI. As the intensity of interventions increase, so should the frequency of progress monitoring. The American Institutes for Research (2011) recommended that universal screenings (i.e., Tier 1) be completed three times per year: once in the fall, winter, and spring. Students receiving supplemental interventions (i.e., Tier 2) should be monitored at least monthly and students receiving intensive services (i.e., Tier 3) should be monitored at least weekly to determine if progress is being made.

Methods for Determining ROI

There are a number of methods to use to determine a student's slope or ROI to determine academic growth. A review of professional literature and practices in the field of RTI reveal several ways to determine a student's ROI. Some methods include visual inspection of data and others use mathematical formulas to calculate a slope statistic. Visual inspections might mean that a data analysis team looks at a student's data to determine if there is a trend in the data or might apply a decision rule based on the location of data points in relation to an aim line. Calculating a slope or statistic that represents the student's trend in progress is a way to quantify the student's progress.

An often-cited rule is the three-point decision rule (Wright, 1992). Either electronically or by hand, someone graphs a student's progress data and then an aim line is set. This is a line drawn from the student's initial data point to the goal at the end of a set period of time. If a student achieves three data points above or below the aim line, an instructional change is made. Lindsley (1990) popularized the use of the standard celeration chart that allows for uniform presentation of academic and behavioral data and includes standard ROI to which student data can be compared.

There are other methods that still require visual inspection of data but also include drawing a trend line. Shinn, Good, and Stein (1989) described the split-middle approach that does not necessarily provide a slope or number association with ROI, but can visually demonstrate a student's progress. The split-middle approach requires a graph to be drawn with the student's data points separated vertically by a line. The middle data points on each side of the graph are connected and considered to be the student's trend line. In a manual for using CBM, Fuchs and Fuchs (n.d.) describe the Tukey method, which further

separates the data into three sections, similar to the split-middle approach. The middle data point from the left and right sections are connected with a line, which is then referred to as the trend line for that set of data. From both the Split Middle and Tukey methods, an ROI statistic can be calculated; however, there is currently no research to support that practice (Flinn & McCrea, 2012).

To achieve an ROI statistic, there are several ways of graphing student progress data for which a slope can be calculated. For instance, the Iris Center, (n.d.) has suggested taking the last student data point minus the first student data point, and dividing that result by the number of weeks of the intervention. Unfortunately, this method does not account for the data points reported between the first and last data points and may not accurately represent the student's overall trend in performance.

Numerous articles that describe student ROI mention the use of ordinary least squares and linear regression (Ardoin, Christ, Morena, Cormier, & Klingbeil, 2013; Christ, 2006; Deno, Fuchs, Marston & Shin, 2001; Good, 1990; Fuchs et al., 1993; Jenkins, Graff, & Miglioretti, 2009; Shinn, Gleason, & Tindal, 1989; Shinn, Good, & Stein, 1989). Christ, Zopluoglu, Long, and Monaghen (2012) directly stated that using ordinary least squares is the accepted method for calculating trend lines from student data.

Hixson et al. (2008) recommended that educators consider multiple sources of variability before graphing data with a trend line such as effects of the intervention being used, controlling for environmental and subject variables, and sources of measurement error. The authors recommended adherence to standardization of administration of assessments and consideration when using goal-level versus grade-level materials. The authors mentioned the use of statistical analysis (i.e., ordinary least squares regression) as

a supplemental piece of information in addition to visual analysis of graphed data because single-subject design studies do not meet required assumptions for linear regression. Consideration should be given to outliers, excessive variability, and type of skill being measured (skill mastery versus general outcome). Educators need to ensure that enough data are collected to minimize standard error of measurement.

Tools for Graphing Progress and ROI

There are many ways to graph an individual student's rate of improvement (ROI). School teams may choose to graph progress by hand, use software on a computer (e.g., Microsoft Excel), use an online graphing tool (e.g., Chart Dog), or utilize charts available through a progress monitoring system (e.g., AIMSweb, STAR Reading). Microsoft Excel is probably the most common and mentioned most often, most likely because of its accessibility to educators (Flinn & McCrea, 2012; Flinn & McCrea, 2013; Kovaleski & Flinn, 2011; Kovaleski, VanDerHeyden, & Shapiro, 2013; McDougal, LeBlanc, & Hintze, 2010; VanDerHeyden & Burns, 2010).

Number of Data Points for a Trend Line

In educational research, 10 data points are recommended in order to have a stable trend line, or line of best fit through data collected over time (Gall, Gall, & Borg, 2007). Wright (1992) suggested that seven to eight data points are recommended when using the Tukey Method for drawing a trend line that is indicative of student performance and McMaster (2011) completed a study that found that eight to nine data points are necessary when determining progress of students on early writing measures to create a stable trend line. Christ et al. (2012) completed a simulation study with a large number of data sets for oral reading fluency and discovered that a very high number of data points are required to

calculate a stable trend line, and that number is dependent upon the set of CBM used. For instance, they found that 18 data points, collected once per week, were necessary when using R-CBM (oral reading fluency) from AIMSweb (www.aimsweb.com), which is far more than previous research had indicated.

There is evidence to suggest that educators can make meaningful instructional changes for students within the general education setting with only a few data points or visual analysis of data that has been graphed (Hixson et al., 2008). However, when the data may be considered as part of an evaluation for special education eligibility, educators need to ensure that the data they have align with recommendations in the literature. For instance, CBM should be administered with fidelity of standardization, there should be a sufficient number of data points, trend lines and ROI statistics (i.e., slope) should be calculated consistent with recommendations in the literature, and sources of error should be eliminated or minimalized. With these conditions met, educators are assured that the student's progress data are meaningful contributions to the determination of a student's eligibility for special education. An example from the medical field is if a patient is not feeling well, the doctoral may do a few simple tests such as temperature, blood pressure, and the patient's description of symptoms before arriving at the conclusion to prescribe an antibiotic. If the patient does not respond to a round of antibiotics, the doctor may need more information to decide on the next treatment.

Considerations for Determining ROI

Instructional decisions are frequently made with the use of CBM, especially in the area of reading. When it comes to making high-stakes decisions such as special education eligibility, the quality of data needs to be higher. Ardoin et al. (2013) sought information

about decision rules and psychometric properties of CBM in review of 102 sources. After a thorough review of research studies and manuals for using CBM, the authors concluded that CBM alone does not currently have the technical adequacy required for making high stakes decisions such as SLD eligibility but is sufficient for making general instructional decisions. This conclusion supports the basis for the current study, which is that many factors need to be considered when determining student ROI and that policy is perhaps including practices that are not yet evidence-based. Ardoin et al. (2013) described the numerous considerations when using data from the monitoring of student progress including standard error of measurement, the number of data points required to graph a stable trend line, and the variability of CBM probes (e.g., equivalency of reading passages). Other studies have questioned the use of CBM for non-English speaking students, the validity of monitoring progress from a student's enrolled grade level, and the comparison to various sets of norms. Additionally, the amount of growth that can be expected from students may also depend on the time of year or their initial performance (Kovaleski et al., 2013).

Studies that have reviewed the standard error of measurement when calculating student ROI have concluded with several recommendations. The stability of a trend line improves when eight to nine data points were entered into a graph using Microsoft Excel (Ditkowsky, 2009). Christ (2006) concluded that more data points gathered over time leads to less error in the trend line, which means that the trend line is more likely to accurately represent the student's performance and is less likely to be explained by error or variability of reading passages. The study also concluded that less chance of error was found in student data that were gathered in optimal assessment conditions such as

adherence to standardization of administration and a quiet room. In a similar study, Christ and Silberglitt (2007) found that the error of measurement surrounding trend lines of student data averaged between five and nine words read correctly per minute. This is problematic for instructional decision-making as explained in the following example. If a student is being monitored for oral reading fluency skills once per week, and makes a gain of six words from week to week, it is possible that the student's gain is just a function of the standard error of measurement, rather than absolute progress.

Another aspect of error comes from the variability between the assessment probes themselves. Ardoin and Christ (2009) reviewed the variability between reading passages for sets of CBM. For instance, the set of reading passages within a system of CBM may include 30 or more different probes that are not exactly the same. Students may naturally perform better on some probes than others, which would contribute to the variability in their data points across time, and skew the trend line as well. Ardoin and Christ concluded that CBM continues to be appropriate for general analysis of progress for groups of students but is not yet at the level of technical adequacy needed to evaluate individual student progress for special education.

Jenkins et al. (2009) compared monitoring of progress schedules for students at varying intervals such as once per week versus once per nine week CBM administrations to determine if the validity of the slope was negatively affected with less frequent progress monitoring. Results indicated that CBM could be administered less frequently if the administrations included multiple probes at a time. For instance, a teacher could effectively monitor student progress by completing two or more reading passages once

every five weeks rather than one passage every week for 10 weeks and yield a more accurate slope of progress.

Jenkins and Terjeson (2011) compared three types of decision-making rules to student data collected over a period of eight weeks to determine if one decision-making rule surpassed the others. Results indicated that decision rules based on the slope estimates (i.e., student ROI) were the best indicators of true progress rather than decision rules based on data points above or below a goal line.

A study seeking to answer the question of how much growth can be expected was completed by Silberglitt and Hintze (2007). The study described rates of growth that are typically found for students performing at various initial levels, which is important when determining student ROI. Students who scored at the higher and lower ends on the benchmark assessments were found to have the least amount of growth over time. In other words, the top performing students and lowest performing students tended to make the smallest gains. Implications for practice are that educators need to be aware of the comparison group(s) they are using when using student ROI for high stakes decisions, knowing that students growth is mediated by initial level of performance. These findings have been replicated in other studies analyzing growth mediated by level using CBM (Fien et al., 2010; Good et al., 2010).

Other considerations regarding the amount of growth than can be expected include differences in performance between semesters within the school year. When analyzing group benchmark data, all but one study (Graney, Missall, Martinez, & Bergstrom, 2009) has found more growth between the fall and winter benchmark assessments and less growth between winter and spring assessments (Ardoin & Christ, 2008; Christ, Silberglitt,

Yeo, & Cormier, 2010; Fien, Park, Smith, & Baker, 2010). Research is also emerging regarding the academic growth for students who are non-English speaking (Abu-Hamour, 2013; Farmer, Swanlund, & Pluymert, 2010).

A final consideration when analyzing individual student progress data is to indicate the group of comparison. There is literature available for which comparison group to use when reviewing student screening data. Several studies have recommended that the analysis of individual student performance on screening data be compared to local norms to provide the best indicator of student need (Hintze & Silberglitt, 2005; McGlinchey & Hixson, 2004; Shapiro, Keller, Lutz, Santoro, & Hintze, 2006; Silberglitt, Burns, Madyun, & Lail, 2006; Stage & Jacobsen, 2001; Stewart & Silberglitt, 2008). Norms provided by the publisher of commercially available CBM or national norms published in research results may not be relevant to the needs of individual students in a given school; therefore, the use of local norms when reviewing screening data is more likely to provide an accurate indication of finding the students who have a true need (Ferchalk, 2013). There does not appear to be consensus with which norm group to use when reviewing individual student ROI. VanDerHeyden and Burns (2010) recommend that school teams use local grade-level norms as an individual student ROI comparison as long as there are at least 75 students per grade level (p. 82). When there are an insufficient number of students per grade level, they recommend combining data from several grades to create a local norms group. It is implied in other literature that school teams will compare student ROI to user norms that include a national sampling of other students by grade level (Kovaleski et al., 2013).

Hixson et al. (2008) summarize the need to be thoughtful about the data and possible sources of error when determining student ROI, stating that "before adding a

trend line, it is important to carefully consider whether the overall pattern in the data is consistent and linear across time, or whether another pattern (nonlinear, curvilinear) better explains the data" (p. 2136). When making high-stakes (individual student) decisions, educators need to move away from what is easy and convenient and towards methods that satisfy statistical assumptions as evidenced in the literature. With an RTI process as a possible method for determining eligibility for SLD and the popularity of using CBM for monitoring student progress, there is a real need for consistency and technical adequacy for determining student ROI to avoid pitfalls of past methods (Flinn & McCrea, 2013). Because of the historical inconsistent use of the ability-achievement discrepancy between states, it is important to track how states are guiding schools to determine student ROI and continually update recommendations and policies consistent with research findings.

Recommended Practices for Determining ROI

Based on the review of key features of monitoring student progress and considerations for determining ROI, a number of recommended practices were gleaned for this study. Table 1 illustrates the connection between recommended practices (as numbered in Appendix A) and supporting literature.

Previous Research of State RTI Guidance Documents

Various researchers have captured incremental changes toward RTI implementation in state regulations and recommendations, particularly in relation to eligibility criteria for SLD. Initial studies focused on survey results from state directors of special education with the assumption that those personnel would be most knowledgeable about changes and proposed plans for implementing RTI (Ahearn, 2008; Hoover, Baca,

Table 1

Summary of Recommended Practices for Determining ROI and Supporting Literature

Recommended Practice	Supporting Literature
1. Use multiple administrations or repeated measures. Monitor student progress over time.	Deno (2003); Christ (2006); VanDerHeyden & Burns (2010)
2. Use measures with alternate forms or multiple assessments that are similar per grade level.	Ardoin et al. (2013); Deno (2003)
3. Use direct measures of academic skill(s).	Deno (1985); VanDerHeyden & Burns (2010)
4. Use brief, short measures.	Deno (1985); VanDerHeyden & Burns (2010)
5. Use measures validated for determining ROI: curriculum-based measurements (CBMs).	Ardoin et al. (2013); Christ et al. (2012); NCRTI (2011)
6. Use measures validated for determining ROI: computer adaptive tests (CATs).	Kovaleski et al. (2013); NCRTI (2011)
7. Examples of assessments for monitoring student progress: AIMSweb, DIBELS, STAR Math. Review of assessments listed on the National Center on Response to Intervention and National Center on Intensive Intervention web sites (http://www.intensiveintervention.org/chart/progress-monitoring.	NCRTI (2011)
8. Use technically adequate (i.e., valid, reliable) measures for skill(s) being monitored.	Ardoin et al. (2013)
9. Use measures sensitive to small increments of growth.	Deno (1985); Fuchs & Fuchs (1999)
10. Use grade-level CBM for monitoring student progress for special education eligibility. (Does not apply to CATs).	Burns & Sensenac (2005); VanDerHeyden & Burns (2010); Vaughn, Wanzek, Linan-Thompson, & Murray (2007)
11. Monitor student progress at Tier 1 at least three to four times a year.	Kovaleski et al. (2013); Vaughn et al. (2007)

Table 1 (continued)

Recommended Practice	Supporting Literature
12. Monitor student progress at Tier 2 at least monthly or every other week.	Kovaleski et al. (2013); Vaughn et al. (2007)
13. Monitor student progress at Tier 3 at least weekly.	Kovaleski et al. (2013); Vaughn et al. (2007)
14. School teams carefully consider and document whether a sufficient number of data points have been gathered to produce a stable trend line.	Ardoin et al. (2013); Christ (2006); Christ et al. (2012)
15. School teams carefully consider and document the number of data points used to make instructional decisions.	Ardoin et al. (2013); Christ (2006); Christ et al. (2012)
16. School teams consider whether assessments used to monitor student progress were administered with fidelity and in accordance with standardization.	Christ (2006); Christ & Silberglitt (2007); Christ et al. (2012)
17. First complete a visual inspection of student response data graphed on an x- and y-axis graph prior to generating a trend line.	Hixson et al. (2008)
18. After student response data are graphed, school teams determine if a linear trend line best describes the student's progress. School teams have a procedure for considering effects of an outlier data point.	Hixson et al. (2008)
19. Use ordinary least squares (OLS) linear regression to generate a trend line for determining ROI.	Ardoin et al. (2013); Christ et al. (2012); Deno et al. (2001)
20. Use a consistent comparison group (i.e., norm group) to determine adequacy of individual student ROI (e.g., national norms, user norms, or local norms).	Kovaleski et al. (2013)

Wexler-Love, & Saenz, 2008; Zirkel & Krohn, 2008). The remaining studies have documented changes in written information provided by the SEA (Hauerwas et al., 2013;

Zirkel, & Thomas 2010a, Zirkel, & Thomas, 2010b). The following is a detailed account of findings from relevant studies.

Hoover et al. (2008) completed a macro-level study of states' emphasis on the implementation of RTI as a multi-tiered framework for improving outcomes for all students and as a potential replacement of the ability-achievement discrepancy model. During the summer of 2007, the researchers sent a survey (field-tested by experts in RTI) to all state directors of special education (including District of Columbia). State directors were chosen as survey respondents because of the RTI publications authored by the National Association of State Directors of Special Education (NASDSE). Surveys were mailed once (63% response), mailed a second time (75% total response), then directors were sent emails or called on the phone (86% total response rate). Seven states (not identified for confidentiality) did not respond. The results of this study provided initial, macro-level data regarding national RTI training and implementation efforts. Some states expressed interest in using RTI to replace the ability-achievement discrepancy model but many reported that they were not yet ready to do so. The survey results indicated that training for culturally-responsive practices within RTI models were less prevalent than overviews of RTI, evidenced-based interventions, monitoring student progress, data-driven decision-making, and roles of educators in RTI. This study did not survey state directors of special education as to the details surrounding the monitoring of student progress or adherence to recommended practices.

Ahearn (2008) intended to gather updated information about state changes to SLD eligibility criteria as a result of IDEA (2006) regulations. Data were gathered via a survey that could be completed electronically or by paper copy and was sent to state and non-state

jurisdictions during the spring of 2008. The response rate was 49 out of 50 states and there was no mention of including or excluding the District of Columbia. Results indicated that 42 of 49 responding states had already made changes and the remaining seven have plans to make changes to SLD eligibility criteria. Of the 42 that changed six required an RTI process for SLD and eliminated ability-achievement discrepancy, 26 allow RTI process or ability-achievement discrepancy, and 10 allowed all three methods (i.e., ability-achievement discrepancy, RTI process, and other research-based process).Regarding training, states reported that a majority of training efforts have been state-wide as written in policies or procedures, state-wide or regional trainings, or as-needed trainings. Researchers noted the necessity of collaboration between general education and special education to implement RTI. Researchers also noted utility of the Internet for disseminating training materials to schools. This survey did not review use of recommended practices for monitoring student progress despite some states requiring the use of an RTI process.

Zirkel and Krohn (2008) completed a study to gather information about changes in state laws regarding SLD eligibility one year after IDEA regulations went into effect in October 2006. State directors of special education were sent an email survey and had a 100%response rate but results were usable for only 47 out of 50 states. Three states reported they were not ready to respond because changes were in process. Approximately half of the states reported that they had not yet finalized changes. Most states reported that they were going to allow districts to choose between the three possible SLD eligibility methods and only a few states were in the process of mandating an RTI process and eliminating the ability-achievement discrepancy altogether. The authors noted that a

paradigm shift is taking place because regular education needs to collaborate with special education in order for RTI (as a framework) to be implemented properly.

Berkeley, Bender, Peaster, and Saunders (2009) reviewed state websites and information provided by state representatives to capture stages of RTI implementation following the 2006 regulations. Results did not specify if data were gathered from state regulations or guidance documents. The authors concluded that many states had started the implementation of RTI but were inconsistent in features such as SLD eligibility criteria, statewide training, and components included in RTI implementation.

Zirkel and Thomas (2010a) intended to build upon previous snapshots of changes in state regulations and guidance documents following IDEA in a more systematic manner. Specifically, the authors reviewed state documents that were both legally binding and not legally binding (guidance documents). The authors systematically searched for case law related to RTI and SLD eligibility, state web sites, and then only if they were unable to find information, they contacted state representatives (i.e., state directors of special education or RTI coordinators). Information was gathered from fall of 2009. The article concluded that 12 states mandated RTI for SLD eligibility and noted that continued systematic review of state regulations and guidance is necessary to capture changes resulting from IDEA.

In a follow-up study, Zirkel and Thomas (2010b) researched whether states had changed their state regulations as a result of the reauthorization of IDEA (e.g., some states have mandated RTI as a process for identifying SLD). Building on previous studies, this study intended to systematically review state regulations and guidance documents for more specific components of RTI frameworks. The authors reviewed state education agency websites and if needed, made direct contact with state officials in education, but did

not specify their positions. Data were gathered from any published written laws or guidelines related to RTI implementation and did not include training materials. Documents were crosschecked with the NCRTI state database from 2009. Results revealed a list of states and source of information (law or guidance documents) and found scant and inconsistent information regarding state requirements or recommendations related to monitoring of student progress. Information regarding monitoring of student progress was not compared to recommended practices for calculating or interpreting (ROI).

Zirkel (2011) authored an article that reviewed previous snapshot data from state documents, described additional RTI implementation features such as the inclusion of behavior, and on a scale created by the author, described the language included in law or guidance documents. Zirkel described 15 states as partially or fully requiring RTI for SLD identification. An example of partially could be using an RTI process for SLD identification for grades kindergarten through four only. The author noted a continued need to systemically review state laws versus state guidelines to canvass incremental changes following IDEA regulations from 2006.

Hauerwas et al. (2013)uncovered the most recent definitions of RTI assessment processes related to SLD eligibility determination and built upon previous national snapshots of changes since IDEA regulations were set forth in 2006. The study was qualitative in nature, and utilized directed content analysis of state regulations and guidance documents related to the use of an RTI process for SLD eligibility determination. Relevant areas of this study included the review of (a) frequency of student RTI data collection, (b) criteria for defining responsiveness, and (c) multidimensional assessment. Information was gathered from state guidance documents during the summer of 2011. The

researchers searched websites of all 50 states for information regarding RTI and SLD identification (regulations or guidance documents). They used previously documented sources (Zirkel, 2010b; NCRTI, 2011) as a starting point for finding state guidance documents. If information was missing, the researchers sent up to three emails to state directors of special education, which resulted in the inclusion of only one additional document. They noted that this information resulted in a snapshot of information from October 2011. The authors found that there was no national consensus regarding parameters for using RTI data to determine SLD eligibility.

The researchers documented several other relevant findings. A total of 25 states included language in their state guidance documents regarding the type of data collected for monitoring student progress. Two states mentioned criteria beyond federal regulations in state regulations and eight states mentioned information beyond federal regulations in state guidance documents. In terms of how much data to gather, some states mentioned a minimum number of data points before changing interventions (i.e., three to four was most common) and other states mentioned a minimum number of data points required before establishing a trend (i.e., six to eight was most common; one state required 10, another state required 12). For defining student responsiveness, the researchers found that 12 states defined responsiveness in state regulations but only seven states defined responsiveness beyond federal regulations. Those states that further defined responsiveness referred to end-of-year benchmarks, performance compared to national or local normative data sets, or quantified a lack of responsiveness to be performance that is 1.25 standard deviations below the mean (but did not specify norms of comparison).

Overall, researchers emphasized lack of consistency and clarity for responsiveness and data collection/interpretation in state regulation and guidance documents. Researchers noted a need for continued canvassing of state documents because they found that documents had continued to evolve during their study. Researchers noted the need to improve use of research findings in state regulations and guidance documents.

CHAPTER III

METHODS AND PROCEDURES

Introduction

In response to research demonstrating numerous problems with the ability-achievement discrepancy model for determining eligibility for specific learning disabilities (SLDs) (Bocian, Beebe, MacMillan, & Grimes, 1999; Gresham & Vellutino, 2010; Vellutino, Scanlon, & Lyon, 2000), the reauthorization of the Individuals with Disabilities Education Act (IDEA, 2004) has required that states allow for an eligibility process demonstrating a student's response "to scientific, research-based intervention" (20 U.S.C. §1414(b)(6)). As more schools adopt this method for determining special education eligibility, recommended practices are emerging in the literature that provides guidance and considerations for analyzing student response data, particularly when using curriculum-based measurement (CBM; Ardoin, Christ, Morena, Cormier, & Klingbeil, 2013; Hixson, Christ, & Bradley-Johnson, 2008) and computer adaptive testing (CAT; Kovaleski, VanDerHeyden & Shapiro, 2013). The purpose of this study is to ascertain whether the language present in state guidance documents (i.e., regulations or recommendations) conforms to recommended practices in the literature for determining student rate of improvement (ROI).

Previous research studies that have documented changes in state policies and recommendations for using response to intervention (RTI)via a survey of state directors of special education (Ahearn, 2008; Hoover, Baca, Wexler-Love, & Saenz, 2008; Zirkel & Krohn, 2008) or a direct examination of state documents (Hauerwas, Brown, & Scott, 2013; Zirkel, & Thomas 2010a, Zirkel, & Thomas, 2010b). A review of previous research of the

changes in state policies since IDEA (2004) supports the methodological approach used in the current study and a need for an updated snapshot of changes since the most recent report of state documents is from the fall of 2011 (Hauerwas et al., 2013). This study will build upon the previous studies that have directly examined state guidance documents.

Research Design

This study used qualitative content analysis and descriptive statistics to explore several research questions related to the inclusion of recommended practices for determining student ROI in state regulatory or guidance documents. Descriptive statistics are used to summarize information in numerical form in an organized way (Bluman, 2012; Gall, Gall, & Borg, 2007). Because information was gathered from state documents and compared to preset categories (i.e., recommended practices) in a checklist format, descriptive statistics were used to illustrate the number and percentage of states that met criteria on the checklist. In addition, the mode and range of the years of publication of state documents were reported to highlight when states had updated regulations or guidance. The mode is referred to as the most common number in a set of numbers and the range is simply a description of the span of a set of numbers (Bluman, 2012).

Qualitative research allows for investigation of information outside of experimental settings and focuses on the depth of information gathered rather than the measurement or assignment of numbers to human behavior (Denzin & Lincoln, 2011). Qualitative research can be a fluid process that infers the meaning in methods of communication with themes or definitions that emerge throughout the research process. In other words, researchers who use qualitative methods may begin a study with a list of research questions and end the study with an expanded list based on the information discovered during the study.

Similar to the Hauerwas et al. (2013) study, this study intended to derive meaning from written information from states, which aligns with defining features of qualitative research and content analysis. Krippendorff (2012) stated that content analysis is qualitative research because the researcher is initially inferring meaning from written information, even if it is to be quantified at some point in the research process. Stemler (2001) reported that qualitative content analysis can be a reliable measure of a large quantity of text to discover themes or compare text to preset, mutually exclusive categories (e.g., a checklist) when categories are well-defined. Krippendorff similarly reported that qualitative content analysis can be exploratory or inferential. Exploratory content analysis may review communication (e.g., presidential speeches) to determine the most common trends or the number of times a phrase is used, whereas inferential content analysis compares trends in information to preset categories, as is the case with this study. The preset categories in this study consist of the items on the checklist of recommended practices for determining ROI (see Appendix A). The content that was analyzed for the presence of the checklist items included state regulatory and guidance documents.

Qualitative content analysis was formed from a constructivist view since it relies on the researcher's interpretation about the meaning of the information gathered (Krippendorff, 2012). With advancements in technology and increased use of web-based dissemination of information, content analysis does not necessarily rely on laborious review by humans. Computer software programs allow for information to be scanned according to preset categories much more efficiently and effectively. Benefits of using content analysis versus a survey are that it is less intrusive and the information is gathered directly from the source, thus reducing the problem of survey respondents responding

differently because they are aware they are in a study. Because data were be gathered from the text within state documents, the method used to analyze state guidance documents was qualitative content analysis.

Population and Sample

The population for this study included written state policies, procedures, and guidelines from 51 state education agencies (SEAs; all 50 states and the District of Columbia). All documents were online and readily available to the researcher. Written regulations and guidance documents were available for download as portable document format commonly known as a pdf while a few others were in the form of published websites endorsed by the state. The sample was limited to information that appeared to be sanctioned by the state, which is consistent with previous research of state documents (Zirkel & Thomas, 2010b). Documents set forth for training purposes were excluded from this study, which is consistent with previous research on this topic (Hauerwas et al., 2013).

Research Tools

This study compared state requirements or recommendations for determining student ROI to a checklist of recommended practices developed by the researcher from a thorough review of related literature (for the checklist, see Appendix A; research review, see Table 1). A checklist of recommended practices for determining ROI did not already exist at the time of this study. The checklist was sent to a panel of three experts for refinement, addition, or deletion of checklist elements and changes were made as necessary. The checklist included items to gather data about recommended practices for determining ROI and whether the state mandates an RTI process for SLD eligibility. Additionally, four checklist items were included to gather information as to the number of

data points needed for making instructional changes, whether a quantifiable magnitude of deficiency in level was specified, and whether procedures were specified by SLD category or grade level.

The final revision occurred after piloting it with a random sample of 16 state documents. The random sample was selected using the Random List Generator (RANDOM.ORG, 2014). During this trial, the researcher kept notes of additional words that could be used to complete a thorough search (see Table 2) of the documents and possible changes to the wording of the checklist items. After this initial trial the following changes were made. First, several key search terms were added to Table 2 to improve chances of finding a recommended practice within a state document. For instance, item number two (see Table 2) initially included key terms of alternate forms, similar assessments, and grade level. However, after reviewing the sample of state documents, other terms were commonly used such as repeatable forms, equivalent measures, and equal forms. Therefore, those terms and phrases were added to the key term search list.

Next, two items were collapsed into one. There were initially two items that separately required searches for requirements of valid (one item) and reliable (second item) assessments. After the initial review of 16 state documents, it was noted that if states suggested the use of valid or reliable assessments, the terms were often used together. In addition, terms such as technically adequate assessments or measures with technical adequacy were used; therefore, those terms were included as key search terms in Table 2.

The final significant change was moving from using a Microsoft Word document to track information to using Google Forms. Google Forms provided a simple online form to score the state documents, from which results were directly imported into a Google

Spreadsheet. The Google Spreadsheet was then downloaded as a Microsoft Excel spreadsheet for final analyses.

Procedures

Data Collection

To gather data for this study, the following steps were completed. The researcher identified the relevant state documents by reviewing past resources (Hauerwas et al., 2013; Zirkel & Thomas, 2010a; Zirkel & Thomas, 2010b) and information provided on the National Center on Response to Intervention state database website (NCRTI, 2010). Specifically, the researcher used the list of state regulations and guidance documents in the Hauerwas et al. study and searched for each using "bing," a popular internet search engine (www.bing.com). If documents were not available from that list, the researcher reviewed the website for the state department of education to look for state requirements or guidance for implementing RTI and/or determining eligibility of SLD that appeared to be endorsed by the state department of education (i.e., had state department of education logo, was listed on the state department of education's website). Using this method revealed regulations for each state and guidance documents for all but two states (District of Columbia and North Carolina). Because all regulations and guidance documents were readily available online, including updated documents since the Hauerwas et al. study, there was no need to contact state directors of special education to confirm accuracy of state documents.

State guidance documents were available in a variety of formats including electronic formats (i.e., portable document format [.pdf], Microsoft Word [.doc, .docx], or text on web pages). Key word in context (KWIC) searches were completed to determine the presence

or absence of recommended practices for determining student ROI within state regulations or guidance documents. The researcher searched for information in each state's documents by recommended practice as indicated in Table 2. In general, if key words or phrases were found, the researcher read the surrounding text to determine if the document was describing the recommended practice being searched. If a recommended practice or disqualifying content was found, the researcher would highlight that section so it could later be reviewed using the checklist to determine the presence or absence of the recommended practice.

The following processes were used to score state documents using the checklist of recommended practices for determining ROI. State documents were reviewed one at a time. The first three items on the checklist served to gather demographic information from the state document being reviewed. To track the state, the two initial were entered into the Google Form (e.g., Pennsylvania was entered as PA). The next item required the year of publication to be typed (e.g., 2008). Documents were then scored as either regulations or guidance. Documents were scored as being regulations when there were SEA specific or federal regulation codes reported sans interpretation. Documents were scored as guidance when the content included recommendations, suggestions, or interpretations of state or federal regulations that were not legally binding.

A majority of checklist items required a score of either "evidence found" or "no evidence found" (items number 1-6, 8-9, 11-14, and 16-21). For these items, the researcher would have the state document and the checklist (via Google Forms) open on a computer, with a paper printout of Table 2. The researcher would search for checklist items one at a time. The researcher would use the KWIC search feature to search the state

Table 2

Search for Recommended Practices by Key Words and Phrases

Recommended Practice	Key Words and Phrases
1. The state guidance document describes use of multiple administrations of assessments; monitor student progress over time; progress monitoring; formative assessment.	repeated assessments; repeated measures; progress monitoring; formative assessment; monitor student progress
2. The state guidance document describes use of measures that have alternate forms or multiple assessments that are similar per grade level.	alternate forms; similar assessments; per grade level; repeatable forms; equivalent; equal;
3. The state guidance document describes use of direct measures of academic skill(s) for monitoring student progress.	direct; directly; measured; academic; skills; achievement
4. The state guidance document describes use of brief, short measures for monitoring student progress; measures that are efficient for monitoring student progress.	brief; short; quick; efficient; minute
5. The state guidance document describes use of measures that are validated for determining ROI: curriculum-based measurements (CBMs).	curriculum based; curriculum-based; CBM
6. The state guidance document describes use of measures that are validated for determining ROI: computer adaptive tests (CATs).	computer adaptive; CAT
7. Critical Feature: The state guidance document cite examples of assessments such as AIMSweb, DIBELS, STAR Math, or suggests a review of assessments listed on the National Center on Response to Intervention and National Center on Intensive Intervention web sites (http://www.intensiveintervention.org/chart/progress-monitoring).	National Center on Response to Intervention; National Center on Intensive Intervention; NCRTI; examples of progress monitoring; examples of curriculum-based measures; examples of computer adaptive tests; example formative Assessments

Table 2 (continued)

Recommended Practice	Key Words and Phrases
Disqualifying Content: The state guidance document cites non-examples including teacher-made tests, locally developed assessments, unit tests, grades, individually administered standardized norm-referenced assessments of achievement (e.g., Woodcock-Johnson Tests of Achievement, Third Edition; Wechsler Individual Achievement Test, Third Edition).	teacher made; unit; grades; standardized; norm-referenced; locally developed
8. The state guidance document directs school teams to use measures that are technically adequate (i.e., valid, reliable) indicators of the skill(s) being monitored.	technical; adequate; valid; reliable; reliability
9. The state guidance document describes use of measures that are sensitive to small increments of growth (progress, change, response to instruction/ intervention).	sensitive; small; change; increment; growth; progress
10. Critical Feature: The state guidance document describes monitoring student academic progress on grade level (in which the student is enrolled) for special education eligibility if CBM is suggested. (Does not apply to CATs since they automatically adjust level of difficulty of items).	grade level; instructional level
Disqualifying Content: The state guidance document describes use of instructional level data for special education eligibility.	
11. The state guidance document describes monitoring of student progress at Tier 1 with a frequency of at least three to four times a year.	universal; screening; tier 1; tier one; three times; four times; per year; benchmark; fall; winter; spring; beginning; middle; end
12. The state guidance document describes monitoring of student progress at Tier 2 with a frequency of at least monthly or every other week.	progress monitoring; tier 2; tier two; month; week
13. The state guidance document describes monitoring of student progress at Tier 3 with a frequency of at least weekly.	progress monitoring; tier 3; tier three; week; at least

Table 2 (continued)

Recommended Practice	Key Words and Phrases
14. The state guidance document directs school teams to carefully consider and document whether a sufficient number of data points have been gathered to produce a stable trend line.	number of; data point; trend; stable; trendline
15. The state guidance document directs school teams to carefully consider and document a specific number of data points for instructional decision-making.	number; data; data point; instruction; change;
16. The state guidance document directs school teams consider whether assessments used to monitor student progress were administered with fidelity and in accordance with standardization.	fidelity; administration; standardization;
17. The state guidance document describes first completing a visual inspection of student response data graphed on an x- and y-axis graph prior to generating a trend line.	visual; graph;
18. The state guidance document describes that after student response data are graphed, school teams should then determine if a linear trend line best describes the student's progress. In other words, school teams have a procedure for considering effects of an outlier data point.	linear; trend; outlier; line of best fit
19. The state guidance document describes use of ordinary least squares (OLS) linear regression to generate a trend line for determining ROI.	least squares; OLS; linear; regression; line of best fit
20. The state guidance document directs school teams to consider the comparison groups(s) (i.e., norm group(s)) to determine adequacy of individual student ROI. School teams should determine which set(s) of norms most closely match the population of the grade level of the student referred for special education evaluation. Examples: -Student ROI should be compared to national norms available through assessment systems or research results. -Student ROI should be compared to the ROI provided by a commercially available set of assessments or assessment tools.	comparison; peers; norm; rate of; slope; national; user; commercial; compared to;

Table 2 (continued)

Recommended Practice	Key Words and Phrases
-Student ROI should be compared to local norms within a school building, grade level, or specific classroom.	
21. The state guidance document indicates that a response to intervention (RTI) process is required or mandated for determining specific learning disability (SLD) eligibility for special education.	response to intervention; RTI; mandated; required; specific learning; SLD; special education; eligibility
22. The state guidance document describes a specific magnitude of deficiency in relation to age- or grade-level standards.	magnitude; deficient; deficiency; in relation to; age-level; grade-level; below peers; compared to peers; deviation/s below; achievement gap; below
23. The state guidance document specifies requirements for using an RTI process for special education eligibility for the following SLD types (Check all that apply): ___Basic reading skill ___Reading fluency skills ___Reading comprehension ___Math Calculation ___Math Reasoning ___Written Expression ___Listening Comprehension ___Oral Expression	specific learning; SLD; reading; math; written; listening; expression
24. The state guidance document specifies the use of an RTI process for only the following grade levels. Check all that apply. Kindergarten 1st Grade 2nd Grade 3rd Grade 4th Grade 5th Grade 6th Grade 7th Grade 8th Grade 9th Grade 10th Grade 11th Grade 12th Grade (or through maximum school-age limit)	grade; response to; 1; first; 2; second; 3; third; 4; fourth; 5; fifth; 6; sixth; 7; seventh; 8; eighth; 9; ninth; 10; tenth; 11; eleventh; 12; twelfth

document for each word, in order, using the key words listed in the right column of Table 2. If a key word or phrase was found within the document, the researcher would read the surrounding text to determine if the key word was in the context of the checklist item being searched. For example, the first key phrase listed for checklist item number eight is "technical." The KWIC search would highlight every instance of the word "technical" or words including the word "technical" within it (e.g., technically). The researcher would read the surrounding text to determine if the word "technical" was used in the context of item number eight, which involves technical adequacy of measures. An example of a score indicating "evidence found" would be a statement such as "when monitoring student progress, the technical adequacy of measures should be considered." A non-example or score of "no evidence found" would be a reference in the document for school teams seeking technical assistance. After reviewing the text surrounding each instance of the key word "technical" in the document and highlighting the sections in which it was found, the researcher would search for the next key word listed on Table 2 for item number eight, "adequate." Only after searching the document for each key term for item number eight would the researcher determine if there was evidence. The sections that included the key term used in the correct context were highlighted in yellow to allow for easy reference in the future if needed and for the spot checks to be completed more efficiently.

For two items (numbers 7 and 10), in addition to the option of scoring as "evidence found" or "no evidence found," there was an option for "disqualifying content." The same search process described above was used but for these items, terms for disqualifying content were included. For example, item number 10 requires a search for "grade level" to determine if evidence was found. To determine if disqualifying content existed for item

number 10, the terms "instructional level" were also searched. If using instructional level data for special education eligibility was required or recommended, the item was scored as having disqualifying content. For items 7 and 10, even if there was information indicating evidence found, if disqualifying content was also found, then the item was scored as having disqualifying content.

For checklist items number 15 and 22, the response required entry of text. The same search process was used but instead of scoring the item as having or not having evidence, information directly taken from the state document needed to be typed. For example, if information was found for item number 15 indicating a required or recommended number of data points for making instructional decisions, then the researcher would type the number or range of data points in the text box (e.g., three data points, four to six data points).

Finally, items number 23 and 24 allowed the researcher to check all items that apply. The search process was the same to determine if information existed related to the description of the item, but instead of evidence found or not found, the researcher checked specific items from a list or checked "did not specify." For instance, if evidence for item number 24 was found where the state specified SLD eligibility criteria for certain grade levels, then the researcher would check next to all of the grade levels listed in the state document. If there were no specifications for eligibility designated by grade level, the researcher checked "did not specify."

For the purpose of this study, the absence of information beyond the adoption of language provided in the IDEA regulations (2006) was recorded as an absence of recommended practices from the SEA. For instance, if the SEA described the use of

"repeated measures" but did not elaborate that this means monitoring of student progress, this was checked as "no evidence" on the checklist.

Determination of Inter-Rater Reliability

Krippendorff (2012) described the need for inter-rater reliability when completing qualitative studies. Inter-rater reliability is determined by having two or more observers review the same event or information and calculating a percentage of agreement to determine if the researchers reached the same conclusion. Landis and Koch (1977; as cited by Hallgren, 2012) provided the guidelines for interpreting Cohen's kappa. Slight agreement falls between 0.0 and 0.2; fair agreement between 0.21 and 0.4; moderate agreement between 0.41 and 0.6; substantial agreement between 0.61 and 0.8; and near perfect to perfect agreement between 0.81 and 1.0. When possible, information being reviewed should yield an inter-rater reliability Cohen's kappa of at least .80 (i.e., 80% agreement; Hallgren, 2012). Establishing inter-rater reliability reduces effects of researcher bias and improves the reliability of the study (Bluman, 2012; Gall et al., 2007).

For this study, inter-rater reliability was established using Cohen's kappa, which was calculated using Microsoft Excel according to guidelines presented by Zaiontz (2014). To establish inter-rater reliability, the researcher utilized one other person to complete spot checks with approximately 10% of states (i.e., five states, which equaled 14 state documents). The states used for spot checks were chosen using the Random List Generator from RANDOM.ORG (2014). All 51 states were entered as a list in alphabetical order into the Random List Generator, were randomized, and the first five states were selected. The second researcher who completed the spot checks was a volunteer who was a doctoral level, practicing school psychologist, familiar with RTI, data analysis, and systems-level

change. In order to minimize the time spent searching documents, the second researcher was given documents that had key words and phrases already highlighted. The second researcher then used the same Google Forms and Table 2 to score the checklist of recommended practices for determining ROI.

Summary

This chapter described the use of a qualitative content analysis and descriptive statistics as the research design for this study. The population and sample were discussed, noting that the population consisted of state regulations and recommendations from all 51 states. The development of and final forms of research tools were discussed. Procedures were reviewed including steps taken for data collection and key search terms used to find recommended practices in state documents. The use of a second researcher for reviewing a subset of state documents was discussed including the determination of inter-rater reliability.

CHAPTER IV

DATA AND ANALYSIS

Introduction

With the passing of regulations from the Individuals with Disabilities Education Act (IDEA; 2006), state education associations (SEAs) may no longer require the antiquated ability-achievement discrepancy for identifying a student with a specific learning disability (SLD). The regulations from IDEA were meant to encourage the use of a response to intervention (RTI) process, which requires frequent monitoring of student progress over time. Evidence in the literature has indicated that there are better ways than others to monitor student progress, especially when making high-stakes decisions such as special education eligibility. For this study, a checklist of such recommended practices was compiled to determine if states are currently requiring or recommending teams to consider the recommended practices found in literature.

This chapter reviews the results of a thorough review of state regulations and guidance documents from all 51 states. In Table 3, the titles of state documents are provided, sorted by state, year published (if available), and document type (i.e., either regulation or guidance). Results of inter-rater reliability for the analysis performed are presented. Checklist items and explicit decision rules are reviewed and results are presented per checklist item in relation to the research questions.

Document Characteristics

Many states had updated or added documents (66 out of 112, or 59%) since the recent Hauerwas et al. (2013) study which had captured information in documents up to the fall of 2011. Eight of the 112 state documents did not have dates and were scored as

Table 3

List of State Documents by Regulation or Guidance, Date of Publication, and Title

State	*Regulation Guidance*	*Date*	*Title of Document*
Alabama	Regulation	2011	Rules of the Alabama State Board of Education Chapter 290-8-9
	Guidance	2009	Response to Instruction (RtI): Alabama's Core Support for All Students
Alaska	Regulation	2013	State Special Education Handbook
	Guidance	2009	Using Response to Instruction/Intervention (RTI) for Alaska's Students
Arizona	Regulation	2012	AZ-TAS Evaluation and Eligibility Processes and Procedures from Referral to Determination of Eligibility
	Guidance	2011	AZ Response to Intervention (RTI): Elementary Technical Assistance Paper
	Guidance	2011	AZ Response to Intervention (RTI): Secondary Technical Assistance Paper
Arkansas	Regulation	2008	Arkansas 6.00 Evaluation-Eligibility Criteria Specific Learning Disability
	Guidance	2011	Integrating the School Prevention, Review, and Intervention Team (SPRINT) and Response-to-Instruction/Intervention (RtI2) Process: A Model Implementation Guidebook for Schools and Districts
California	Regulation	2014	California Code of Education, Section 3030
	Guidance	2009	Determining Specific Learning Disability Eligibility Using Response to Instruction and Intervention (RTI^2)
Colorado	Regulation	2013	ECEA Rules for the Administration of the Exceptional Children's Education Act 1 CCR 301-8
	Guidance	2011	Guidelines for Identifying Students with Specific Learning Disabilities
Connecticut	Regulation	2012	State of Connecticut Regulations State Department Of Education (Update)
	Guidance	2010	Guidelines for Identifying Children with Learning Disabilities
	Guidance	2008	Using Scientific Research-Based Interventions for Improving Education for All Students: Connecticut's Framework for RTI
Delaware	Regulation	2011	Title 14 Education Delaware Administrative Code
	Guidance	2008	Response to Instruction Implementation 2008-2009 School Year: Who, What, and When
District of Columbia	Regulation	2011	3000 Special Education Policy

Table 3 (continued)

State	*Regulation Guidance*	*Date*	*Title of Document*
Florida	Regulation	2013	Volume I-B: Florida Statutes and State Board of Education Rules: Excerpts Related to Exceptional Student Education
	Guidance	2011	Guiding Tools for Instructional Problem Solving
Georgia	Regulation	2010	Eligibility Determination and Categories of Eligibility
	Guidance	2011	Response to Intervention: Georgia's Student Achievement Pyramid of Interventions
Hawaii	Regulation	n.d.	Hawaii Administrative Rules: Chapter 60 Guidelines
	Guidance	2012	2012 Comprehensive Student Support System Introductory Guide
Idaho	Regulation	2009	Idaho Special Education Manual
	Guidance	2009	Response to Intervention-Idaho: Connecting the Pieces
Illinois	Regulation	2007	Illinois Administrative Code 226
	Guidance	2012	Illinois Special Education Eligibility and Entitlement Procedures and Criteria within a Response to Intervention (RTI) Framework: A Guidance Document
Indiana	Regulation	2012	Indiana Administrative Code, Article 7
	Guidance	2011	Considerations in SLD Evaluations and Eligibility Determination
Iowa	Regulation	2011	Title VII Special Education Chapter 41 Special Education (and 2011 Amendments to Iowa's Administrative Rules of Special Education)
	Guidance	2011	Iowa Department of Education: Guidance Document Response to Intervention
Kansas	Regulation	2011	Kansas Special Education Services Process Handbook (sections 91-40-11)
	Guidance	2013	Kansas Multi-Tier System of Supports Mathematics
	Guidance	2014	Kansas Multi-Tier System of Supports Reading
Kentucky	Regulation	2008	707 KAR 1:310. Determination of eligibility.
	Guidance	2012	A guide to the Kentucky System of Interventions
	Guidance	2014	Specific Learning Disabilities Eligibility Guidance Document (2014)
Louisiana	Regulation	2009	Bulletin 1508 – Pupil Appraisal Handbook
	Guidance	n.d.	Response to Intervention (RTI) in Louisiana
Maine	Regulation	2013	05-071 Chapter 101: Maine Unified Special Education Regulation Birth to Age Twenty
	Guidance	2009	Response to Intervention Guide

Table 3 (continued)

State	*Regulation Guidance*	*Date*	*Title of Document*
Maryland	Regulation	n.d.	COMAR 13A.05.01.06
	Guidance	2008	A Tiered Instructional Approach to Support Achievement for All Students: Maryland's Response to Intervention Framework
Massachusetts	Regulation	2008	Specific Learning Disability Team Determination of Eligibility Mandated Forms 28/10
	Guidance	2008	Specific Learning Disability: Eligibility Requirements
	Guidance	2011	Massachusetts' System of Tiered Instruction Draft
Michigan	Regulation	2013	Michigan Administrative Rules for Special Education Supplemented with IDEA Federal Regulations (Part 1)
	Guidance	2010	Michigan Criteria for Determining the Existence of a Specific Learning Disability
Minnesota	Regulation	2008	3525.1341 Specific Learning Disability
	Guidance	2011-2012	Determining the Eligibility of Students with Learning Disabilities: A Technical Manual (multiple documents)
Mississippi	Regulation	2009	State Policies Regarding Children with Disabilities Under the Individuals with Disabilities Education Act Amendments of 2004
	Guidance	2010	Response to Intervention Best Practices Handbook
Missouri	Regulation	n.d.	III: Identification and Evaluation
	Guidance	2008	State of Missouri RtI Guidelines: Appendix A
Montana	Regulation	2007	Montana State Plan Under Part B of the Individuals with Disabilities Education Act
	Guidance	2008	Montana Response to Intervention (RTI) Framework
	Guidance	2008	RTI Documentation
Nebraska	Regulation	2014	Nebraska Department of Education Rule 51 Regulations and Standards for Special Education Programs Title 92, Nebraska Administrative Code, Chapter 51
	Guidance	n.d.	Learn About RTI
Nevada	Regulation	2008	Adopted Regulation R064-08 Adopted Regulation of the State Board of Education LCB File No. R064-08
	Guidance	n.d.	Response to Intervention
New Hampshire	Regulation	2008	New Hampshire Rules for the Education of Children with Disabilities
	Guidance	2009	Interactive Guide to RtI in New Hampshire

Table 3 (continued)

State	*Regulation Guidance*	*Date*	*Title of Document*
New Jersey	Regulation	2013	N.J.A.C. 6A:14, Special Education
	Guidance	2008	Resource Manual for Intervention and Referral Services (I&RS)
New Mexico	Regulation	2013	Policies and Procedures for the Provision of Special Education Services for Students with Disabilities and Gifted Students Chapter 4
	Guidance	2011	Technical Evaluation and Assessment Manual
	Guidance	2009	The Student Assistance Team and the Three-Tier Model of Student Intervention: New Mexico's Response to Intervention Framework
New York	Regulation	2013	Regulations of the Commissioner of Education. Part 200 Students with Disabilities
	Guidance	2010	Response to Intervention Guidance for NY State School Districts
North Carolina	Regulation	2013	Policies Governing Services for Children with Disabilities
North Dakota	Regulation	2012	Chapter 67-23-06 Response to Intervention
	Guidance	n.d.	Response to Intervention in a Unified North Educational System
	Guidance	2007	Guidelines: Identification and Evaluation of Students with Specific Learning Disabilities
Ohio	Regulation	2013	Chapter 6.7 Specific Learning Disabilities – 3301-51-06(H)
	Guidance	2013	Guidance: Eligibility Determination – Specific Learning Disabilities
Oklahoma	Regulation	2014	Special Education Handbook
	Guidance	2010	Response to Intervention (RtI) Guidance Document
Oregon	Regulation	2013	Oregon Administrative Rules for Special Education
	Guidance	2013	Oregon RTI Handbook: A Model for Implementation and Documentation of RTI Practices
Pennsylvania	Regulation	2008	§14.125. Criteria for the Determination of Specific Learning Disabilities
	Guidance	2008	PA Guidelines for Identifying Students with Specific Learning Disabilities
Rhode Island	Regulation	2013	Rhode Island Regulations Governing the Education of Children with Disabilities
	Guidance	2010	Rhode Island Criteria and Guidance for the Identification of SLD

Table 3 (continued)

State	*Regulation Guidance*	*Date*	*Title of Document*
South Carolina	Regulation	2013	South Carolina Code of Regulations Unannotated
	Guidance	2013	Special Education Process Guide for South Carolina
	Guidance	2011	SC Response to Intervention: A Framework and Technical Assistance Guide for Districts and Schools
South Dakota	Regulation	2013	Eligibility Guide
	Guidance	n.d.	Specific Learning Disability
	Guidance	2013	Response to Intervention Implementation Guide: The South Dakota Model
Tennessee	Regulation	2014	Rules of the State Board of Education Chapter 0520-01-09 Special Education Programs and Services
	Guidance	2010	Special Education Manual
Texas	Regulation	2012	Special Education Rules and Regulations
	Guidance	2013	Evaluation of Learning Disability (LD) Eligibility
	Guidance	2008	2008-2009 Response to Intervention Guidance
Utah	Regulation	2013	Utah State Board of Education Special Education Rules
	Guidance	2008	USOE Specific Learning Disability Eligibility Guidelines
Vermont	Regulation	2013	State of Vermont Special Education Rules
	Guidance	2008	Field Memo: FM#08-07
Virginia	Regulation	2010	Regulations Governing Special Education Programs for Children with Disabilities in Virginia
	Guidance	2007	Responsive Instruction: Refining Our Work of Teaching All Children
Washington	Regulation	2013	State of Washington Rules for the Provision of Special Education
	Guidance	2006	Using Response to Intervention (RTI) for Washington's Students
West Virginia	Regulation	2012	Policy 2419 Regulations for the Education of Students with Exceptionalities
	Guidance	2012	Support for Personalized Learning: Guidance For West Virginia Schools and Districts
Wisconsin	Regulation	2010*	Chapter PI 11.36(6) *Updated each month
	Guidance	2013	Wisconsin's Specific Learning Disabilities (SLD) Rule: A Technical Guide for Determining the Eligibility of Students with Specific Learning Disabilities

Table 3 (continued)

State	*Regulation Guidance*	*Date*	*Title of Document*
Wyoming	Regulation	2010	Wyoming Department of Education Special Programs Division Policy and Procedure Manual for Special Education
	Guidance	2010	Chapter 7, Services for Children with Disabilities
	Guidance	2011	A Model Response to Intervention (RtI) Learning Disabilities

n.d. Of the remaining documents, the dates ranged from 2006 to 2014 with a median year of 2011 and most frequent year (i.e., mode) being 2013. A regulation document was found for each state. All states except the District of Columbia and North Carolina had at least one guidance document, with some states having two. Guidance documents provided recommendations for identifying students with SLD, others suggested implementation strategies for an RTI framework, and some documents covered eligibility for special education within an RTI framework.

Inter-Rater Reliability

In order to establish reliability of the ROI checklist, another researcher was used to conduct spot checks on a subset of all state documents, which consisted of documents from approximately 10% of the 51 states, and equaled 14 out of a total of 112 documents. Regulations and guidance documents from the five states were used as opposed to randomly selecting documents from a pool of 112 total documents to ensure that both regulations and guidance documents were proportionately included in the subset of documents reviewed. The states chosen for spot checks were selected through the Random List Generator (RANDOM.ORG, 2014) and included Connecticut, New Mexico, North Dakota, South Carolina, and Vermont.

Cohen's kappa was used to determine agreement between researchers and was calculated in Microsoft Excel. The percent of agreement is calculated by adding the cells diagonally from left to right and dividing by the total. Using the data is Table 4 for example, the researcher would compute (4 + 9) / 14 for a percentage of agreement equaling 0.93. The chance of agreement, also known as the probability of agreement if researchers randomly scored items, is calculated as follows. Researcher 1 found evidence a total of four times out of 14 documents which equates to 0.29. Researcher 2 found evidence five times out of 14 documents which equates to 0.36. Therefore the probability that both researchers would find evidence is 0.29 x 0.36 = 0.10. The probability that both researchers would not find evidence is (1 – 0.29) x (1 – 0.36) = 0.45. The probability of both researchers agreeing in general is 0.10 + 0.45 = 0.55. Then using the kappa formula in Equation 1, inter-rater reliability was (0.93 – 0.55) / (1 – 0.55) = 0.84.

Table 4

Example Data for Calculating Inter-Rater Reliability

		Researcher 1		
		Evidence	No Evidence	Total
	Evidence	4	1	5
Researcher 2	No Evidence	0	9	9
		4	10	14

Inter-rater reliability was calculated per checklist item, spanning all 14 documents. Results are reported in Table 5. Item numbers 15, 22, 23, and 24 were not included in the calculation for inter-rater reliability because they were open-ended questions used to gather supplemental information.

Table 5

Inter-Rater Reliability by Checklist Item

Checklist Item Number	Cohen's kappa	PABAK
1. Monitoring progress	0.39	
2. Alternate forms	0.43	
3. Direct measures	0.46	
4. Brief measures	0.55	
5. Curriculum-based measurement	0.85	
6. Computer adaptive test	0.00	1.00
7. Assessment examples	0.42	
8. Technically adequate measures	0.84	
9. Sensitive to growth	0.46	
10. Grade-level monitoring	0.72	
11. Tier 1 monitoring	0.81	
12. Tier 2 monitoring	0.42	
13. Tier 3 monitoring	0.43	
14. Stable trend line	0.44	
16. Fidelity of assessment administration	0.76	
17. Visual inspection, graph	1.00	
18. Consider outlier data	0.00	0.86
19. Least squares, linear regression	0.00	0.86
20. Comparison, norm group	0.46	
21. RTI as mandated eligibility process	0.70	

As data were entered into Excel for Cohen's kappa, there were a few items (numbers 6, 18, and 19) where this statistic was not appropriate and did not explain the level of agreement. When items resulted in a Cohen's kappa of zero (0), the raw data were studied in more detail for explanation. Byrt, Bishop, and Carlin (1993) explained that two common phenomenon can result in misleading Cohen's kappa metrics: effects of bias or prevalence. Bias effects occur when one researcher demonstrates a pattern of responses that favor certain categories when compared to responses of the other researcher. Prevalence effects interfere with Cohen's kappa when, even despite a high percentage of agreement, the

distribution of scores leans toward one direction only. For purposes of illustration, results are considered to be in a positive direction if both researchers agree that evidence exists.

In contrast, if both researchers agree that evidence is lacking, results are reported in a negative direction. An example of results falling in a negative direction only is illustrated in Table 6 where both researchers agreed for 13 out of 14 documents there was no evidence found for considering outlier data (item 18). Although 13 instances of agreement out of 14 equates to a high percentage of agreement (92.9%), a Cohen's kappa was calculated to be zero because all agreements occurred under "No Evidence" leaving the top left corner with a zero. As reported by Sim and Wright (2005), when researchers agree on items that fall only in one direction, Cohen's kappa should be interpreted with caution. One

Table 6

Example of Prevalence Effect Using Checklist Item Number 18 Data

		Researcher 1		
		Evidence	No Evidence	Total
	Evidence	0	1	1
Researcher 2	No Evidence	0	13	13
	Total	0	14	14

way of adjusting the Cohen's kappa is to take average the two cells used to determine agreement. This method is termed the prevalence-adjusted bias-adjusted kappa (PABAK; Sim& Wright, 2005). This method is not a perfect solution since it is not representing the actual scores from the researchers and should be a supplemental source of agreement in addition to Cohen's kappa. Table 7 illustrates how data from checklist item number 18 was calculated using PABAK. Instead of the original, uninterpretable Cohen's kappa of zero, the adjusted PABAK is 0.86.

Table 7

Example Data for Calculating a PABAK Score

		Researcher 1		
		Evidence	No Evidence	Total
	Evidence	6.5	1	7.5
Researcher 2	No Evidence	0	6.5	6.5
	Total	6.5	7.5	14

Using the adjusted list of Cohen's kappa scores (kappa and PABAK results), inter-rater reliability for the ROI checklist in this study yielded a range between fair agreement of 0.39 and perfect agreement 1.0. The average of kappa results was considered substantial agreement at 0.64 and the median kappa was similar at 0.63. The most common kappa, or mode, was moderate at 0.46.

Research Question 1

The first research questions was: Do state regulation and guidance documents include recommended practices for monitoring student progress? Out of the 24 items included on the checklist, one checklist item (item 21) was not included in the answer to this research question because it was used to determine if evidence existed for mandating an RTI process rather than the inclusion of a recommended practice. Four checklist items were interpreted as the number of state documents that specified procedures rather than the inclusion of recommended practices. Those items referred to the number of data points gathered for making instructional decisions (item 15), the magnitude of deficiency in level (item 22), and whether procedures were specified for SLD category (item 23) or grade level (item 24). Organized by each recommended practice, Table 8 describes the number of states that included a recommended practice in regulations only, both

regulations and guidance, guidance only, or neither types of documents. Some states had more than one guidance document. If evidence for a recommended practice was included in at least one regulations or guidance document, that state was scored as "evidence found." Items 7 and 10 allowed for a score of "disqualifying content." The recommended practice of having assessment examples (item 7) was scored as having disqualifying content for four states (Texas regulations; Minnesota, North Dakota, and New Jersey guidance). There were no states scored as having disqualifying content for the recommended practice of monitoring student progress on grade level (item 10).

The first level of analysis addressed which practices were most frequently required by states (i.e., which items appeared in regulations). Evidence was found in regulations (as reflected in either column two or three in Table 8) most often for the recommended practices of monitoring progress (item 1), using direct measures of academic skills (item 3), and establishing a norm group (item 20), with inclusion in 17, nine, and seven sets of regulatory and guidance documents respectively.

The second level of analysis addressed which practices were most frequently mentioned, but not necessarily required by states (i.e., which items appeared in either regulations or guidance documents). A review of specific items included in either or both regulations and guidance documents indicated that monitoring progress (item 1) was included most often an in a total of 49 states. The second most commonly found recommended practice was the use of CBM (item 5) in 38 states followed by 36 states including both the use of direct measures of academic skills (item 3) and the frequency of monitoring at Tier 1 (item 11). The recommended practices included least often were considering outlier data (item 18) which was found in only two states, specifications by

Table 8

Number of States That Included a Recommended Practice in Either Regulations Only, Guidance Only, Both Regulations and Guidance, or Neither Types of State Documents

Recommended Practice	Regulations Only	Regulations and Guidance	Guidance Only	Neither
1. Progress Monitoring	1	16	32	2
2. Alternate forms	2	0	13	36
3. Direct measures	0	9	27	15
4. Brief measures	0	1	24	26
5. Curriculum-based measurement	0	6	32	13
6. Computer adaptive tests	0	0	5	46
7. Assessment examples	1	2	32	16
8. Technically adequate measures	0	4	28	19
9. Sensitive to growth	0	0	24	27
10. Grade-level monitoring	4	1	20	26
11. Tier 1 monitoring	0	4	32	15
12. Tier 2 monitoring	1	1	26	23
13. Tier 3 monitoring	0	1	26	24
14. Stable trend line	0	1	17	33
15. Data Points	1	1	22	27
16. Fidelity of assessment administration	0	0	20	31
17. Visual inspection, graph	0	1	14	36
18. Consider outlier data	0	0	2	49
19. Least squares, linear regression	0	0	3	48
20. Comparison, norm group	3	4	15	29
22. Magnitude of deficiency	1	4	5	41
23. SLD type	1	1	4	45
24. Grade levels	0	2	0	49

grade level (item 24) in only two states, and the use of linear regression to establish a trend line (item 19) in only three states.

A review of which recommended practices in regulations revealed that not many are required. Progress monitoring (item 1) was most often required with evidence found in 17

states followed by nine states requiring the direct measurement of academic skills (item 3). Establishing a comparison group (item 20) was required in seven states but the remaining recommended practices were scarcely found to be required. In fact, there were several recommended practices that were not required at all such as using CATs (item 6), using measures that are sensitive to growth (item 9), administering assessments with fidelity (item 16), considering outlier data (item 18), and using linear regression to establish a trend line (item 19).

Evidence was found most often in only regulation documents for the recommended practices of monitoring student progress on grade level (item 10), establishing a comparison or norm group (item 20), and using alternate forms of assessment (item 2) with frequencies of four, three, and two, respectively. In general, there were very few recommended practices that were included in only regulations and it appears that most states are likely to provide further details in guidance documents.

Although evidence was scant for required recommended practices, more evidence existed in guidance documents. Evidence was found in guidance documents for 48 states for monitoring student progress (item 1), seconded by 38 states guiding schools to use CBM (item 5). Recommendations in guidance documents existed in 36 states each for two recommended practices: that schools use direct measures of academic skills (item 3) and monitor student progress at Tier 1 (item 11).

Procedures specified by grade level (item 24) and the recommended practice of considering outlier data (item 18) were absent from regulations and guidance documents for 49 of the 51 states. The recommended practice of using linear regression to establish a trend line (item 19) was absent from regulation and guidance documents for 48 out of the

51 states. Overall, states are not requiring many of the recommended practices nor do they outline specific requirements in regulations. Recommended practices tended to be located in guidance documents or not at all.

Research Question 2

The second research question was: What are the differences, if any, for including recommended practices of monitoring student progress for states that mandate RTI versus states that do not mandate an RTI process for identifying SLD eligibility? The first step to reporting results for this question was to determine which states require an RTI process for SLD eligibility. Item number 21 on the ROI checklist was used to gather this information. Some states required an RTI process in addition to other procedures, and were coded as "evidence found" on the checklist. Table 9 identifies the 14 states that mandate an RTI process and the 37 states that do not. Note that federal regulations have allowed for an RTI process since 2006 with the reauthorization of IDEA but this table depicts which states have further required such a process.

Table 10 displays evidence found per recommended practice as numbers and percentages in states that mandate and RTI process and states that do not. Results are

Table 9

List of States That Mandate and Do Not Mandate an RTI Process for SLD Eligibility

Mandated	Not Mandated
CO, CT, DE, FL, GA, ID, IL, LA, ME, NM, NY, RI, WI, WV	AK, AL, AR, AZ, CA, DC, HI, IA, IN, KS, KY, MA, MD, MI, MN, MO, MS, MT, NC, ND, NE, NH, NJ, NV, OH, OK, OR, PA, SC, SD, TN, TX, UT, VA, VT, WA, WY

reported for evidence found in state regulations only, both regulations and guidance, guidance only, or neither types of documents. Out of the 24 items included on the checklist,

one checklist item (item 21) was not included in the answer to this research question because it was used to determine if evidence existed for mandating an RTI process rather than the inclusion of a recommended practice. Four checklist items were interpreted as the number and percentage of state documents that specified procedures rather than the inclusion of recommended practices. Those items referred to the number of data points gathered for making instructional decisions (item 15), the magnitude of deficiency in level (item 22), and whether procedures were specified for SLD category (item 23) or grade level (item 24). Items 7 and 10 allowed for a score of "disqualifying content." The recommended practice of having assessment examples (item 7) was scored as having disqualifying content for four states (Texas regulations; Minnesota, North Dakota, and New Jersey guidance). There were no states scored as having disqualifying content for the recommended practice of monitoring student progress on grade level (item 10).

A review of individual recommended practices revealed several patterns. Some recommended practices are included more often in documents from states that mandate an RTI process than states that do not. Some recommended practices are included in similar percentages of state documents regardless of whether RTI is mandated, and some recommended practices are scarcely included at all.

Evidence for the inclusion of monitoring student progress (item 1) at least at the guidance level was found in all states that mandate an RTI process and in 95% of documents from states that do not mandate RTI. While monitoring progress is required in half of states that mandate RTI, it is also required in 27% of states that do not mandate RTI, indicating that it is seen as an important practice by most states regardless of whether RTI is required.

Table 10

A Comparison of the Number and Percentage of Evidence Found for Each Recommended Practice by States That Mandate and States That Do Not Mandate RTI

	Regulations Only		Regulations and Guidance		Guidance Only		Neither	
Recommended Practice	*n*	%	*n*	%	*n*	%	*n*	%
1. Progress Monitoring								
Mandated	0	0%	7	50%	7	50%	0	0%
Not Mandated	1	3%	9	24%	25	68%	2	5%
2. Alternate forms								
Mandated	2	14%	0	0%	3	21%	9	64%
Not Mandated	0	0%	0	0%	10	27%	27	73%
3. Direct measures								
Mandated	0	0%	4	29%	10	71%	0	0%
Not Mandated	0	0%	5	14%	17	46%	15	40%
4. Brief measures								
Mandated	0	0%	1	7%	9	64%	4	29%
Not Mandated	0	0%	0	0%	15	41%	22	59%
5. Curriculum-based measurement								
Mandated	0	0%	2	14%	11	79%	1	7%
Not Mandated	0	0%	4	11%	21	57%	12	32%
6. Computer adaptive tests								
Mandated	0	0%	0	0%	2	14%	12	86%
Not Mandated	0	0%	0	0%	3	8%	34	92%
7. Assessment examples								
Mandated	1	7%	0	0%	11	79%	2	14%
Not Mandated	0	0%	2	5%	21	57%	14	38%
8. Technically adequate measures								
Mandated	0	0%	1	7%	11	79%	2	14%
Not Mandated	1	3%	3	8%	17	46%	17	46%
9. Sensitive to growth								
Mandated	0	0%	0	0%	7	50%	7	50%
Not Mandated	0	0%	0	0%	17	46%	20	54%

Table 10 (continued)

Recommended Practice	Regulations Only		Regulations and Guidance		Guidance Only		Neither	
	n	%	*n*	%	*n*	%	*n*	%
10. Grade-level monitoring								
Mandated	3	21%	1	7%	6	43%	4	29%
Not Mandated	1	3%	0	0%	17	46%	22	59%
11. Tier 1 monitoring								
Mandated	0	0%	2	14%	10	71%	2	14%
Not Mandated	0	0%	2	5%	22	59%	13	35%
12. Tier 2 monitoring								
Mandated	1	7%	1	7%	6	43%	6	43%
Not Mandated	0	0%	0	0%	20	54%	17	46%
13. Tier 3 monitoring								
Mandated	0	0%	1	7%	6	43%	7	50%
Not Mandated	0	0%	0	0%	20	54%	17	46%
14. Stable trend line								
Mandated	0	0%	0	0%	8	57%	6	43%
Not Mandated	0	0%	1	2%	9	24%	27	73%
15. Data Points								
Mandated	1	7%	0	0%	8	57%	5	36%
Not Mandated	0	0%	1	2%	14	38%	22	59%
16. Fidelity of assessment administration								
Mandated	0	0%	0	0%	6	43%	8	57%
Not Mandated	0	0%	0	0%	14	38%	23	62%
17. Visual inspection, graph								
Mandated	0	0%	1	7%	5	36%	8	57%
Not Mandated	0	0%	0	0%	9	24%	28	76%
18. Consider outlier data								
Mandated	0	0%	0	0%	1	7%	13	93%
Not Mandated	0	0%	0	0%	1	2%	36	97%
19. Least squares, linear regression								
Mandated	0	0%	0	0%	1	7%	13	93%
Not Mandated	0	0%	0	0%	2	5%	35	95%

Table 10 (continued)

Recommended Practice	Regulations Only		Regulations and Guidance		Guidance Only		Neither	
	n	%	*n*	%	*n*	%	*n*	%
20. Comparison, norm group								
Mandated	1	7%	3	21%	4	29%	6	43%
Not Mandated	2	5%	1	2%	11	30%	23	62%
22. Magnitude of deficiency								
Mandated	1	7%	2	14%	1	7%	10	71%
Not Mandated	0	0%	4	11%	2	5%	31	84%
23. SLD type								
Mandated	1	7%	1	7%	2	14%	10	71%
Not Mandated	0	0%	0	0%	2	5%	35	95%
24. Grade levels								
Mandated	0	0%	2	14%	0	0%	12	86%
Not Mandated	0	0%	0	0%	0	0%	37	100%

Note: Mandated (*n*=14), Not Mandated (*n*=37)

The use of alternate forms (item 2) was not included in many state requirements regardless of whether the state mandates or does not mandate RTI (only two mandated states require it). This practice was mentioned in guidance only documents in 21% of states that mandate RTI and 27% of states that do not. These results indicate that this practice has not been found to be important by states when determining student ROI.

Evidence for the use of direct measures for monitoring academic skills (item 3) was found in all states that require RTI (four states require it and 10 states make recommendations guidance documents) compared to 60% of states that do not mandate RTI including this practice, with the great majority of these states mentioning it only in guidance documents.

The use of brief measures (item 4) was evident in 71% of states that mandate RTI, but only in one state's regulations. There were no states that required this practice in

states that do no mandate RTI, but 15 did suggest using brief measures in guidance documents.

The use of CBM (item 5) was found in the majority (93%) of states that mandate RTI, but was required in only two of these states (14%). Similarly, evidence for a requirement of CBM was found in four states (11%) that do not mandate RTI but was not included as often in guidance documents (57%) compared to states that mandate RTI (79%). In contrast, the use of CATs (item 6) was far less prevalent in any states, regardless of whether RTI was mandated or not. There were no states in which evidence was found for the requirement of CATs. Only two states that mandate RTI and three states that do not mandate RTI included the use of CATs in guidance documents. The use of CBM is clearly more evident in state regulations and guidance documents than CATs.

The majority of states that mandate RTI (86%) included requirements or recommendations citing appropriate examples of assessments for monitoring student progress (item 7) compared to states that do not mandate RTI (62%). Only one state that mandated RTI and two states that do not mandate RTI demonstrated evidence of assessment examples in regulatory documents.

Most states were found to have evidence recommending school teams to use technically adequate measures (item 8) and measures sensitive to growth (item 9) in guidance documents only. One state that mandated RTI and four states that do not mandate RTI required the use of technically adequate measures (item 8) and no states, regardless of mandating RTI required measures sensitive to small changes in growth (item 9).

Only four states that mandate an RTI process and one state that does not required that school teams monitor students on grade level (item 10). There was roughly equivalent evidence found for states that mandate RTI (43%) and do not (46%) in guidance documents; however, there was a higher percentage overall of evidence for this practice across states that mandate RTI (71%) versus states that do not (41%).

In terms of evidence found for the frequency of monitoring progress at Tiers 1, 2, and 3, there were a few noticeable trends. Two states that mandate RTI and two states that do not mandate RTI required an appropriate frequency for monitoring student progress at Tier 1 (item 11). More evidence was found for states that mandate RTI in guidance documents (71%) compared to states that do not (59%) for Tier 1 monitoring, which therefore indicates that states that mandate RTI had more evidence overall (86%) compared to states that do not mandate RTI (65%). Overall, less evidence was found for monitoring student progress at Tier 2 (item 12) with only two states that mandate RTI requiring this practice and no requirements found for states that do not mandate RTI. If evidence existed for monitoring at Tier 2, it was included in guidance documents, but there was not a large difference between states that mandate RTI (57%) and states that do not (54%). Even less evidence was found overall for the inclusion of monitoring at the Tier 3 level (item 13). In state documents that mandate RTI, only one was found to include Tier 3 monitoring requirements and six states provided guidance, for a total of half of states demonstrating evidence. Just over half (54%) of states that do not mandate RTI included this practice, though no evidence was found in regulatory documents, only in guidance. Evidence was found largely for monitoring at Tier 1 and faded with Tiers 2 and 3.

There was no evidence found for requiring a sufficient number of data points to establish a trend line (item 14) in states that mandate RTI, although one state that does not mandate RTI detailed specific requirements in regulations. Evidence in guidance documents was more prevalent in states that mandate RTI (57%) for this practice compared to states that do not (26%).

Only one state that mandated RTI and one state that did not mandate RTI specified a number or range of data points for making instructional decisions (item 15). A higher percentage of guidance documents from states that mandate RTI included specifications (57%) compared to states that do not mandate RTI (40%). There were no states overall that required fidelity of assessment administration (item 16) and evidence was found for a similar percentage of states that require RTI (43%) and do not require RTI (38%).

One state that mandated RTI required a visual inspection to determine if a linear trend line appropriately described the data (item 17) but there was no evidence found requiring this practice in states that do no mandate RTI. If evidence was found, it was generally in guidance documents with a higher percentage found for states that mandate RTI (43%) compared to states that do not (24%).

There were no states that required consideration of outlier data (item 18). One state that mandated RTI and one state that did not mandate RTI that provided evidence in guidance documents. Similarly, no states required the use of linear regression when establishing a trend line to illustrate student progress (item 19). Only three states, one that mandated RTI and two that do not, included recommendations for this practice in guidance documents. Evidence for establishing a norm or comparison group (item 20) was found in four states that mandate RTI compared to three states that do not. However, more state

guidance documents where RTI was mandated included evidence (50%) compared to states that do not (32%).

Many states did not require or recommend specifications for a magnitude of deficiency (item 22), procedures for SLD categories (item 23), or procedures for specific grade levels (item 24). Three states that mandate RTI and four states that do not required a specific magnitude of deficiency and only slightly more states that mandate RTI included specifications overall (29%) compared to states that do not (16%). Two states required procedures for identifying students by SLD categories and both were states that mandate an RTI process. Four states (28%) that mandate RTI demonstrated evidence of procedures for SLD categories in guidance documents, which is a higher percentage compared to states that do not mandate RTI (5%). There was no evidence found for procedures by grade level for states that do not mandate RTI and only two states (14%) out of the 14 that mandate RTI required and recommended procedures by grade level.

Progress monitoring (item 1) was included in all states that mandate an RTI process and 95% of states that do not mandate RTI. All states that mandate an RTI process included the recommended practice of using direct measures of academic skills (item 3) whereas only 60% of states that do not mandate RTI included this item. The majority of states that mandate RTI (93%) required or recommended the use of CBM (item 5) compared to 68% of states that do not mandate RTI. There was one instance where there were more states that included a recommended practice that do not mandate RTI compared to states that do mandate RTI. The recommended practice of monitoring progress at Tier 3 (item 13) was found in 54% of states that do not mandate RTI versus 50% of states that do.

Several items were significantly absent from both states that mandate and do not mandate RTI. Most states did not require or recommend that school teams use CATs (item 6), that school teams consider outlier data points (item 18), a specific a magnitude of deficiency (item 22), specific procedure by SLD category (item 23), or specific procedures by grade level (item 24). However, these items were still included in a higher percentage of states that mandate RTI versus states that do not.

Each recommended practice absent from both regulation and guidance documents were from states that do not mandate an RTI process. Overall, there were significantly more recommended practices included and in states that mandate an RTI process compared to states that do not and most recommendations were found in guidance documents or not at all.

Research Question 3

The third research question was: What are states requiring or recommending in terms of the number of data points needed to make instructional changes? Hixson et al. (2008) reported that instructional change can be made confidently with very few data points. Zirkel and Thomas (2010b) reported various suggestions or requirements for the number of data points used for instructional decision-making. Item 15 on the checklist referred to the number of data points states require or recommend using before making an instructional decision or change. Table 11 illustrates the number or range of data points suggested or required for states. There were sometimes multiple guidance documents per state with one document specifying a number or range of data points and another document not specifying any. Out of 51 states, 23 specified a number or range of student response data points recommended for making an instructional change in guidance

Table 11

States by Guidance and Regulation That Specified a Number or Range of Data Points for Making Instructional Decisions

Data Points	State
Guidance	
3	AL, CT, SC
3-4	GA, OK, VA
3-6	ME
4	KS (Reading), WA
4-5	NM
4-6	OR
6	CO, KS (Math), MI
6-8	AK, NY, RI
7-9	MS
8	ID, VT, WY
8-12	IN
9	SD
12	MN
Not Specified	AR, AZ, CA, DE, FL, HI, IL, KY, MA, MD, MO, MT, ND, NE, NH, NJ, NM, NV, OH, PA, SC, SD, TN, TX,VT, WI, WV, WY
Regulations	
6	DE
12	MN
Not Specified*	*The remaining 49 states did not specify.

documents, whereas only two states depicted specific requirements in regulations.

Research Question 4

The fourth research question was: Do states have specific, quantifiable requirements or recommendations for defining a magnitude of deficiency in level when a student is performing significantly below specified expectations? Item number 22 on the ROI checklist was an open-ended item to gather information about the magnitude of deficiency in level that states suggest between a student and the comparison group. Only 10 states (14 of the 112 documents) included a specified magnitude of deficiency. Guidance documents from nine states and regulation documents from five states were found to include specified magnitudes of deficiency. More documents from states where an RTI

process was not mandated (six states) included specific information compared to states where an RTI process was mandated (four states). Table 12 outlines the specific magnitudes of deficiency as depicted by states and document type.

Research Question 5

The fifth research question was: Do states specify requirements for identifying students with SLD for any or all of the eight categories of SLD? Item number 23 on the ROI checklist allowed for multiple responses to be checked to gather information about RTI procedures specified by SLD category. Table 13 lists the states and SLD categories for which procedures are stated. Only six of the 51 states specified requirements by SLD categories. Two were regulatory and five were guidance documents. Five of the states that specify criteria by SLD categories require an RTI process for SLD eligibility compared to two states that do not. If states did specify categories of SLD, it was typically either reading, math, or both reading and math. If states did not delimit the categories of SLD, it was assumed that requirements or recommendations were intended for all categories of SLD identification.

Research Question 6

The sixth research question was: Do states specify requirements for identifying students with SLD by specific grade levels? Item number 24 on the ROI checklist allowed for multiple responses to be checked to gather information about RTI procedures specified by grade level. Only two states (five of the 112 documents) specified procedures by grade level. The two states were New York and New Mexico where RTI is mandated for these grade levels only. Table 14 lists the state and grade levels for which procedures are stated. If states did not delimit the grade levels, it was assumed that requirements or recommendations were intended for all grade levels.

Table 12

List of States That Specified a Magnitude of Deficiency by Document Type

State	Magnitude
	Guidance
CO	At or below 12th percentile on national norms
KS	9th percentile for AIMSweb and STAR Math; 19th percentile for easyCBM
MI	At or below 9th percentile
MN	At or below 5th percentile
NM	1.5 standard deviations below grade level peers
PA	At or below 10th percentile
SD	2 of 3 criteria: 1.5 standard deviations from the mean; CBM scores in lowest 10% of child's peer group; Performance is 2 or more grade levels below age or grade level
WA	2 of 4 criteria: CBM scores at or below 7th percentile on grade level or at or below 16th percentile of previous grade level; A standardized assessment score 1.75 standard deviations below the mean, consistent with test protocols; CBM scores and other data with student's median performance below grade-placement peers with a discrepancy ratio of at least 2.0 (the discrepancy ratio is calculated by dividing the peers' median performance by the target student's median performance); Student's instructional performance level is two or more grade levels below current grade placement determined by CBM scores
WI	1.25 standard deviations below the mean
	Regulations
LA	Greater than 1.5 standard deviations below the mean in grades 1 and 2; Greater than 2 standard deviations for grades 3-12 using age norms
MN	At or below 5th percentile
NM	1.5 standard deviations below grade level peers
SD	2 of 3 criteria: 1.5 standard deviations from the mean; CBM scores in lowest 10% of child's peer group; Performance is 2 or more grade levels below age or grade level
WI	1.25 standard deviations below the mean

Table 13

List of States With Specified Procedures by SLD Category

State	SLD Category
Regulations	
DE	Basic Reading Skill, Reading Fluency Skills, Reading Comprehension, Mathematic Calculation, Mathematic Reasoning
NY	Basic Reading Skill, Reading Fluency Skills, Reading Comprehension
Guidance	
CO	Basic Reading Skill, Reading Fluency Skills, Reading Comprehension
CT	Basic Reading Skill, Reading Fluency Skills, Reading Comprehension
KS	Basic Reading Skill, Reading Fluency Skills, Reading Comprehension, Mathematic Calculation, Mathematic Reasoning
NY	Basic Reading Skill, Reading Fluency Skills, Reading Comprehension
OR	Basic Reading Skill, Reading Fluency Skills, Reading Comprehension

Table 14

States With Specified Procedures by Grade Level

State	Grade Level
NM	
Regulations	Kindergarten, first, second, and third grades
Guidance	Kindergarten, first, second, and third grades
NY	
Regulations	Kindergarten, first, second, third, and fourth grades
Guidance	Kindergarten, first, second, third, and fourth grades

Summary

This chapter first presented results for inter-rater reliability between two researchers for the checklist of recommended practices and results for the six research questions. Inter-rater reliability results from having a second researcher review a subset of state documents were presented. Cohen's kappa and PABAK results for calculating inter-rater reliability indicated fair to perfect agreement for checklist items.

Overall, evidence was found for 14 states for requiring an RTI process for SLD eligibility whereas evidence indicated that 37 states did not require RTI.

Results from the first and second research questions indicated that more recommended practices were found in states that mandate RTI versus states that do not, and that recommended practices were generally included in guidance documents rather than regulations.

Results for the third research question indicated that a variety of specifications were described in state documents for the number or range of data points required or recommended for instructional changes or decisions, with most states not having evidence of specifications.

The fourth research question addressed whether states required or recommended a specific, quantifiable magnitude of deficiency in level to consider students significantly below their peers. Results indicated that if states did have requirements, they were not consistent with each other. If states included a specific magnitude of deficiency, they were generally included in guidance documents versus regulations.

Evidence for the fifth research question indicated that the majority of states do not specify requirements or recommendations by SLD category and if they did it was for math and/or reading. Results for the sixth research question addressed whether states specified requirements for determining SLD eligibility by grade level. The few states that included grade-level specifications typically referred to elementary grade levels only, with one state requiring an RTI process for grades kindergarten through third and another state requiring an RTI process for grades kindergarten through fourth.

CHAPTER V

DISCUSSION

With the reauthorization of the IDEA of 2004, state education agencies (SEAs) were given the option to use a response to intervention (RTI) process for identifying students with a specific learning disability (SLD). Given that special education eligibility determination is a high-stakes decision, it is important that the components of such evaluations are aligned with recommended practices in the literature. The goal of this study was to provide an updated snapshot of requirements by state regulation and guidance documents with increased specificity for requirements related to monitoring student progress and determining student rate of improvement (ROI).

The researcher for this study developed a checklist of recommended practices gleaned from the literature and used this tool to review 112 state documents from all 51 SEAs including regulations and guidance. A second researcher was used to establish inter-rater reliability with a subset of state documents.

This chapter discusses the results of this study, examining similarities and differences between the current study and previous research. Statements of limitations to the current study are discussed followed by implications for future research and practice in education and school psychology. Prior to gathering data to describe results for the research questions, inter-rater reliability, a description of state documents, and determination of which states mandate and RTI process is presented.

Inter-Rater Reliability

Inter-rater reliability was established between two researchers for a subset totaling 14 of the 112 state documents from five states, which equates to approximately 10% of all

states. Cohen's kappa was used to calculate agreement in all but three instances, where a prevalence effect did not allow for such calculations. For those items, a PABAK score was calculated per recommendations from Sim and Wright (2005). Agreement ranged from fair to perfect. The item with the least amount of agreement was for monitoring student progress (item 1), which yielded a Cohen's kappa of 0.39 and fell within the fair range according to parameters of interpreting kappa set forth by Landis and Koch (1977; as cited by Hallgren, 2012). The items with perfect agreement included the use of CATs (item 6) and visual inspection of data (item 17). Items with less agreement may have had less clear descriptions or terms used for searching state documents than items with high levels of agreement. To improve the inter-rater reliability of items that were found to have less agreement (i.e., monitoring progress, item 1), future research could have two researchers independently review a sample of documents until a minimum agreement is reached (i.e., Cohen's kappa of .80 or higher). Overall, the procedure for inter-rater reliability worked well and might be recommended for future studies of this type.

State Documents Reviewed

A total of 112 documents were reviewed for this study. A majority of documents were found searching for the titles listed in the Hauerwas et al. (2013) study. Other documents were found by searching state department of education or technical assistance websites. Hauerwas et al. had suggested that during their study, documents were in the process of being revised. Since the Hauerwas et al. study, 59% of state documents were revised, updated, or rewritten, most commonly in the year 2013. Zirkel and Thomas noted in their updated study (2010b) that four states (AR, DC, NJ, and WY) did not provide guidance documents for RTI whereas Hauerwas et al. reported the absence of guidance

documents for five states (DE, HI, NV, NJ, and WY). Hauerwas et al. did not include regulation or guidance documents for the District of Columbia but Zirkel and Thomas and the current study included regulatory documents for District of Columbia. Regulation documents were found for all 51 states. However, the current study included guidance documents for all SEAs except the District of Columbia and North Carolina. One explanation set forth by Zirkel and Thomas and Hauerwas et al. for the differences between studies is that states have continued to reevaluate the need to provide guidance for using an RTI process since the reauthorization of the IDEA. Another explanation is that states use terms other than RTI that may not have been used for document searches in previous studies. For instance, Hawaii's guidance document was entitled *Comprehensive Student Support System*. New Jersey's lengthy guidance document was entitled *Resource Manual for Intervention and Referral Services*. More recent guidance documents have adopted the terms multi-tiered system of supports (MTSS) to describe both academic RTI and positive behavior intervention and supports (i.e., Kansas). If the authors from previous studies searched only for narrow variations of the term RTI, the documents described likely would not have been found.

There is a notable difference in which documents were ultimately analyzed in the current study, the Zirkel and Thomas (2010b) study, and the study by Hauerwas et al. (2013). All three studies reviewed all available documents; however, Hauerwas et al. reviewed only the regulations and guidance documents from 38 states that described an RTI process in more detail than federal regulations for a total of 29 regulations and 27 guidance documents. The difference in inclusion criteria can be attributed to a difference in research questions. Zirkel and Thomas and the current study aimed to systematically

capture sets of information across all states, whereas Hauerwas et al. set out to gather information about states that described an RTI process and related components to uncover themes in prevalent RTI components.

RTI Mandated States

In addition to the documents and terms being updated, a few states updated requirements for using an RTI process for SLD eligibility. The current study found evidence of 14 SEAs mandating an RTI process for determining eligibility for SLDs. Previous studies have found differing numbers of states with this mandate. One of the first studies reviewing changes to state regulations (by survey) after the IDEA regulations were published in 2006 found that six states required an RTI process for SLD eligibility (Ahearn, 2008). Although Hoover et al. (2008) also completed a review of changes to regulations, they did not specify the number of states requiring an RTI process. Zirkel and Krohn (2008) had concluded that at the time of their survey, state directors were leaning toward allowing LEAs to make decisions about eligibility requirements rather than mandate an RTI process at the level of state regulation. Berkeley et al. (2009), though one of the first to directly review information from state websites, did not specify if their review was of regulations or guidance documents but had suggested that as many as 15 states required an RTI process. Zirkel and Thomas (2010a) were the first to attempt to systematically review published documents and differentiate between regulations and guidance documents. Therefore, the researchers were able to conclude that 12 states required an RTI process for determining SLD eligibility. In their research update, Zirkel and Thomas (2010b) reported that 13 states required an RTI process for SLD eligibility but noted that Wisconsin was still in the process of revising their requirements. In the most recent

snapshot of state regulations and guidance documents, Hauerwas et al. (2013) had found requirements for RTI data in 17 states but noted that these states did not solely require an RTI process for special education eligibility.

Further examination of the current study compared to previous studies yielded similarities and differences among findings. Two previous studies (Hauerwas et al., 2013; Zirkel & Thomas, 2010b) provide for the most direct comparisons of which states require or allow an RTI process because of similar methodologies and recent publication dates.

Table 15

A Comparison of Findings for States That Mandate RTI Between the Current Study and Two Previous Studies in Chronological Order

Zirkel & Thomas (2010b)	Hauerwas et al. (2013)	Current Study (2014)
CO	CO	CO
CT	CT	CT
DE	DE (reading, math)	DE (reading, math)
FL	FL	FL
GA	GA	GA
	IA	
ID	ID	ID
IL	IL	IL
LA	LA	LA
ME	ME	ME
	MS	
NM	NM (K-3)	NM (K-3)
NY	NY (K-4, reading)	NY (K-4, reading)
RI	RI	RI
	WI	WI
WV	WV	WV
	WY	
Total *13*	*17*	*14*

Reviewing in chronological order, Table 19 illustrates the similarities and differences in findings for which states require an RTI process for SLD eligibility. The majority of findings are similar across studies; however, there were a few inconsistencies. The Hauerwas et al. study included states that required "RTI data" as part of the eligibility process whereas the Zirkel and Thomas (2010b) study and the current study specified an RTI process as being mandated for SLD eligibility. The difference in qualifying as an RTI-mandated state is most likely in the difference in definitions or criteria set forth by the researchers. Compared to the Hauerwas et al. (2013) study, the current study did not find evidence that an RTI process was mandated for SLD eligibility for Iowa, Mississippi, or Wyoming. A review of Iowa regulations revealed the following statement, which the researcher determined to mean that two possible avenues for identifying SLD were available to school teams in this state:

> a. Requires the use of a process based on the child's response to scientific, research-based intervention *or* the use of other alternative research-based procedures for determining whether a child has a specific learning disability, as defined in subrule 41.50(10); and
>
> b. Prohibits the use of a severe discrepancy between intellectual ability and achievement for determining whether a child is an eligible individual on the basis of a specific learning disability. [*Italics added for emphasis*]. (Iowa Administrative Code, Section 281—41.307(1), 2011, p. 49)

Mississippi adopted criteria set forth by federal regulations, which the researcher for the current study interpreted as allowing an RTI process, rather than requiring one. Mississippi's regulations are defined as follows:

Public agencies—

(1) May use a severe discrepancy between intellectual ability and achievement for determining whether a child has a specific learning disability, as defined in §300.8(c)(10);

(2) May use a process based on the child's response to scientific, research-based intervention; and/or

(3) May use other alternative research-based procedures for determining whether a child has a specific learning disability, as defined in §300.8(c)(10). (Mississippi Department of Education, 2009, p. 152)

Hauerwas et al. (2013) likely interpreted this guidance section in Mississippi's regulations to mean that "RTI data" were required as part of an evaluation to determine SLD eligibility:

PRE-REFERRAL REQUIREMENTS

To ensure that underachievement in a child suspected of having a specific learning disability is not due to a lack of appropriate instruction in reading or math, the multidisciplinary evaluation team must consider, as part of the evaluation:

A. Data that demonstrate that prior to, or as a part of, the referral process, the child was provided appropriate instruction in general education settings, delivered by qualified personnel; and

B. Data-based documentation of repeated assessments of achievement at reasonable intervals, reflecting formal assessment of student progress during instruction, which was provided to the child's parents. (Mississippi Department of Education, 2009, pp. 301-302)

Wyoming's regulations (Wyoming Department of Education, 2010) indicate the following in terms of SLD eligibility: "The group shall use either the Wyoming Severe Discrepancy Formula or a response to intervention process when determining whether a child is not making sufficient progress to meet age or Wyoming grade-level standards" (p. 11). The researcher for the current study interpreted this regulation as permitting an RTI process or severe discrepancy to determine SLD eligibility.

The addition of one state between the Zirkel and Thomas (2010b) study and the current study is due to updated regulations, specifically for Wisconsin. The list of states in Table 19 is defined as states that require (a) solely an RTI process for SLD eligibility only, (b) an RTI process in addition to other requirements (e.g., a pattern of strengths and weaknesses), and (c) an RTI process for only certain grade levels or SLD areas (noted in table). The data from Zirkel and Thomas (2010b) was captured in May 2010, Hauerwas et al. (2013) was a snapshot from October 2011, and the current study reviewed documents in August 2014. Over the course of approximately four years, it appears that very few states have changed their regulation documents to require an RTI process for SLD. The trend since Zirkel and Thomas's study is consistent with their statement that the requirement or inclusion of an RTI process has been predominantly a local decision rather than a state mandate.

Research Question 1

The first research question was: Do state regulation and guidance documents include recommended practices for monitoring student progress? There were very few instances in which recommended practices were included in regulations but not guidance documents. If states included recommended practices, they tended to be in guidance

documents. This finding is consistent with findings from Zirkel and Thomas (2010b), who had reasoned that it is perhaps easier to influence changes in guidance documents, which are not legally-binding, compared to regulations.

The checklist items could be arbitrarily categorized into groups of those found most often (i.e., in 26 or more of the 51 states or roughly 50%), less often (i.e., 13 to 25 states, or roughly 25-49%), and items that were scarce among states (i.e., 0 to 12 states, or less than 25%). Those found in most states included monitoring of student progress (item 1), using CBM (item 5), using direct measures of academic skills (item 3), citing examples of assessments for monitoring progress (item 7), using technically adequate measures (item 8), and frequency of monitoring progress at Tiers 1, 2, and 3 (items 11, 12, and 13, respectively). Evidence found in fewer states (25-49%) included using alternate and equal forms (item 2), using brief measures (item 4), using measures sensitive to growth (item 9), monitoring on a student's grade level (10), ensuring a stable trend line (item 14), specifying a number or range of data points (item 15), administering assessments with fidelity (item 16), graphing student data (item 17), and establishing a comparison or norm group (item 20). There was infrequent evidence found (i.e., in less than 25% of states) for several important recommended practices including the use of CATs (item 6), considering outlier data (item 18), using linear regression to establish a trend line (item 19), and specifying procedures for a magnitude of deficiency in level (item 22), by SLD category (item 23), and by grade level (item 24). Evidence of the latter items was inconsistent across states yet accurately monitoring student progress is essential for high-stakes decisions such as special education eligibility (Ardoin et al., 2013).

Monitoring student progress over time (item 1), beyond the federal requirement of repeated measures for SLD eligibility, was the most commonly found recommended practice (49 of 51). This is likely because of the prevalence of information in the literature that monitoring student progress is a major component for determining a student's RTI (Ardoin et al., 2013; Fuchs, 2004; Stecker, Fuchs, & Fuchs, 2008). The inclusion of evidence for using CBM to monitoring student progress (item 5) was found in 38 states and was the second most commonly found recommended practice. Similar to monitoring student progress, the substantial evidence found for using CBM can be attributed to popularity in the literature for educators and school psychologists (Ardoin et al., 2013; Deno, 1982; Fuchs, 2004; Shinn, 2007). Despite increasing popularity on the NCRTI site (2011) listing tools for monitoring student progress and inclusion in the literature (Kovaleski et al., 2013; Stone & Davey, 2011), CATs were found in only one state that mandates an RTI process and five in states that do not mandate RTI.

Two recommended practices on the checklist included recommended comparisons for student response data. One recommended practice was for student progress to be monitored on the student's grade-level, especially when using CBM and when determining special education eligibility. This item was included in less than half of states, yet is a strong recommendation in the literature (Kovaleski et al., 2013) and is used to determine eligibility for SLD per federal requirements (IDEA, 2004). A second recommended practice regarding comparison groups was that states specify the comparison or norm group to which student data is compared in order to determine whether a student is making sufficient progress. There was no evidence found in over half of the states. To describe the comparison group, more often than not state documents used vague descriptions such as

"compared to peers" or "compared to expectations." This may be purposeful to allow local school districts to choose a more specific comparison group relevant to their population of students, which is consistent with recommendations to use local norms in the literature (Burns & VanDerHeyden, 2010); however, state documents should specify a consistent comparison group for making high-stakes decisions (Hintze & Silberglitt, 2005).

The frequency of monitoring student progress throughout the various tiers of an RTI framework was outlined in three different checklist items. If specifications were included in state documents for frequency of monitoring, the specifications were usually aligned with recommendations in the literature for Tier 1, Tier 2, and Tier 3 (American Institutes for Research, 2011). The American Institutes for Research recommends that students be screened at Tier 1 at least three times per year, monitored at Tier 2 at least once a month, and monitored at Tier 3 at least once a week. More state documents included specifications for Tier 1 (36 states), fewer documents specified frequency for Tier 2 (28 states), and fewer still specified frequency for Tier 3 (27 states). These results are categorized differently for this study compared to Zirkel and Thomas (2010b) but are similar in that more states have specifications for Tier 1, compared to Tiers 2 and 3. Hauerwas et al. (2013) did not review documents for the frequency of Tier 1 monitoring, possible because it is already a widely accepted practice (Gleason, 2013), but documented evidence of frequencies for Tiers 2 and 3. A review of the table in the Hauerwas et al. study (p. 113) indicated that if state documents did include specific frequencies of monitoring, they were aligned with recommendations in existing literature for Tier 2 (monthly or every other week) and Tier 3 (weekly or more often). With 14 states mandating an RTI process, and the remaining 37 allowing for an RTI process, it is concerning that more states do not

have specified frequencies for monitoring student progress at Tiers 2 and 3 that align with suggestions in the literature since these data play a large role in determining special education eligibility for SLDs (Kovaleski et al., 2013).

Several checklist items specified recommended practices for reviewing and interpreting student response data. Although student ROI is important data for determining eligibility for special education (Fuchs & Fuchs, 1998; Gresham, 2001; Kovaleski et al., 2013), these practices were not included in many documents. State documents were searched for evidence of school teams graphing student response data (15 states), determining if a linear trend line is appropriate to describe the data (considering outlier data, 2 states), and if so, using linear regression to establish a trend line (3 states). Only two state documents suggested that teams detail a process for considering outlier data, Minnesota and Wisconsin. It is notable that a prolific author in the fields of education and school psychology for the stability of trend lines for student response data, Theodore Christ, co-authored the state document from Minnesota, his state of residence. Only three total documents (one regulation, two guidance) suggested the use of linear regression for establishing a trend line: Kentucky, Minnesota, and Wisconsin. All three documents were guidance documents, not regulatory. Of all 51 states, only 18 required or recommended that school teams consider the stability of the trend line. This statistic seems inadequate given the increasing amount of literature cautioning educators to consider sources of variability when establishing a trend line or slope of progress (Christ, 2006; Christ & Silberglitt, 2007; Christ, Zopluoglu, Long, & Monaghen, 2012; Ditkowsky, 2009; Hixson, Christ, & Bradley-Johnson, 2008). Although instructional changes can be made confidently with few data points (Ardoin et al., 2013; Hixson et al., 2008), the number of data points

required for a stable trend line is much higher (Christ et al., 2012). Initial studies on this topic suggested that more data points are better (Christ, 2006), then 8-9 data points (Ditkowsky, 2009), and most recently upwards of 18 weekly data points with a technically adequate measurement system (Christ et al., 2012). Because monitoring and documenting student progress is a key component of an RTI process (Fuchs & Fuchs, 1998; Kovaleski et al., 2013; VanDerHeyden & Burns, 2010), it is imperative that more than the 18 of states currently making suggestions for stable trend lines need to include this recommended practice.

Overall, the more specific recommended practices that influence the quality of data used to determine eligibility were not present in the majority of states. The inconsistent inclusion of recommended practices is on par with past studies that have reviewed changes to regulations and state guidance documents either by surveying state directors of special education (Ahearn, 2008; Hoover, Baca, Wexler-Love, & Saenz, 2008; Zirkel & Krohn, 2008) or directly reviewing state websites or published documents (Berkeley, Bender, Peaster, & Saunders, 2009; Hauerwas et al., 2013; Zirkel, 2011; Zirkel & Thomas, 2010a; Zirkel & Thomas, 2010b).

Research Question 2

The second research question was: What are the differences, if any, for including recommended practices of monitoring student progress for states that mandate RTI versus states that do not mandate an RTI process for identifying SLD eligibility? The percentage of evidence for recommended practices found in RTI mandated was higher than in states that do not mandate RTI except for the frequency of monitoring at Tier 3 (item 13). The difference in this item was only 4%, however. Monitoring student progress (item 1) and

using direct measures of academic skills (item 3) were found in all states that mandate RTI. The high percentage of evidence found for these items can be attributed to the prevalence in the literature of monitoring student academic skills over time (Fuchs & Fuchs, 1998; Kovaleski et al., 2013; VanDerHeyden & Burns, 2010).The majority of states that mandate RTI (93%) required or recommended the use of CBM (item 5) compared to 68% of states that do not mandate RTI, which is in line with the literature that suggests CBM is an adequate tool for monitoring student response to instruction and intervention (Ardoin et al., Fuchs, 2004, NCRTI, 2011; Shinn, 2007).

Several items were significantly absent from both states that mandate and do not mandate RTI. Most states did not require or recommend that school teams use CATs (item 6), that school teams consider outlier data points (item 18), a specific a magnitude of deficiency (item 22), specific procedure by SLD category (item 23), or specific procedures by grade level (item 24). However, these items were still included in a higher percentage of states that mandate RTI versus states that do not. A reasonable hypothesis that can be developed from these findings is that states that mandate RTI tend to include more recommended practices from the literature for monitoring student progress than states that do not mandate RTI.

Research Question 3

The third research question was: What are states requiring or recommending in terms of the number of data points needed to make instructional changes? Hixson et al. (2008) reported that instructional change can be made confidently with very few data points. Zirkel and Thomas (2010b) reported various suggestions or requirements for the number of data points used for instructional decision-making. A direct comparison can be

made between the current study and the Hauerwas et al. (2013) study in terms of the number of data points states recommend or require for making an instructional change. Hauerwas et al. reported that the most commonly cited range of data points recommended was 3 to 4. The current study found that the most common numbers and ranges cited were 3, 3-4, 6, 6-8, and 8. These numbers and ranges were each found for three states. Although 23 states specified a number or range of data points for making instructional changes, only two states demonstrated evidence in regulations: Delaware (6 data points) and Minnesota (12 data points). Hixson et al. (2008) stated that instructional change can be made with confidence based on only a few data points; however, when making high-stakes decisions such as special education eligibility, far more data points are required to account for standard error of measurement associated with extraneous variables such as assessment probe variations (Christ & Silberglitt, 2007), testing conditions (Hixson et al., 2008), and the expected level of growth based on initial student performance (Fien et al., 2010). The Hauerwas et al. study did not list all findings for the number or ranges of data points found, only the general trend. It is possible that more state documents have included language regarding the number or range of data points used to make instructional changes since the Hauerwas et al. study since, as stated previously, the majority of documents have been updated since their study and this may account for the difference between findings from their study and the current results.

Research Question 4

The fourth research question was: Do states have specific, quantifiable requirements or recommendations for defining a magnitude of deficiency in level when a student is performing significantly below specified expectations? Item number 22 on the

checklist was used to capture information to determine if states are recommending or requiring specific magnitudes of deficiency for comparing individual student response data to a norm or comparison group. Federal regulations require that school teams document that "the child does not make sufficient progress to meet age or to meet State-approved grade-level standards" (Federal Register, 2006, p. 46544,) but does not define "sufficient progress" or lack thereof. A majority of state documents did not quantify or define a magnitude of deficiency in level of performance. Only 10 states included specific criteria. Of the states that defined a magnitude of deficiency, there were various requirements or recommendations ranging from specific percentile ranks (e.g., at or below the 12th percentile), or a specific standard deviations from a norm group (e.g., 1.25 standard deviations from the mean), or meeting two criteria out of three or four possible criteria. Kansas even specified requirements specific to commercially available assessments (i.e., 19th percentile for easyCBM). The range of requirements is concerning and leads to inconsistency in eligibility across states, which is an echo of past research regarding the consistency of eligibility using the ability-achievement discrepancy model of SLD identification (Bocian et al., 1999; Ysseldyke, Gresham, & Bocian, 1983).

Research Question 5

The fifth research question was: Do states specify requirements for identifying students with SLD by any or all of the eight categories of SLD? Because there are eight categories of SLDs, the researcher was interested if state regulations or guidance documents specified criteria by SLD category. The majority of documents did not specify if requirements were for only certain SLD category. If an SLD category was not specified, it was inferred that the state criteria was intended for all category of SLDs. For instance, New

York requires an RTI process for SLD eligibility only for reading. For the purpose of this study, this was interpreted to include the three SLD categories that relate to reading including basic reading skill, reading fluency skills, and reading comprehension. Only six states discussed SLD categories. Two documents were regulatory and five were guidance documents; one state, New York, specified requirements in both guidance and regulations. Five of the states that specify criteria by SLD category require an RTI process for SLD eligibility compared to two states that do not. If states did specify categories of SLD, it was typically either reading, math, or both reading and math. Consistent with a previous study (Hauerwas et al., 2013), Delaware and New York were the only two states that required an RTI process for a specific academic area, and both were for reading only. Five guidance documents (Colorado, Connecticut, Kansas, New York, and Oregon) specified criteria by SLD categories and all were for reading except Kansas, which included mathematics (i.e., SLD types would be mathematic calculation and mathematic problem-solving). According to these findings, the states that specify requirements by academic area or SLD categories have remained consistent since the Hauerwas et al. study, which had reported on data from October of 2011. Essentially, there have been no additional states that specify criteria by academic area or SLD categories for the past three years.

Research Question 6

The sixth research question was: Do states specify requirements for identifying students with SLD by specific grade levels? The researcher was interested to determine if states differentiated requirements or recommendations by specific grade levels. If a state did not specify requirements by grade level, it was inferred that requirements for the state were intended for all grade levels. One state, Arizona, differentiated guidance documents

by elementary and secondary but did not specify what grade levels to attribute to the documents. Only two states, New Mexico and New York, specified that an RTI process was mandated for specific grade levels. For both states, specific grade levels were indicated in both regulatory and guidance documents. Consistent with grade-level specifications in Hauerwas et al., New York mandated an RTI process for grades kindergarten through fourth and New Mexico mandated an RTI process for grades kindergarten through third. According to these findings, the states that specify requirements by academic area or SLD category have remained consistent since data was collected in 2011 for the Hauerwas et al. (2013) study.

Limitations

Despite a comprehensive review of state regulation and guidance documents using a checklist developed from an extensive review of existing literature, there are limitations to the current study. One limitation parallels concerns of previous researchers who have completed systematic reviews of state documents. With evidence of 59% of state documents having been updated since the Hauerwas et al. study, which included documents in October 2011, it is noted that states are continuously updating and changing regulations and recommendations. The data presented in this study are limited to the documents reviewed in August 2014 and it is assumed that the researcher found the most up-to-date and accurate state documents using the search methods described in the third chapter. Another assumption is that school teams have access to these same documents and use them for implementing an RTI framework and SLD eligibility. It is possible that LEAs use additional recommended practices than those required or provided in guidance documents at the state level. Zirkel and Thomas (2010b) alluded to this possibility in their

study by highlighting which states required an application process in order to use an RTI process for SLD eligibility. Hence, while a state may allow multiple methods for SLD eligibility, the state requires careful consideration of proof of RTI implementation before approving an LEA to use an RTI process and it is possible that those LEAs that have been approved would incorporate more recommended practices than what the state requires. For instance, Pennsylvania requires approval of an application to use an RTI process but does not require an RTI process for SLD eligibility. Consequently, a closer look at individual LEA use of recommended practices would be more indicative of practices used by LEAs in the state.

A second limitation is the use of a newly-developed research tool. In order to document the reliability of the research tool (i.e., checklist) used for this study, inter-rater reliability was calculated between the two researchers' scores for a randomized subset of approximately 10% of all states was completed. The use of a randomly chosen subset of documents was for calculating inter-rater reliability to reduce the chance that the first researcher would choose documents that were more easily scored using the research tool. Inter-rater reliability results for Cohen's kappa and PABAK on individual checklist items ranged from fair to perfect. Inter-rater reliability results indicate that some checklist items were more explicitly defined than others, which may influence the results of the current study. Both researchers used the search procedures defined in Table 2 to allow for consistent methods of reviewing state documents; however, the results from the study were based on the first researcher's review of all 112 state documents and there is a possibility for researcher-bias. In addition, both researchers have extensive background knowledge in the implementation of RTI and recommended practices for determining ROI.

Should the checklist and search procedures in Table 2 be used by other researchers without such background knowledge, the results could vary, indicating that the checklist items and search procedures were not well-defined for general use.

Implications for Research

A checklist of recommended practices for monitoring student progress and determining student ROI did not exist at the time of this study; therefore, one was developed based on a thorough review of existing literature. Additional research is needed to establish reliability and validity of the checklist since results of inter-rater reliability indicated that some items were better defined than others. Also, not all of the practices in the literature review are based on significant empirical evidence and many are practices that have been set forth by experts in the field. Therefore, more research is needed to determine if the list of recommended practices is an accurate list of practices for monitoring student progress and determining student ROI. An example is that CBM is widely used for monitoring student progress (Fuchs, 2004) and was found to be frequently mentioned in state documents in this study, but was not designed for nor has the technical adequacy required to make high-stakes decisions such as special education eligibility in isolation from other data or procedures (Ardoin et al., 2013). Research is needed to determine the benefit of states requiring or recommending all items on the checklist. For instance, if states were to require all of the recommended practices, would it lead to better or more consistent instructional decisions and outcomes for students?

One of the assumptions of the current study is that the documents reviewed are the same documents that school teams within LEAs use to implement an RTI framework for SLD eligibility. Research is needed to determine if LEAs have access to and regularly refer

to the documents used in this study. Similarly, since the majority of states have allowed for, rather than require, an RTI process and are therefore allowing LEAs to make decisions for eligibility requirements, research is needed to determine if this flexibility between LEAs within a state leads to better decisions-making in regards to SLD identification as well as better outcomes for students or if the inconsistency in state requirements is problematic as is has been historically for the ability-achievement discrepancy. Future research could address whether there are meaningful differences in state documents that are authored solely by the SEA or through a collaboration of groups such as the SEA, the SEA's technical assistance network, or other groups.

A problem with the findings for the first research question in this study is that disparate inclusion of recommended practices, which have been indicated in the literature as meaningful practices for determining student ROI, perpetuates inconsistent interpretation of student response data. Furthermore, it is problematic that the majority of recommended practices are found within guidance documents, and are not legally binding, adding to the range of instructional decisions possible from student response data. Research is needed to determine the extent to which LEAs adhere to guidance documents to determine if these non-legally binding documents influence better and more consistent data-based decision making.

Research that expands on the various magnitudes of deficiency set forth by states is needed to determine if different numbers or profiles of students are identified as having SLD, similar to previous studies of SLD prevalence with various criteria for the ability-achievement discrepancy (Bocian et al., 1999; Ysseldyke, Gresham, & Bocian, 1983).

Similarly, the effect of instructional decisions should be empirically explored based on various number and ranges of student progress data points.

Additional research is needed to determine if recommended practices should be differentiated between grade levels since there may be different grade-level requirements and assessments/norms used. In addition, studies at the pre-school level may be beneficial. A similar research need is to continue to review requirements by SLD category since some (e.g., listening comprehension) have not been as widely researched and discussed in the literature as others (i.e., reading). As indicated in the Zirkel and Thomas (2010b) study, more information is also needed for determining SLD eligibility for diverse learners such as English language learners.

Studies comparing SLD eligibility decisions made with a consistent set of norms (i.e., national) versus decisions made with varying norms (local grade-level norms for one school, end-of-year benchmark comparisons for another) would be beneficial to determine if there is a best set of norms for comparison or if it is best to let school teams determine which norms make sense for their population of students.

Not only can procedures differ by state, but the fields of education and school psychology may grapple with the long-standing definition of SLD. In the federal definition, SLD is a neurologically-based disorder, yet there is room to identify students with SLD without addressing neurological bases at all when using an RTI model. Therefore, research is needed to determine how to improve the assessment of SLD. Literature in the fields of education and school psychology have begun to tackle this question with criteria such as dual-discrepancy (Fuchs & Fuchs, 1998) or determining that, based on a student's established ROI, he or she will not progress to the expected level "within a reasonable

amount of time" (p. 150). Research is needed to determine if there are best practices for identifying students with SLDs using various criteria within an RTI process.

Implications for the Practice

Educators at the federal and state level need to determine if flexibility of criteria for SLD is helpful or harmful. After several decades of the ability-achievement discrepancy model for identifying SLD, it is apparent that lack of specific requirements has likely contributed to variations in the numbers of students identified as eligible for special education (Bocian et al., 1999; Ysseldyke, Gresham, & Bocian, 1983). Furthermore, loosely defined criteria allow for differences in allocation of special education resources between states and even between neighboring school districts (Bocian, Beebe, MacMillan, &Gresham, 1999). Now that an RTI process is allowed, it is important for educators to consider the benefits of clearly defined and consistent procedures for identifying SLD.

Educators also need to continue to stay current with recommended and best practices found in the literature. In general, the current study noted that guidelines included more recommended practices or details for implementing RTI than regulations, likely because guidelines are easier to change and are not legally binding. This finding is consistent with the findings in Zirkel and Thomas (2010b). Similarly, many states have been lenient in their requirement for using an RTI process, allowing LEAs to make final decisions. This may be intentional since there are aspects of implementing RTI and using such a process for SLD that are not well-researched to date. In that case, it is wise to wait for further clarification from research; however, in the meantime school teams still need to make the best possible decisions based on available information.

Not only should educators remain current with empirical findings and recommendations in the literature, but SEAs need to be aware of changes in the research literature as well. One way for SEAs to stay current is to align with national and state associations when constructing regulations and guidance documents to gather input from those working in the field and conducting research.

For both the fields of education and school psychology, training institutions need to stay current with best practices in the literature for working in teams to monitor and interpret student response data. Educators, and especially school psychologists, need to stay current with changes in regulations and guidance documents set forth by the state to ensure their instructional decisions are aligned with state requirements and recommendations. School psychologists may require updated training for identifying SLDs and documenting data that align with regulations and recommendations in their respective state.

School psychologists, intervention specialists, administrators, and teachers can serve as leaders and facilitators of teams implementing RTI with their training in assessment, teaming, and data-analysis (National Association of School Psychologists, 2006). In a leadership position, they can help keep school teams informed of changes in the literature for best practices for monitoring student progress and interpreting ROI, not just for SLD eligibility but for all instructional decisions. School psychologists may even be able to influence local requirements for their school building or district. On a broader scale, school psychologists and their national association could be advocates for change in the regulation and guidance provided by states. In addition to being responsible, savvy consumers of current research, school psychologists can conduct research to help improve

practices related to monitoring student progress and determining ROI, contribute to existing literature, and influence practices on a larger scale.

Conclusions

Several years following the reauthorization of IDEA has evidenced changes in state regulations in terms of requiring or allowing an RTI process for identifying students with SLDs. With federal regulations encouraging a shift from the problematic ability-achievement discrepancy model to an RTI process, it is important to ensure that states are utilizing recommended practices from the literature for SLD eligibility, specifically for the important component of interpreting student progress and ROI. A chronology of past research has documented changes in regulations and guidance documents since the IDEA regulations were adopted in 2006. For instance, one of the earliest studies (Ahearn, 2008) reported that as many as six states required an RTI process within less than two years of the 2006 IDEA regulations. Most recently, Zirkel and Thomas (2010b) and Hauerwas et al. (2013) documents 13 ad 17 states requiring RTI, respectively, in some form for SLD eligibility. The current study found 14 states to have evidence of mandating an RTI process, with the difference being explained by changes in regulations over time, specifically the inclusion of Wisconsin since the Zirkel and Thomas study, and the difference in definitions used between the Hauerwas et al. and current study.

The purpose of this study was to systematically review state regulations and documents to gather information about which recommended practices are included in state regulations and guidance documents, and to compare the inclusion of evidence of these practices in states that mandate RTI versus states that do not. Additional research questions addressed specific requirements or recommendations from states such as the

number of data points needed before making an instructional decision, specifications for a magnitude of deficiency for a student to be considered significantly below expectations, and specifications by SLD category and grade level. Findings yielded inconsistent inclusion of recommended practices across states, which is consistent with findings from similar studies (Hauerwas et al., 2013; Zirkel & Thomas, 2010b). Reasons for this finding are attributed to states historically and continually, to date, taking a liberal approach in regulations, often allowing various methods for identifying SLD.

There were more recommended practices included for states that mandate RTI than states that do not. Reasons for this finding are attributed to the logic that states requiring RTI for SLD eligibility would have components aligned with the literature compared to states that have chosen to allow an RTI process. Recommended practices were found more often in guidance documents than in regulations. Requirements and recommendations for the number of data points indicated that there is a variety of numbers and ranges suggested in state documents, which was similar to results reported in Hauerwas et al. (2013). Only 10 states set forth specific requirements for a magnitude of deficiency and they, too, varied widely. Very few states specified criteria by SLD category and grade level, and the states that did have remained the same since the data gathered in 2011 as reported in the Hauerwas et al. study.

In conclusion, many of the key features for responsibly determining student ROI, especially for high-stakes decisions of special education eligibility were not included in regulatory or guidance documents, leaving the decisions largely to LEAs and opening the field of education and school psychology to repeating problems of inconsistent eligibility requirements from the past.

Summary

The research questions in this study were exploratory and were chosen to provide an updated snapshot of requirements and recommendations in state guidance documents and to gather more specific information about practices related to monitoring student progress and determining student ROI. Previous comparable studies (Hauerwas et al., 2013; Zirkel & Thomas, 2010b) reported best practices that existed in state guidance documents but did not systematically compare information in state documents to a preset list of recommended practices found in a thorough review of literature specific to monitoring student progress.

A randomized sample of approximately 10% of states was used to calculate inter-rater reliability between two researchers. Results per checklist item ranged from fair to perfect agreement, which is sufficient for analysis (Hallgren, 2012) but also indicates room for improvement in clarifying items and procedures for searching for recommended practices in state documents.

Regulations were found for each state including the District of Columbia. Guidance documents, sometimes more than one, were found for all states except the District of Columbia or North Carolina, which only had one paragraph on their department of education website indicating that RTI was a beneficial framework for student outcomes. Compared to previous studies, more documents were reviewed in the current study. This finding is attributed to a wider use of search terms and because states continue to update, add, and refine state documents. For instance, this study found that 59% of documents had been updated since the Hauerwas et al. (2013) study, which had included a review of documents through October 2011.

The results for the first research questions indicated that states are inconsistent with the inclusion of recommended practices. The second research question found evidence that there are more recommended practices included in states that mandate RTI compared to states that do not. Evidence was found most often for the recommended practices of monitoring student progress, using CBM, monitoring student progress at Tier 1, and using direct measures of academic skills. Evidence was found least often for the recommended practices or specifications of considering outlier data, specifying grade level procedures, and using linear regression to establish a trend line. Monitoring student progress and using direct measures of academic skills were found in all states that mandate an RTI process.

The findings of this study indicate that there is room for continued research to replicate the current study, further study the validity of research tool used, update the list of recommended practices as research advances in the areas of monitoring student progress and determining ROI, and updating changes in regulations and guidance documents related to RTI requirements. Additional research is warranted for determining if the inclusion of recommended practices leads to more accurate decision-making and better student outcomes and whether consistency between eligibility criteria is beneficial to students.

Implications for the field of education include the need for pre-service training institutions to teach skills for working in teams, implementing components of an RTI framework, analyzing and interpreting data, and staying informed with best practices in the literature and changes in state regulations and recommendations. For the field of school psychology, implications for practice include pre-service training institutions

staying current with changing state regulations and recommendations for implementing an RTI framework and identifying students with SLDs. School psychologists have the opportunity to not only be responsible consumers of current best practices, but to contribute to the literature base with research related to implementing RTI, monitoring student progress, and determining student ROI.

This study, though it was an in-depth review of existing literature and state regulations and guidance documents, provides only a small picture and snapshot in time of regulatory and recommended practices for monitoring student progress and determining student ROI. As states continue to change regulations and guidance documents, and as the literature-base increases for implementing RTI and identifying SLD, there a significant need to expand this study and previous comparable studies (Hauerwas et al., 2013; Zirkel & Thomas, 2010b).

References

Abu-Hamour, B. (2013). Using curriculum-based measurement to improve student achievement in a special education university course. *International Journal of Humanities and Social Science, 3,* 163-172.

Ahern, E. M. (2008, July). State eligibility requirements for specific learning disabilities. inForum (Brief Policy Analysis). Retrieved from http://nasdse.org/DesktopModules/DNNspot-Store/ProductFiles/109_2c8774ff-500d-4b96-a8ee-f29fd07c522f.pdf

American Institutes for Research. (2011). RTI essential components integrity rubric. Retrieved from http://www.rti4success.org/pdf/RTI_Framework_Integrity_Rubric_08-25-11.pdf

Ardoin, S. P., & Christ, T. J. (2009). Curriculum-based measurement of oral reading: Standard errors associated with progress monitoring outcomes from DIBELS, AIMSweb, and an experimental passage set. *School Psychology Review, 38*, 266-283.

Ardoin, S. P., & Christ, T. J. (2008). Evaluating curriculum-based measurement slope estimates using triannual universal screenings. *School Psychology Review, 37*, 109-125.

Ardoin, S. P., Christ, T. J., Morena, L. S., Cormier, D. C., & Klingbeil, D. A. (2013). A systematic review and summarization of the recommendations and research surrounding curriculum-based measurement of oral reading fluency (CBM-R) decision rules. *Journal of School Psychology, 51*, 1-18.

Averill, O. H., & Rinaldi, C. (2011, September). Multi-tier system of supports: The RTI and PBIS approaches involve targeting specific areas in which students are struggling.

Retrieved from http://www.districtadministration.com/article/multi-tier-system-supports

Batsche, G., Elliott, J., Graden, J. L., Grimes. J., Kovaleski, J. F., R, Prasse, D., …Tilly, W. D., III (2005). *Response to intervention: Policy considerations and implementation.* Alexandria, VA: National Association of State Directors of Special Education.

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

Bluman, A. G. (2012). *Elementary statistics: A step-by-step approach* (8th ed.). New York, NY: McGraw-Hill.

Bocian, K. M., Beebe, M.E., MacMillan, D. L., & Gresham, F. M. (1999). Competing paradigms in learning disabilities classification by schools and the variations in the meaning of discrepant achievement. *Journal of Disabilities Research & Practice, 14,* 1-14.

Buffman, A., Mattos, M., & Weber, C. (2009). *Pyramid response to intervention: RTI, professional learning communities, and how to respond when kids don't learn*. Bloomington, IN: Solution Tree Press.

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

Burns, M. K., & Sensenac, B. V. (2005). Comparison of dual discrepancy criteria to assess response to intervention. *Journal of School Psychology, 43,* 393-406.

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

Bradley, R., Danielson, L., & Hallahan, D. (Eds.). (2002). *Identification of learning disabilities: Research to practice.* Mahwah, NJ: Lawrence Erlbaum.

Brown-Chidsey, R., & Steege, M. W. (2010). *Response to intervention: Principles and strategies for effective practice* (2nd Ed.). New York, NY: Guilford.

Christ, T. J. (2006). Short-term estimates of growth using curriculum based measurement of oral reading fluency: Estimating standard error of the slope to construct confidence intervals. *School Psychology Review, 35*, 128-133.

Christ, T. J., & Silberglitt, B. (2007). Estimates of the standard error of measurement for curriculum-based measures of oral reading fluency. *School Psychology Review, 36*, 130-146.

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

Christ, T. J., Zopluoglu, C. Long, J. D., & Monaghen, B. D. (2012). Curriculum-based measurement of oral reading: Quality of progress monitoring outcomes. *Exceptional Children, 78,* 356-373.

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

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

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review Special Issue: Mini-series: Issues in Data-Based Decision Making in Special Education, 30*, 507– 524.

Denzin, N. K., & Lincoln, Y. S. (2011). Introduction: The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (pp. 1-19). Thousand Oaks, CA: Sage.

Ditkowski, B. (2009). Progress monitor with confidence. Retrieved from http://www.measuredeffects.com/index.php?id=files_and_templates

Education of the Handicapped Act, Public Law No. 94-142 (1975).

Elliot, S. N., & Fuchs, L. S. (1997). The utility of curriculum-based measurement and performance assessment as alternatives to traditional intelligence and achievement tests. *School Psychology Review, 26,* 224-233.

Farmer, E., Swanlund, L., & Pluymert, K. (2010, March). *Curriculum based measurement and language proficiency in English language learners.* Paper presented at the National Association of School Psychologists Annual Convention, Chicago, IL.

Federal Register. (2006). Rules and regulations, *71*(156), 46540-46843.

Ferchalk, M. R. (2013). *Test nationally, benchmark locally: Using local DIBELS benchmarks to predict performance on the PSSA* (Doctoral dissertation). Retrieved from ProQuest. (UMI No. 3589383)

Fien, H., Park, Y., Baker, S. K., Smith, J. L. M., Stoolmiller, M., & Kame'enui, E. J. (2010). An examination of the relation of nonsense word fluency initial status and gains to reading outcomes for beginning readers. *School Psychology Review, 39*, 631-653.

Flinn, C. S. (2011). Defining rate of improvement. Retrieved from http://rateofimprovement.com/roi/category/definitions/

Flinn, C. S., & McCrea, A. E. (2012, February). *Determining eligibility for special education in an RTI system: Advanced applications of determining rate of improvement.* Paper presented at the meeting of the National Association of School Psychologists Annual Convention, Philadelphia, PA.

Flinn, C. S., & McCrea, A. E. (2013, June). *Graphing, calculating, and interpreting rate of improvement*. Paper presented at the annual meeting of the PaTTAN RtII Implementers Forum, Hershey, PA.

Fuchs, D., & Fuchs, L. S. (2001). Responsiveness-to-intervention: A blueprint for practitioners, policymakers, and parents. *Teaching Exceptional Children, 38,* 57-61.

Fuchs, D., & Young, C. L. (2006). On the irrelevance of intelligence in predicting responsiveness to reading instruction. *Council for Exceptional Children, 73,* 8-30.

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

Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review, 33*, 189-192.

Fuchs, L. S., & Fuchs, D. (1999). Monitoring student progress toward the development of reading competence: A review of three forms of classroom-based assessment. *School Psychology Review, 28*, 659-671.

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

Fuchs, L. S., & Fuchs, D. (n.d.) Using CBM for progress monitoring. Retrieved from http://www.studentprogress.org

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22,* 27-48.

Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research and Practice, 13*, 204-219.

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction (8th ed.).* New York, NY: Pearson.

Gibbons, K. (2008). Evaluating RTIs effectiveness over the long term. *School Administrator, 65*(8), 13.

Gleason, T. (2013). *The impact of RTI practices on student reading achievement.* Retrieved from ProQuest Digital Dissertations (AAT 3558733).

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

Good, R. H. (1990). Forecasting accuracy of slope estimates for reading curriculum based measurement: Empirical evidence. *Behavioral Assessment, 12*, 179-193.

Good, R. H., Wheeler, C. E., Cummings, K. D., Baker, S. K., Fien, H., & Kame'enui, E. J. (2010, March). *Rigorous RtI decisions: Normative growth rates for oral reading fluency.* Paper presented at the meeting of the National Association of School Psychologists Annual Convention in Chicago, IL.

Google. (n.d.). Google drive. Retrieved from http://drive.google.com

Graney, S. B., Missall, K. N., Martinez, R. S., & Bergstrom, M. (2009). A preliminary investigation of within-year growth patterns in reading and mathematics curriculum-based measures. *Journal of School Psychology, 47*, 121-142.

Gresham, F. (2001). Responsiveness to intervention: An alternative approach to the identification of learning disabilities. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.). *Identification of learning disabilities: Research in practice* (pp. 467–519). Mahwah, NJ: Erlbaum.

Gresham, F. M., & Vellutino, F. R. (2010). What is the role of intelligence in the identification of specific learning disabilities? Issues and clarifications. *Learning Disabilities Research & Practice, 25,* 194-206.

Griffiths, A., Parson, L. B., Burns, M. K., VanDerHeyden, A., & Tilly, W. A. (2007). *Response to intervention: Research for practice*. Alexandria, VA: National Association of State Directors of Special Education.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology, 8,* 23-34.

Hauerwas, L. B., Brown, R., & Scott, A. N. (2013). Specific learning disability and response to intervention: State level guidance. *Council for Exceptional Children, 80*, (101-120).

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

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

Hixson, M., Christ, T. J., & Bradley-Johnson, S. (2008). Best practices in the analysis of progress-monitoring data and decision-making. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 2133-2146). Bethesda, MD: National Association of School Psychologists.

Hoover, J. J., Baca, L., Wexler-Love, E., & Saenz, L. (2008, August). *National implementation of response to intervention (RTI): Research summary.* Retrieved from http://www.nasdse.org/Portals/0/NationalImplementationofRTI-ResearchSummary.pdf

Individuals with Disabilities Education Act Amendments (IDEA) of 1997, Public Law No. 105-117. (1997).

Individuals with Disabilities Education Improvement Act of 2004 (IDEA), Pub. L. No. 108-446. (2004).

Individual with Disabilities Education Improvement Act of 2004 Regulations, 34. CFR Parts 300-301. (2004).

Iowa Administrative Code (2011). Title VII special education chapter 41 special education. Retrieved from https://www.educateiowa.gov/sites/files/ed/documents/Iowa%20Administrative%20Rules%20of%20Special%20Education%202010.pdf

Iris Center. (n.d.) Slope calculator. Retrieved from http://iris.peabody.vanderbilt.edu/resources.html

Jenkins, J. R., Graff, J. J., & Miglioretti, D.L. (2009). Estimating reading growth using intermittent CBM progress monitoring. *Exceptional Children, 75*, 151-163.

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

Kovaleski, J. F. (2007). Potential pitfalls of response to intervention. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of Response to Intervention: The science and practice of assessment and intervention.* New York, NY: Springer.

Kovaleski, J. F., & Flinn, C. S. (2011, February). *Determining special education eligibility in an RTI model.* Paper presented at the meeting of the National Association of School Psychologists Annual Convention, San Francisco, CA.

Kovaleski, J. F., VanDerHeyden, A. M., & Shapiro, E. S. (2013). *The RTI approach to evaluating learning disabilities.* New York, NY; Guilford Press.

Kratochwill, T. R., Volpiansky, P., Clements, M., & Ball, C. (2007). Professional development in implementing and sustaining multitier prevention models: Implications for response to intervention. *School Psychology Review, 36,* 618-631.

Krippendorff, K. (2012). *Content analysis: An introduction to its methodology (3rd ed.)*. Washington, DC: Sage Publications.

Lindsley, O. R. (1990). Precision teaching: By teachers for children. *Teaching Exceptional Children, 22*(3), 10-15.

Lyon, G. R., Fletcher, J. M., Fuchs, L. S., & Chhabra, V. (2006). Learning disabilities. In E.J. Mash & R. Barkley (Eds.). *Treatment of childhood disorders* (3rd ed., pp. 512-594). Hillsdale, NJ: Erlbaum.

MacMillan, D. L., Gresham, F. M., & Bocian, K. M. (1998). Discrepancy between definitions of learning disabilities and school practices: An empirical investigation. *Journal of Learning Disabilities, 31,* 314-326.

Marston, D. & Magnusson, D. (1985). Implementing curriculum-based measurement in special and regular education settings. *Exceptional Children, 52,* 266-276.

McDougal, J. L., LeBlanc, M., & Hintze, J. (2010). Graphing student academic data for universal screening and progress monitoring. In A. S. Canter, L. Z. Paige, & S. Shaw (Eds.) *Helping children at home and school III: Handouts for families and educators* (pp. S3H15–1 - S3H15–4). Bethesda, MD: National Association of School Psychologists.

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

McMaster, K. (2011). Curriculum-based measures of beginning writing: Technical features of the slope. Retrieved from http://readperiodicals.com/201101/2251251451.html

McMaster, K. L., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2005). Responding to nonresponders: An experimental field of trial of identification and intervention methods. *Council for Exceptional Children, 71,* 445-463.

Mellard, D. F., McKnight, M., & Woods, K. (2009). Response to intervention screening and progress-monitoring practices in 41 local schools. *Learning Disabilities Research & Practice, 24,* 186-195.

Meyer, M. S. (2000). The ability-achievement discrepancy: Does it contribute to an understanding of learning disabilities? *Educational Psychology Review, 12*, 315-337.

Mississippi Department of Education. (2009). State policies regarding children with disabilities under the individuals with disabilities education act amendments of 2004. Retrieved from http://www.mde.k12.ms.us/docs/sped-july-2009-policies/Policy_06-17-09.pdf?sfvrsn=2

Moss, M., Jacob, R., Boulay, B., Horst, M., & Poulos, J. (2006). Reading First implementation evaluation: Interim Report. Retrieved from http://www2.ed.gov/rschstat/eval/other/readingfirst-interim/readingfirst.pdf

National Association of School Psychologists. (2006). New roles in response to intervention: Creating success for schools and children. Retrieved from http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCAQFjAA&url=http%3A%2F%2Fwww.nasponline.org%2Fadvocacy%2FNew%2520Roles%2520in%2520RTI.pdf&ei=tnFUVIrXDcqZgwTr6oL4Cg&usg=AFQjCNFdniUHJjEHNqw9RJc8AOwpZTPo6Q

National Center on Response to Intervention. (2010). States chart. Retrieved from http://state.rti4success.org/

National Center on Response to Intervention. (2011). Progress monitoring tools. Retrieved from http://www.rti4success.org/progressMonitoringTools

National Center on Response to Intervention. (n.d.). Universal screening. Retrieved from http://www.rti4success.org/categorycontents/universal_screening

National Joint Committee on Learning Disabilities. (2005). Responsiveness to intervention and learning disabilities. Retrieved from http://www.ldaamerica.us/pdf/rti2005.pdf

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

National Reading Panel. (2000). *Report of the national reading panel: Teaching students to read: An evidence-based assessment of the scientific literature on reading and its implications for reading instructions.* Washington, DC: National Institute of Child Health and Human Development, National Institutes of Health, U.S. Government Printing Office.

No Child Left Behind Act of 2001. Pub. L. No. 107-110, 115 Stat. 1425 (2002).

Office of Special Education Programs. (2007). Questions and answers on response to intervention and early intervening services. Retrieved from http://www.rti4success.org/pdf/qaonrtiandeis.pdf

RANDOM.ORG. (2014). True random number generator. Retrieved from http://www.random.org/

Renaissance Learning. (2013). *STAR reading*. Wisconsin Rapids, WI: Author.

Reschly, D. J., & Hosp, J. L. (2004). State SLD identification policies and practices. *Learning Disability Quarterly, 27,* 197-213.

Restori, A. F., Katz, G. S., & Lee, H. B. (2009). A critique of the IQ / achievement discrepancy model for identifying learning disabilities. *Europe's Journal of Psychology, 4*, 128-145.

Shapiro, E. S. (2008). Best practices in setting progress monitoring goals for academic skill improvement. In A. Thomas and J. Grimes (Eds.), *Best practices in school psychology V.* (Vol. 2, pp. 141-157). Bethesda, MD: National Association of School Psychologists.

Shapiro, E., Keller, M., Lutz, J., Santoro, L., & Hintze, J. (2006). Curriculum-Based Measures and Performance on State Assessment and Standardized Tests: Reading and Math Performance in Pennsylvania. *Journal of Psychoeducational Assessment, 24*, 19-35.

Shinn, M. R. (2008). Best practices in using curriculum-based measurement in a problem-solving model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 243-362). Bethesda, MD: National Association of School Psychologists.

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

Shinn, M. R., Gleason, M. M., & Tindal, G. (1989). Varying the difficulty of testing materials: Implications for curriculum-based measurement. *The Journal of Special Education, 23,* 223-233.

Shinn, M. R., Good, R. H., & Stein, S. (1989). Summarizing trend in student achievement: A comparison of methods. *School Psychology Review, 18*, 356-370.

Silberglitt, B., Burns, M., Madyun, N., & Lail, K. (2006). Relationship of reading fluency assessment data with state accountability test scores: A longitudinal comparison of grade levels. *Psychology in the Schools, 43*, 527-535.

Silberglitt, B., & Hintze, J. M. (2007). How much growth can we expect? A conditional analysis of R-CBM growth rates by level of performance. *Exceptional Children, 74*, 71-84.

Speece, D. L., Case, L. P., & Malloy, D. E. (2003). Responsiveness to general education instruction as the first gate to learning disabilities identification. *Learning Disabilities Research & Practice, 18*, 147-156.

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

Stecker, P. M., Fuchs, D., & Fuchs, L. S. (2008). Progress monitoring as essential practice within response to intervention. *Rural Special Education Quarterly, 27*(4), 10-17.

Stemler, S. (2001). An *introduction to content analysis.* Retrieved from ERIC database. (ED458218)

Stewart, L. H., & Silberglitt, B. (2008). Best practices in developing academic local norms. In A. Thomas and J. Grimes (Eds.) *Best practices in school psychology V*. (Vol. 2, pp. 225-242). Bethesda, MD: National Association of School Psychologists.

Stuebing, K. K., Barth, A. E., Molfese, P. J., Weiss, B., & Fletcher, J. M. (2009). IQ is not strongly related to response to reading instruction: A meta-analytic interpretation. *Council for Exceptional Children, 76,* 31-55.

Stuebing, K. K., Fletcher, J., LeDoux, J., Lyon, G. R., Shaywitz, S., & Shaywitz, B. (2002). Validity of IQ-discrepancy classifications of reading disabilities: A meta-analysis. *American Educational Research Journal, 39,* 469–518.

Stone, E., & Davey, T. (2011). Computer-adaptive testing for students with disabilities: A review of the literature. Research Report. ETS RR-11-32. Retrieved from http://www.ets.org/Media/Research/pdf/RR-11-32.pdf

Technical Assistance Coordination Center (2012). *Historical state-level IDEA data files: Part B child count 2011* [data file]. Retrieved from http://tadnet.public.tadnet.org/pages/712

Thornblad, S. C., & Christ, T. J. (2014). Curriculum-based measurement of reading: Is 6 weeks of daily progress monitoring enough? *School Psychology Review, 43*, 19-29.

Torgesen, J. K., (2009). The response to intervention instructional model: Some outcomes from a large-scale implementation in Reading First schools. *Child Development Perspectives, 3,* 38-40.

U.S. Department of Education. (n.d.). Definitions. Retrieved from http://www.ed.gov/race-top/district-competition/definitions

U.S. Department of Education. (2006, August, 14). Assistance to the states for the education of children with disabilities. Federal Register, 71(156). Retrieved from http://idea.ed.gov/download/finalregulations.pdf

US Legal. (2013). State education agency law and legal definition. Retrieved from http://definitions.uslegal.com/s/state-educational-agency-education/

VanDerHeyden, A. M., & Burns, M. K. (2010). *Essentials of response to intervention.* Hoboken, NJ: John Wiley & Sons, Inc.

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

Vaughn, S., & Klinger, J. (2007). Overview of the three-tier model of reading intervention. In D. Haager, J. Klinger, & S. Vaughn (Eds.), *Evidenced-based reading practices for response to intervention* (pp. 3-10). Baltimore, MD: Paul H. Brookes Publishing Co.

Vaughn, S., Wanzek, J., Linan-Thompson, S., & Murray, C. S. (2007). Monitoring response to supplemental services for students at risk for reading difficulties: High and low responders. In S. R. Jimerson, M. K. Burns & A. M. VanDerHeyden (Eds.), H*andbook of*

response to intervention: The science and practice of assessment and intervention (pp. 236-245). New York, NY: Springer.

Vaughn, S., Wanzek, J., Woodruff, A. L., & Linan-Thompson, S. (2007). Prevention and early identification of students with reading disabilities: A research review of the three-tier model. In D. H. Haager, S. Vaughn & J. K. Klingner (Eds.), *Evidence-based reading practices for response to intervention* (pp 11-28). Baltimore, MD: Brookes.

Vellutino, F. R., Scanlon, D. M., & Lyon, G. R. (2000). Differentiating between difficult-to-remediate and readily-remediated poor readers: More evidence against the IQ-achievement discrepancy definition of reading disability. *Journal of Learning Disabilities, 33,* 223-238.

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

Wright, J. (1992). CBM manual for teachers. Retrieved from http://www.cbmnow.com/documents/cbaManualhand.pdf

Wright, P. W. D., Hale, J. B., Backenson, E. M., Eusebio, E. C., & Dixon, S. G. (2013). Forest Grove v. T.A. rejoinder to Zirkel: an attempt to profit from malfeasance? *Journal of Psychoeducational Assessment, 31*, 318-325.

Wyoming Department of Education. (2010). Wyoming department of education special programs division policy and procedure manual for special education. Retrieved from http://edu.wyoming.gov/downloads/special-ed/SpecEd_Policy_and_Procedure_Manual_v__1_1FINAL_8-20-2010.pdf

Ysseldyke, J., Algozzine, B., & Epps, S. (1983). A logical and empirical analysis of current practice in classifying students as handicapped. *Exceptional Children, 50*, 160-166.

Zaiontz, C. (2014). Real statistics using Excel: Cohen's kappa. Retrieved from http://www.real-statistics.com/reliability/cohens-kappa/

Zirkel, P. A. (2011, May). State laws and guidelines for RTI: Additional implementation features. *Communiqué, 39*(7), 30-32.

Zirkel, P. A., & Krohn, N. (2008). RTI after IDEA: A survey of state laws. *Teaching Exceptional Children, 40*, 71-73.

Zirkel, P. A., & Thomas, L. B. (2010a). State laws for RTI: An updated snapshot. *Council for Exceptional Children, 42,* 56-63.

Zirkel, P. A., & Thomas, L. B. (2010b). State laws and guidelines for implementing RTI. *Teaching Exceptional Children, 43*, 60-73.

Zirkel, P. A. (2013). The trend in SLD enrollments and the role of RTI. *Journal of Learning Disabilities, 46,* 473-479.

Appendix A

Checklist of Recommended Practices for Determining Rate of Improvement

Directions: Use one checklist form per state document. Note that some items may require you to enter text, choose from preset options, or check all items that apply. Some items may include disqualifying content. Read those items carefully for requirements.

State Being Reviewed
Use the initials of the state being reviewed (e.g., PA)

Date of Document or Update (If no date, write "n.d.")

Regulations or Guidance Document?
Choose whether the document being reviewed includes regulations or guidance

- Regulations
- Guidance

1. The state guidance document describes use of multiple administrations of assessments; monitor student progress over time; progress monitoring; formative assessment.
Note: This item must be in addition to the federal requirement of "repeated assessments." If "repeated assessments" is the only information found, then record a response of "No Evidence Found."

- Evidence Found
- No Evidence Found

2. The state guidance document describes use of measures that have alternate forms or multiple assessments that are similar per grade level.

- Evidence Found
- No Evidence Found

3. The state guidance document describes use of direct measures of academic skill(s) for monitoring student progress / response to intervention.

- Evidence Found
- No Evidence Found

4. The state guidance document describes use of brief, short measures for monitoring student progress; measures that are efficient for monitoring student progress.

- Evidence Found

- No Evidence Found

5. The state guidance document describes use of measures that are validated for determining ROI: curriculum-based measurements (CBMs).
- Evidence Found
- No Evidence Found

6. The state guidance document describes use of measures that are validated for determining ROI: computer adaptive tests (CATs).
- Evidence Found
- No Evidence Found

7. The state guidance document cite examples of assessments for monitoring student progress such as AIMSweb, DIBELS, STAR Math, or suggests a review of assessments listed on the National Center on Response to Intervention and National Center on Intensive Intervention web sites (http://www.intensiveintervention.org/chart/progress-monitoring).
Note: If the state documents suggests locally developed tests (or others as listed in the disqualifying content below) AND acceptable measures as listed above, the disqualifying content takes precedence. In this case, select Disqualifying Content.
- Evidence Found
- No Evidence Found
- Disqualifying Content Found: The state guidance document cites non-examples including teacher-made tests, locally developed assessments, unit tests, grades, individually administered standardized norm-referenced assessments of achievement (e.g., Woodcock-Johnson Tests of Achievement, Third Edition; Wechsler Individual Achievement Test, Third Edition).

8. The state guidance document directs school teams to use measures that are technically adequate (i.e., valid, reliable) indicators of the skill(s) being monitored.
- Evidence Found
- No Evidence Found

9. The state guidance document describes use of measures that are sensitive to small increments of growth (progress, change, response to instruction/intervention).
- Evidence Found
- No Evidence Found

10. The state guidance document describes monitoring student academic progress on grade level (in which the student is enrolled) for special education eligibility if CBM is suggested. (Does not apply to CATs since they automatically adjust level of difficulty of items).
Note: If both grade level and instructional level are part of the requirement for special education eligibility, choose Disqualifying Content Found.
- Evidence Found
- No Evidence Found

- Disqualifying Content Found: The state guidance document describes use of instructional level data for special education eligibility.

11. The state guidance document describes monitoring of student progress at Tier 1 with a frequency of at least three to four times a year.
Note: Documents that describe screening taking place in the fall, winter, and spring but do not specify three to four times per year is still considered Evidence Found.

- Evidence Found
- No Evidence Found

12. The state guidance document describes monitoring of student progress at Tier 2 with a frequency of at least monthly or every other week.
Note: Documents that specify bi-monthly only (no other clarifying descriptions) for Tier 2 should be considered No Evidence Found since bi-monthly is ambiguous (i.e., could mean twice a month or once every two months).

- Evidence Found
- No Evidence Found

13. The state guidance document describes monitoring of student progress at Tier 3 with a frequency of at least weekly.
Note: Documents that specify bi-weekly only (no other clarifying descriptions) for Tier 2 should be considered No Evidence Found since bi-weekly is ambiguous (i.e., could mean twice a week or once every two weeks).

- Evidence Found
- No Evidence Found

14. The state guidance document directs school teams to carefully consider and document whether a sufficient number of data points have been gathered to produce a stable trend line.

- Evidence Found
- No Evidence Found

15. The state guidance document directs school teams to carefully consider and document a specific number of data points for instructional decision-making.
Note: Record the number suggested in the state document below. If a range is suggested, record the range (i.e., 3-4 data points). If there is no suggested number of data points, type "Not Specified."

16. The state guidance document directs school teams consider whether assessments used to monitor student progress were administered with fidelity and in accordance with standardization.

- Evidence Found
- No Evidence Found

17. The state guidance document describes first completing a visual inspection of student response data graphed on an x- and y-axis graph prior to generating a trend line.

- Evidence Found
- No Evidence Found

18. The state guidance document describes that after student response data are graphed, school teams should then determine if a linear trend line best describes the student's progress. In other words, school teams have a procedure for considering effects of an outlier data point.

- Evidence Found
- No Evidence Found

19. The state guidance document describes use of ordinary least squares (OLS) / linear regression to generate a trend line for determining ROI.

- Evidence Found
- No Evidence Found

20. The state guidance document directs school teams to consider the comparison group(s) (i.e., norm group(s)) to determine adequacy of individual student ROI. School teams should determine which set(s) of norms most closely match the population of the grade level of the student referred for special education evaluation.

Examples: -Student ROI should be compared to national norms available through assessment systems or research results. -Student ROI should be compared to the ROI provided by a commercially available set of assessments or assessment tools. -Student ROI should be compared to local norms within a school building, grade level, or specific classroom.

- Evidence Found
- No Evidence Found

21. The state guidance document indicates that a response to intervention (RTI) process is required or mandated for determining specific learning disability (SLD) eligibility for special education.

Note: If other methods are allowed in addition to an RTI process, then the response is "No Evidence Found."

- Evidence Found
- No Evidence Found

22. The state guidance document defines a specific magnitude of deficiency in relation to age- or grade-level standards.

Record the stated magnitude below. If a specific quantifiable magnitude of deficiency is stated, record the exact wording. Examples: (a) 1.25 standard deviations below the mean, (b) at or below the 10th percentile. If a general description is used to describe the magnitude, record the exact phrase. Examples: (a) significantly below peers, (b) below

reasonable expectations for grade or age. If the magnitude of deficiency is not described at all, type "Not Specified."

23. The state guidance document differentiates between eligibility requirements using an RTI process for each SLD type (i.e., reading fluency). Check all that apply.
If the state document does not specify requirements for SLD types, check "Does Not Specify."

- Basic Reading Skill
- Reading Fluency Skills
- Reading Comprehension
- Mathematics Calculation
- Mathematics Reasoning
- Written Expression
- Listening Comprehension
- Oral Expression
- Does Not Specify

24. The state guidance document specifies the use of an RTI process for only the following grade levels. Check all that apply.
If state document does not specify the use of an RTI process for specific grade levels only, then check "Does Not Specify."

- Kindergarten
- 1st Grade
- 2nd Grade
- 3rd Grade
- 4th Grade
- 5th Grade
- 6th Grade
- 7th Grade
- 8th Grade
- 9th Grade
- 10th Grade
- 11th Grade
- 12th Grade (or through maximum school-age limit)
- Does Not Specify