

5-2019

Using Response to Intervention to Determine Eligibility for Specific Learning Disabilities in Reading: What Determines a Non-Responder?

Drew Hunter

Follow this and additional works at: <https://knowledge.library.iup.edu/etd>

Recommended Citation

Hunter, Drew, "Using Response to Intervention to Determine Eligibility for Specific Learning Disabilities in Reading: What Determines a Non-Responder?" (2019). *Theses and Dissertations (All)*. 1685.
<https://knowledge.library.iup.edu/etd/1685>

This Dissertation is brought to you for free and open access by Knowledge Repository @ IUP. It has been accepted for inclusion in Theses and Dissertations (All) by an authorized administrator of Knowledge Repository @ IUP. For more information, please contact cclouser@iup.edu, sara.parme@iup.edu, edzimmer@iup.edu.

USING RESPONSE TO INTERVENTION TO DETERMINE ELIGIBILITY FOR SPECIFIC
LEARNING DISABILITIES IN READING: WHAT DETERMINES A NON-RESPONDER?

A Dissertation

Submitted to the School of Graduate Studies and Research

in Partial Fulfillment of the

Requirements for the Degree

Doctor of Education

Drew W. Hunter

Indiana University of Pennsylvania

May 2019

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

We hereby approve the dissertation of

Drew W. Hunter

Candidate for the degree of Doctor of Education

Timothy Runge, Ph.D.
Professor of Educational and School Psychology,
Chair

Roger L. Briscoe, Ph.D.
Professor of Educational and School Psychology

Daniel R. Wissinger, Ph.D.
Assistant Professor of Professional Studies in
Education

Joseph F. Kovaleski, D.Ed.
Professor Emeritus of Educational and School
Psychology, Retired

ACCEPTED

Randy L. Martin, Ph.D.
Dean
School of Graduate Studies and Research

Title: Using Response to Intervention to Determine Eligibility for Specific Learning Disabilities in Reading: What Determines a Non-Responder?

Author: Drew W. Hunter

Dissertation Chair: Dr. Timothy Runge

Dissertation Committee Members: Dr. Roger L. Briscoe
Dr. Daniel R. Wissinger
Dr. Joseph F. Kovaleski

The purpose of this study was to examine the dual-discrepancy model (Fuchs & Fuchs, 1998) and its application to the identification of specific learning disabilities (SLD) in reading when using a response to intervention (RTI) decision-making framework. Curriculum-based measurement-reading (CBM-R) data and demographic data were obtained for 163 students in grades two through four from one Mid-Atlantic school district. The data were analyzed to determine if students who were referred for multidisciplinary evaluation (MDE) and determined eligible, referred for MDE and determined ineligible, and students who received Tier 3 reading support and were not referred for MDE were different on their level of performance or rate of improvement (ROI) using CBM-R data. The data were also analyzed to determine if level of performance and ROI predicated eligibility for special education more than other student demographic data. The results indicated that students referred for MDE and determined eligible had significantly lower levels of performance than students receiving Tier 3 support and not referred for evaluation. Student ROI was not meaningfully different among the three levels of the dependent variable. The results also revealed that level of performance significantly predicted student eligibility for special education but ROI did not. Implications for research, practice, and policy related to using RTI for SLD eligibility decisions are discussed.

TABLE OF CONTENTS

Chapter		Page
1	INTRODUCTION	1
	Statement of the Problem	3
	Research Questions and Hypotheses	6
	Assumptions	7
	Limitations.....	8
	Definition of Terms	8
2	REVIEW OF LITERATURE.....	11
	Multi-Tiered Systems of Support and Response to Intervention	12
	Research on Non-Responders.....	15
	Comparison of Different Methods for Determining Intervention Response.....	19
	Assessing Student Growth using Rate-of-Improvement	24
	Passage Equivalency	27
	The Effect of Passage Equivalency on ROI	30
	Using RTI for Determining SLD Eligibility: Current Recommendations, Unanswered Questions, and Future Direction	33
	Statement of the Problem	39
	Summary.....	40
3	METHODS AND PROCEDURES	41
	Introduction	41
	Research Questions and Hypotheses	41
	Design.....	42
	Population.....	43
	Sample	43
	Measurement	44
	Research Question 1	44
	Research Question 2	46
	Procedures	48
	Statistical Analysis	49
	Summary.....	51
4	RESULTS.....	52
	Sample.....	52
	Results of Analyses.....	54
	Research Question 1	54

Chapter		Page
	Research Question 2	60
	Conclusion	61
5	DISCUSSION.....	64
	Summary of Findings	64
	Research Question 1	65
	Research Question 2	68
	Limitations.....	71
	Implications for Research.....	73
	Implications for Policy	76
	Implications for Practice.....	79
	Summary.....	81
	REFERENCES	84
	APPENDICES	95
	Appendix A - Institutional Review Board Approval.....	95
	Appendix B - Institutional Review Board Modification of Project	97
	Appendix C - Normal Q-Q Plots for the Dependent Variable of Level of Performance of Research Question 1	99
	Appendix D - Detrended Q-Q Plots for the Dependent Variable of Level of Performance for Research Question 1.....	100
	Appendix E - Box and Whisker Plots for the Dependent Variable of Level of Performance for Research Question 1.....	101
	Appendix F - Normal Q-Q Plots for the Dependent Variable of ROI in Research Question 1	102
	Appendix G - Detrended Q-Q Plots for the Dependent Variable of ROI for Research Question 1.....	103
	Appendix H - Box and Whisker Plots for the Dependent Variable of ROI for Research Question 1	104
	Appendix I - Variance Inflation Factor for the Independent Variables of Research Questions 2.....	105
	Appendix J - Studentized Residuals for Binomial Logistic Regression	106

LIST OF TABLES

Table	Page
1	Number of Student Cases by Sex..... 53
2	Number of Student Cases by Grade..... 53
3	Number of Student Cases by Race..... 54
4	Number of Student Cases by Free and Reduced Meal Status..... 54
5	Number of Student Cases by Special Education Eligibility Status..... 55
6	Descriptive Statistics for Level of Performance and ROI for Total Sample 57
7	Descriptive Statistics for Level of Performance and ROI by Disability Status 57
8	MANOVA Results for the Effects of Eligibility Status on Level of Performance and ROI 58
9	ANOVA Results for the Effects of Eligibility Status on Level of Performance and ROI 58
10	Post Hoc Analyses for Main Effect Using Tukey HSD..... 59
11	Binomial Logistic Regression Predicting Special Education Eligibility Based on Sex, Grade, Race, Free and Reduced Meal Status, Level of Performance, and ROI 62

CHAPTER 1

INTRODUCTION

The 2004 reauthorization of the Individuals with Disabilities Education Act (IDEA) provided revisions allowing for the use of a response to intervention (RTI) model for making eligibility determinations for specific learning disabilities (SLD). IDEA 2004 also prohibited state mandates requiring schools to use an ability-achievement discrepancy model for SLD determination (IDEA, 2004). Although there is much research to support the use of RTI as a method of intervening early for academic problems (Burns, Appleton, & Stehouwer, 2005; Fuchs, Mock, Morgan, & Young, 2003), there are still many questions that need to be answered regarding its implementation and use as a method of identifying SLD.

SLD has historically been conceptualized as an unexpectedly poor academic performance evidenced by low achievement in reading, written expression, or math (Fletcher, Lyons, Fuchs, & Barnes, 2007). This underachievement is not attributable to environmental or intellectual factors. The passage of the Education for All Handicapped Children Act of 1975, which has been reauthorized several times since as IDEA, was the first law to establish federal regulations for the identification of students with SLD and the provision of special education for these students. In 1977, the federal government included regulations for the identification of SLD using the ability-achievement discrepancy method. In this model, intellectual functioning is compared to achievement in reading, mathematics, written expression, oral expression, and listening comprehension on standardized measures of achievement. Poor academic functioning that is below what would be expected when compared to intellectual functioning is used to identify students with SLD. However, attempts to empirically validate the ability-achievement discrepancy model have not been successful. Research has not been able to differentiate a group

of students who are ability-achievement discrepant from students who are low ability-low achievers in terms of their skill level and the underlying cognitive processes believed to cause the disorders (Hoskyn & Swanson, 2000; Stuebing et al., 2002).

RTI is a system of early intervention and a possible method for identifying SLD that uses data-based decision making and a system of tiered interventions of increasing intensity designed to improve student learning outcomes. It is a multi-tiered system of support (MTSS) used in conjunction with universal screening, the identification of specific goals and objectives to meet standards for academic and behavioral outcomes, the identification of specific student needs and scientifically-validated interventions to bring students to desired levels of performance, frequent progress monitoring, and problem-solving using school, classroom, and student-level data (Reschly & Bergstrom, 2009). What distinguishes RTI from MTSS in general is that RTI refers to the process of measuring a student's response to robust, research-based interventions and using this response, or lack thereof, as the basis for making decisions regarding intervention effectiveness, movement across tiers, and special education eligibility decisions for SLD. So, MTSS refers to the framework for providing tiered interventions according to student needs and RTI is the process within MTSS systems that measures the student's response to these efforts (Torgesen, 2007).

MTSS models that use RTI as part of SLD determination decisions have been shown to be effective at improving outcomes for students and schools while simultaneously resulting in fewer students being identified as SLD (Burns, Appleton, & Stehouwer, 2005). However, no single model of RTI for diagnostic purposes exists in research or state guidelines (Zirkel & Thomas, 2010b). As a result, critics of RTI argue that there is a lack of empirically-based

guidelines for the implementation of RTI systems as well as guidelines for using RTI as a means of disability determination (Reynolds & Shaywitz, 2009a; 2009b).

Statement of the Problem

Following the passage of IDEA 2004, proponents of RTI identified the need for more research into key aspects of using RTI for SLD determination including what constitutes an adequate response to an intervention, which measures and what cut points to use when determining adequate intervention response, and more specific guidelines for determining disability status when using RTI (Fuchs & Deshler, 2007). A review of the literature shows five different methods for determining adequate or inadequate response-to-intervention: split median, final normalization, final benchmark, slope discrepancy, and dual discrepancy (Fuchs, Compton, Fuchs, Bryant, & Davis, 2008). Research shows that using different methods of determining adequate intervention response and different cut-points for operationalizing adequate intervention response results in different groups of students being identified as non-responsive, each with different levels of sensitivity, specificity, and other technical adequacies (Barth, et al., 2008; Brown-Waesche, Schatschneider, Maner, Ahmed, & Wagner, 2011; Burns, Scholin, Kosciolk, & Livingston, 2010; Burns & Senesac, 2005). If RTI is to be used to determine special education eligibility, more research as to what constitutes a non-responder is needed to ensure consistency in the students identified as SLD.

To date, the dual-discrepancy model proposed by Fuchs and Fuchs (1998) has the most empirical support. This model requires that a student must be discrepant in both level of performance and growth when compared to some benchmark of performance. Research has shown that the dual-discrepancy model outperforms ability-achievement discrepancies and simple low achievement methods for identifying SLD by identifying a group of students who are

more consistent with estimated prevalence rates for SLD, more impaired on tasks of phonological processing, and more representative of the general population (Speece & Case, 2001; Speece, Case, & Malloy, 2003). The dual-discrepancy model also outperforms other RTI methods for determining adequate intervention response (Fuchs, Compton, Fuchs, Bryant, & Davis, 2008).

The dual-discrepancy model often makes use of data from curriculum-based measures to determine level of performance and growth rates. Curriculum-based measures are brief, reliable and valid measures of general reading, math, and writing ability (Deno, 1985). The properties of curriculum-based measurement (CBM) make it well suited for use within the dual-discrepancy model, particularly its ability to be administered repeatedly and frequently, the ability to display the results graphically, and the ability to use peer or norm-referencing to determine a student's level of performance. Research has also shown that CBM is sensitive to differences in student performance and can differentiate students without disabilities from students with SLD (Shinn & Marston, 1985; Shinn, Ysseldyke, Deno, & Tindal, 1986). It is also sensitive to student growth and is ideal for the measuring students' progress in response to instructional efforts (Marston, Fuchs, & Deno, 1986).

When using the dual-discrepancy model to determine intervention response, CBM data are graphed in a time-series display. The slope derived from this process, also called rate of improvement (ROI), is used to determine student growth in response to instructional efforts. Research by Shinn, Good, and Stein (1989) determined that using an ordinary least squares (OLS) regression line is the most technically adequate way to summarize ROI when using CBM data graphed in a time-series manner. Using an OLS method, Fuchs, Fuchs, Hamlett, Waltz, and Germann (1993) estimated yearly ROI for a sample of first through sixth-grade students

using CBM in reading, math, and spelling to determine typical growth rates for students at each grade. Growth trajectories as students advance through the grades have also been researched and established (Nese et al., 2013). This research provides a reference of comparison when examining student growth at various grades or across a student's education. Deno, Fuchs, Marston, and Shin (2001) have also determined growth rates for students with SLD receiving typical special education programming using curriculum-based measurement in reading (CBM-R), demonstrating that ROI can differentiate students with SLD from their non-identified peers. Although much is known about the ROI of typically progressing students, there are no clear guidelines for what ROIs constitute an inadequate response to an intervention or the presence of SLD. Recent research has also called into question the reliability of slopes derived from CBM-R data at the individual level (Ardoyn & Christ, 2009; Ardoyn, Christ, Morena, Cormier, & Klingbeil, 2013; Christ, 2006; Christ, Zopluoglu, Long, & Monaghan, 2012; Christ, Zopluoglu, Monaghan & Van Norman, 2013).

Several research-based models have been proposed for implementing RTI that incorporate a dual-discrepancy approach for making special education eligibility decisions (Johnson, Mellard, Fuchs, & McKnight, 2006; Kovaleski, VanDerHeyden, & Shapiro, 2013;). However, neither has provided empirically-based guidelines for how deficient a student should be in level of performance and/or ROI to before special education eligibility is warranted. States have provided guidance on using RTI for SLD identification, as well, but have given varied recommendations and requirements regarding progress monitoring practices and the level of deficiency required for SLD eligibility, both of which are central to determining adequate intervention response (Flinn, 2014). Given the available research and guidance on RTI,

determining adequate intervention response and its use in identification practices is an area where more research is still needed.

Research Questions and Hypotheses

Research Question 1: Do students who are identified as SLD in reading using RTI procedures /methods differ from those students who are determined ineligible for special education and those students receiving supplemental intervention but not referred for evaluation on their ROI and/or level of performance? A single-factor experimental design will be used to compare the ROI and the level of performance of students determined eligible for special education to those students who are determined to be ineligible, and to those students receiving intervention but not referred for evaluation to establish if significant differences between these groups exist. In this design, the students' status in relation to their referral for, or identification of, SLD (eligible, ineligible, not referred) is the independent variable and ROI and level of performance using CMB-R data are the dependent variables. It is hypothesized that students identified as SLD will have significantly lower levels of performance and lower ROIs than students not identified as SLD and students not referred for evaluation. Peterson and Shinn (2002) found that low achievement relative to local standards best explains which students are identified as SLD, regardless of what eligibility criteria are in place. Research has also shown that using both ROI and level of performance in SLD determination identifies a group of students that are more impaired on reading and reading related tasks than other methods of SLD identification (Speece & Case, 2001). Based on this research, it is believed that using both ROI and level of performance will identify a group of students who are significantly impaired compared to their peers and differentiate students determined to be eligible for special education from those determined to be ineligible and those not referred.

Research Question 2: What student attributes best predict eligibility for special education using a dual discrepancy approach? A multinomial logistic regression will be used with ROI, level of performance, sex, race, free and reduced meal status, and grade as the independent variables. The dependent variable is the students' status in relation to their referral for or identification of SLD (eligible, ineligible, not referred). It is hypothesized that ROI and level of performance will predict student categorization as eligible, not eligible, and not referred significantly more than sex, race, free and reduced meal status, and grade. As previously stated, significantly poor achievement compared to local standards best explains the identification of students as SLD (Peterson & Shinn, 2002). Research has also shown that using ROI results in increased sensitivity and specificity when determining SLD (Fuchs, Compton, Fuchs, Bryant, & Davis, 2008) and more accurately discriminates responders from non-responders (Fuchs, Fuchs, & Compton, 2004). Additionally, using CBM data as the basis for special education referrals results in a group of students that better meet established eligibility criteria and are less biased in terms of sex and other characteristics (Marston, Mirkin, & Deno, 1984). Based on this research, it is believed that level of performance and ROI will explain student classification as eligible, ineligible, or not referred more than other student characteristics.

Assumptions

It is assumed that the CBM-R data used in this study were collected in accordance with the standardized administration procedures. In regard to the variable of eligibility, it is acknowledged that SLD is a psychological construct with no definitive test or markers on which to base eligibility determinations on. Therefore, this variable is not based on a definitive test for SLD but rather on decisions made by multidisciplinary (MDE) teams based on the data that were available to them. It is also assumed that the MDE teams making the eligibility decisions being

used for this research made appropriate decisions according to established practices in the field. However, research has shown 52% to 70% of the decisions reached by MDE teams do not correspond to the state and federal criteria for SLD (MacMillan & Speece, 1999).

Limitations

The limitations of this study include the variability of how each participating school district is implementing MTSS and RTI. Even if it is assumed that all of the schools used for data collection are indeed implementing MTSS and RTI with fidelity, there is no way to ensure that all schools have operationalized RTI the same way. Schools may differ on aspects of MTSS/RTI implementation such as instructional conditions and interventions used at each tier, the quality of the CBM-R passages used for progress monitoring, the duration of progress monitoring schedules required before making eligibility decisions, etc.

Definition of Terms

Dual-Discrepancy Model: The dual-discrepancy model is a process originally proposed by Fuchs and Fuchs (1998) as a means for determining students who were non-responsive to instruction and intervention and for the identification of learning disabilities using curriculum-based measures. For the purpose of this study, the dual discrepancy refers to a decision-making framework using both a student's level of performance and growth rate to determine adequate intervention response.

Level of Performance: Level of performance refers to the level of academic functioning as measured by CBM-R data. For this study, level of performance will be operationalized as the median performance on the final three CBM-R data points from a student's progress monitoring data. Using the median of the final three data points helps to account for measurement error caused by non-equivalent forms of the probes used to monitor student progress.

Growth Rates: Growth rates or growth refers to the measurement of student improvement over time as determined by comparing performance at one point in time to performance at a previous point in time.

Multi-Tiered Systems of Support (MTSS): An MTSS is a framework for meeting students' needs, either academic or behavioral, using universal supports and interventions of increasing intensity to match the severity of student needs as they progress through a hierarchy of tiered support. MTSS uses evidence-based practices at the the universal and targeted level and uses assessment and data-based decision making to create a comprehensive system of prevention and intervention designed to improve student outcomes (Stoiber, 2014).

Oral Reading Fluency: ORF is the number of words read correctly per minute on CBM-R probes.

Rate of Improvement (ROI): ROI refers a mathematically calculated trend line or slope that summarizes student gains in performance per some increment of time.

Response to Intervention (RTI): RTI is the process of measuring a student's response to research-based, academic interventions for the purpose of making decisions regarding intervention effectiveness, movement across tiers, and special education eligibility decisions for SLD. RTI systems are embedded in an MTSS that provides evidence-based curricula to all students with interventions of increasing intensity to match student needs/deficits. This is coupled with assessment, progress monitoring, and data-based decision making.

Tier 1: Within the context of a three-tier MTSS, tier 1 refers to the use of high-quality, scientifically-based curricula made available to all students within the general education setting and delivered by the general education teacher (Johnson, Mellard, Fuchs, & McKnight, 2006).

Tier 2: Within the context of a three-tier MTSS, tier 2 refers to supplemental, targeted instruction provided in addition to tier 1 instruction in a small group setting for students who are not meeting tier1 standards of learning (Johnson, Mellard, Fuchs, & McKnight, 2006).

Tier 3: Within the context of a three-tier MTSS, tier 3 refers to the most intensive level of supplemental instruction provided within a three-tiered RTI model for students who are not making sufficient progress in tier 2. Students who are not successful in tier 3 may be considered for special education eligibility (Kovaleski et al., 2013).

CHAPTER 2

REVIEW OF LITERATURE

The passage of the Education for All Handicapped Students Act in 1975 was the culmination of years of research and advocacy concerning the education of students with specific learning disabilities (SLD). This law provided the federal definition of SLD and established the system of special education in the United States. Later in 1977, the federal government introduced guidelines for the identification of students as SLD using the ability-achievement discrepancy approach. Controversy has surrounded the definition and identification of SLD since its inception (Fletcher, Lyon, Fuchs, & Barnes, 2007).

Per the 1977 recommendation, students who exhibited a severe discrepancy between academic achievement and intellectual ability in any one of seven areas met the eligibility criteria for SLD (U.S. Office of Education, 1977). Concerns with the validity of the ability-achievement discrepancy model in the coming years led to research examining this issue. Hoskyn and Swanson (2000) and Stuebing et al. (2002) conducted meta-analyses of this research. They concluded there were no meaningful differences in reading and reading-related processes between students who displayed a discrepancy between intellectual functioning and academic achievement and those students with no discrepancy (low ability and low achievement).

The 2004 revisions of Individuals with Disabilities Education Act (IDEA) made allowable for the first time the use of “a process based on the child’s response to scientific, research-based intervention” to determine eligibility for SLD (300/D/300.37/a). The 2006 revisions to IDEA 2004 defined four criteria central to determining SLD eligibility. This includes two inclusionary criteria: inadequate achievement relative to age or grade-level

standards and either a discrepancy between intellectual functioning and achievement or a lack of response to scientifically validated instruction. The latter of these two becomes central in determining SLD eligibility using response to intervention (RTI). The other two criteria are exclusionary criteria which require that the presence of visual, motor, or hearing problems, intellectual disabilities, an emotional disturbance, cultural or environmental issues, limited English proficiency, and lack of appropriate instruction are determined not to be the cause of the inadequate achievement. These exclusionary criteria are part of the comprehensive multidisciplinary evaluation process and must be addressed regardless of which evaluation method is being used (Kovaleski et al., 2013).

Multi-Tiered Systems of Support and Response to Intervention

Multi-tiered system of support (MTSS) is a collection of evidence-based practices including the problem-solving process, behavior assessment, CBM, formative assessment, and scientifically-based instruction and intervention. A key and defining element of both RTI and MTSS is the use of interventions that increase with intensity and measurement frequency as the intensity of student needs increase (Reschly & Bergstrom, 2009). RTI, however, specifically refers to the system of measurement embedded within an MTSS that is used to measure students' response to empirically-based interventions and make decisions regarding intervention effectiveness, movement across tiers, and potentially eligibility determinations for SLD (Torgesen, 2007).

Ensuring that all students have access to high-quality curriculum and instruction is the main feature of Tier 1 of an MTSS system. Other central features at Tier 1 include universal screening for students at risk for academic failure, data analysis designed to inform instruction and improve outcomes for all students, and the use of evidence-based instruction for core

curriculum (Johnson et al., 2006). At the advanced tiers, the intensity and length of instruction increase as well as the frequency that student progress is monitored. Specific goals and objectives for desired levels of student performance are developed and assessment data are used to identify students' needs and monitor their progress in response to instructional efforts. Although different MTSS models may make use of a different amount of tiers, a three-tiered model is common. When RTI is used for SLD eligibility decisions, the lack of progress in response to scientifically-validated instruction, the second of the four federal criteria needed for SLD eligibility decisions, becomes central to the eligibility process (IDEA, 2006).

The use of MTSS at various stages in the special education process is not as new as the 2004 revisions of IDEA. One of the earliest proposed methods for using RTI as part of special education eligibility decisions is from the work of Germann and Tindal (1985). Termed the Total Special Education System, they outlined a process that involves the use of curriculum-based measurement (CBM) data at the universal and individual level to identify and intervene with students at risk for academic difficulties, and potentially identify students with SLD. In the Total Special Education System, CBM data are used to develop local norms in various academic domains. Baseline data are collected when students are referred for problems. The median of this baseline data were then compared to the median performance of same-aged peers within the district by calculating a discrepancy ratio. To do this, the referred student's level of performance is divided into the expected level of performance. The value indicates how far below peers the student is (e.g., x 1.5 discrepant from peers/ expected performance). A theoretically-derived discrepancy ratio of x 2.0 or greater is considered the eligibility criteria for special education services in the proposed model. Although the process is outlined in the article, Germann and Tindal do not go on to provide empirical support for their model but do demonstrate that the use

of CBM data as the basis for referrals results in a group of students being referred that is less biased than teacher referral methods.

Research shows that MTSS results in improved student-level and systems-level outcomes and leads to fewer students being identified as SLD than traditional refer-test-place models (Burns et al., 2005). However, few studies address the use and validity of RTI for special education eligibility purposes. As a result, critics of RTI have cited lack of sufficient research into key aspects of using RTI for identifying SLD as a main area of concern with using RTI for making special education eligibility determinations (Reynolds & Shaywitz, 2009a: 2009b). Specifically, they argue that there is lack of clear criteria for determining non-responders, measurement issues with determining student growth, and a lack of clearly defined standards against which to compare a student's intervention response. Other criticisms include that a lack of procedural guidance from the government concerning RTI implementation will result in inconsistencies among students identified as SLD between different states. Zirkel and Thomas (2010a, 2010b) have documented the slow response from state educational agencies in providing procedural guidance for the use of RTI in SLD eligibility determinations and Flinn (2014) has documented inconsistencies in key aspects of RTI implementation and decision making in the available procedural guidance provided to date. However, these criticisms are also true of ability-achievement models of SLD identification (MacMillan & Speece, 1999). Even with these shortcomings, RTI has superior empirical support when compared to ability-achievement discrepancy models. One area where further research is still needed, particularly as it relates to using RTI for SLD identification, is in determining what is a non-responder and how poor of a response is required before identifying a student as having an SLD.

Research on Non-Responders

Although lack of response to scientifically-validated intervention is a key feature in determining SLD, few studies have examined this facet of RTI. Those that do have failed to provide clear decision rules for practitioners to implement. As a result, there are no firm empirical guidelines for determining what is or is not an adequate response to an intervention as noted by Reynolds and Shaywitz (2009a: 2009b). Nor is it clear how poor an intervention response should be to meet the eligibility criteria for SLD. The available literature has produced multiple methods for measuring intervention response including the following methods: split median, dual discrepancy, final normalization, final benchmark, and slope discrepancy (Fuchs, Compton, Fuchs, Bryant, & Davis, 2008). Research has converged on the dual-discrepancy model as being the most technically adequate and valid method for determining intervention response. However, research to date has still failed to adequately define what a non-responder is and how poor of a response is indicative of SLD.

Vellutino et al. (1996) conducted one of the earliest studies examining intervention response. In this study, a sample of students was identified in November of first grade as manifesting the early signs of reading problems based upon teacher ratings and a performance at or below the 15th percentile on either the Word Identification or Word Attack subtest of the Woodcock Reading Mastery Tests- Revised (WRMT-R). The students then received daily tutoring and were administered the WRMT-R several times over the course of the study. The data were used to estimate slopes for student growth using a linear regression analysis. The slopes were then rank ordered and divided into four groups based on performance: very limited growth, limited growth, good growth, and very good growth. Using a split median method, the lowest performing two groups, or those performing below the median, were deemed non-

responsive. Although the purpose of this study was to investigate cognitive correlates of difficult-to-remediate reading problems, it represents an early attempt to distinguish responders from non-responders.

The dual-discrepancy model originally proposed by Fuchs and Fuchs (1998) is perhaps the most comprehensive and conceptually complete model for determining intervention response and eligibility for SLD using RTI. In this model, only students who are below peers in level of performance and growth are deemed non-responsive to instruction. The original model involves four phases. Phase I is conducted at the classroom level. During this part of the process, all students were repeatedly assessed with CBM. The entire class' response to the instructional environment was determined to verify that generally effective instruction was being provided at the classroom level. Phase II then used these data to identify students whose response to the instructional environment was significantly below peers in both level of performance and growth. Phase III introduced modifications and adaptations to core instruction and then monitored the student's response to these changes. If the response was determined to be inadequate, the student moved on to Phase IV which was a trial period of special education. In the original model, only students who showed a response to the trial period of special education instruction were then determined eligible for special education. If the effectiveness of special education could not be established, the authors argued that there was no basis for removing the student from general classroom instruction.

Fuchs and Fuchs (1998) also provided some preliminary evidence for the validity of the dual-discrepancy model and guidelines for determining adequate intervention response. CBM math data for 469 students were examined. The use of dual-discrepancy criteria identified a group of students who were the lowest performing in their class. Adding the criterion for growth

eliminated 29-71% of the students deemed non-responsive by using performance level alone. Using both growth and level of performance also identified a group of students consistent with the percentage of students in the school already identified as SLD. Additionally, the percentage of students deemed nonresponsive using the following criteria was compared to estimated prevalence rates for SLD: discrepancies of 1.5 standard deviations on both slope and performance level, discrepancies of 1.5 standard deviations either slope or performance level, discrepancies of one standard deviation on either slope or performance level, and discrepancies of one standard deviation on both slope and performance level. The authors concluded that a discrepancy of one standard deviation on both slope and performance level was best for identifying non-responders. Although the criterion of 1 standard deviation in both slope and performance level had higher rates of false positives than the other criteria examined, this was deemed acceptable within the context of this model because these criteria were for determining which students needed to move onto the next stage of intervention (from Phase I and II to Phase III) as opposed to being identified for learning disabilities. The criterion of one standard deviation below peers in level of performance and growth also identified a group of students (9.4%) that was very similar to national prevalence rates for SLD (8.7%).

Research conducted by Torgesen et al. (2001) produced another method for determining intervention responsiveness called final normalization. This study was designed to compare two instructional approaches on the reading outcomes of severely disabled readers. One instructional approach placed more time and emphasis on phonemic awareness and the other placed more time and emphasis on phonics. Students received daily, individual tutoring for two, 50-minute sessions per day for 8 to 9 weeks and a 50-minute generalization session for the following 8 weeks. Students who achieved a standard score of 90 or greater on various measures of word

reading, reading fluency, and comprehension at the end of the intervention period were deemed responsive to the instruction. Although the intention of this study was to determine which of two instructional approaches resulted in better reading outcomes for students with reading disabilities, it provides another prospective method for distinguishing responders from non-responders within an RTI model.

Good, Simmons, and Kame'enui (2001) provided the basis for the final benchmark approach for determining intervention response. This study examined the predictive power of established benchmarks of proficiency for fluency-based measures of early literacy skills and CBM-R. Short-term, longitudinal analyses of CBM-R data for four cohorts of students in grades kindergarten through third-grade were conducted using fluency-based measure of early reading skills and oral reading fluency (ORF) probes. The results supported that benchmarks of proficiency for early literacy skills and curriculum-based measurement in reading (CBM-R) predicted continued success in reading development and proficiency on a high-stakes state accountability assessment. Again, the purpose of this study was not to examine adequate intervention response. However, it provides an evidence-based target for proficiency when providing intervention for struggling readers.

A fifth way of determining adequate intervention response is the slope discrepancy method (Fuchs, Fuchs, & Compton, 2004). When using the slope discrepancy method, a student's growth is depicted using a line of best fit from data that are graphed in a time-series manner. The slope of this line is compared to some predetermined benchmark of adequate growth. Students who fail to meet this performance benchmark are considered non-responsive.

Comparison of Different Methods for Determining Intervention Response

Research has shown that using different methods for determining intervention response results in different groups of students being identified as non-responders. Barth et al. (2008) examined the agreement of different operationalizations of RTI and its impact on identifying responders and non-responders. Differences in methods of determining response or non-response, the cut points used to differentiate between responders and non-responders, and the measures used to determine responders and non-responders were compared. The methods examined for determining intervention response included the dual-discrepancy model, slope discrepancy, and final benchmark, and normalization approaches. Agreement among the different methods for determining intervention response was generally poor. Methods that used a final benchmark/normalization approach tended to identify more non-responders than methods that incorporate a measure of growth. The use of fluency benchmarks alone identified the highest incidence of non-responders. The superiority of one method over another was not demonstrated (e.g., dual discrepancy versus slope discrepancy). In the instances where agreement was high, this tended to be driven by agreement on determining responders as opposed to non-responders.

Brown-Waesche, Schatschneider, Maner, Ahmed, and Wagner (2011) also examined the agreement of several models for determining learning disabilities: the traditional ability-achievement discrepancy model and three RTI models. The RTI models included the final normalization, slope discrepancy, and dual-discrepancy approaches. The chance-corrected affected-status agreement statistic was used as opposed to Cohen's Kappa to reduce the effect of agreement in determining adequate readers on the overall level of agreement. Since agreement on adequate readers has a higher base rate and makes up the majority of the samples examined,

Cohen's Kappa tends to overestimate agreement on non-responders, or in this case, disabled readers. Agreement was strongest among the dual-discrepancy method and the other two RTI methods with poorer agreement was observed between final normalization and slope discrepancy. Longitudinal stability was poor for all models (affected-status agreement rates of below 50%).

Of the various methods for determining intervention response, the dual discrepancy approach has the most empirical support. Speece and Case (2001) compared the validity of the dual-discrepancy model proposed by Fuchs and Fuchs (1998) to a traditional ability-achievement discrepancy and simple low achievement methods for identifying SLD. Low achievement was defined as a standard score of less than 90 on the Basic Reading Skill Cluster of the Woodcock-Johnson Psychoeducational Battery-Revised (WJ-R). In this study, a sample of first and second grade students at risk for reading failure were identified. Of this group, students who met the criteria of the dual discrepancy (1 SD below peers in slope and level of performance) were compared to a group of students who were ability-reading achievement discrepant and a group of students with just low achievement in reading. The results supported the validity of the dual-discrepancy model as it identified a group of students that was consistent with estimated prevalence rates (8.1%), who were more impaired on tasks of phonological processing and teacher ratings, and who were more representative of the general population than the other two methods.

Speece, Case, and Malloy (2003) conducted several analyses extending the research of the previous study conducted by Speece and Case (2001) and provided further empirical validation of the dual-discrepancy model. They conducted a three-year study in three suburban schools in the Mid-Atlantic United States. Children were identified as at risk based on a

performance that was below the 25th percentile on a Letter Sound Fluency probe (fall of first grade) or an ORF probe (fall of second grade). A purposive sample of children performing between the 30th and 90th percentiles was used as the comparison group to determine discrepancy status. All children were administered the Basic Reading Skills cluster of the WJ-R and an abbreviated battery from the Wechsler Intelligence Scale for Children-Revised to estimate Full-Scale IQ (FSIQ). The at-risk group was administered measures of phonological processing, word reading efficiency, and teacher behavior rating scales for problem behavior and academic competence. The validity of the dual-discrepancy model was again supported as dual-discrepant students, at-risk students, and students from the purposive sample differed in terms of reading scores and performance on measures of phonological processing, word reading efficiency, and ratings of academic competency with the dual-discrepant group being the most impaired. Also, the dual-discrepancy model identified a sample of students more consistent with the general population in terms of age, gender, and race.

In a second set of analyses, students who were identified using the dual-discrepancy criteria were given individualized programming which was periodically reviewed throughout the course of the study. Children who were persistently dual discrepant despite intervention (met the dual-discrepancy criteria four or more times over the course of the study) performed worse on measures of reading, phonological awareness, and ratings of academic competence and were rated as having more problem behaviors than those students who were infrequently or never dual discrepant (three or less times). These students also demonstrated slopes consistent with or worse than typical special education students from the Deno et al. (2001) study. This study also provided evidence that the dual-discrepancy model improves outcomes for at-risk students. Students who received full implementation of the dual-discrepancy model (e.g., intervention

following dual-discrepancy status and periodic review) performed better than control groups on reading level and growth at three years follow up.

The previous studies established that the dual-discrepancy model is a valid method for applying RTI to the identification of SLD. Fuchs, Fuchs, and Compton (2004) provided further support for the dual-discrepancy model by establishing that it is psychometrically superior for the purpose of identifying SLD than other methods for determining intervention response. They compared different methods for determining intervention responsiveness on the following variables: the number of students identified as non-responsive and how well each method differentiated non-responders in terms of growth and final outcomes. In this study, a sample of first and second grade students were screened to determine at-risk status. For one semester, Peer Assisted Learning Strategies was implemented in the general education environment. The progress of these at-risk students was monitored to determine their response to this scientifically-validated intervention. At the end of the semester, students were identified using a dual-discrepancy model based on a performance that was substantially lower than peers (.5 standard deviations below the reference group in slope and performance level). These students then received intensive tutoring and had their progress monitored.

In first grade, four different methods for determining intervention response were evaluated: split median using nonsense-word fluency, split median using Dolch high frequency word lists, final normalization based upon a word reading standard score of less than 90 on the WRMT (Word Identification and Word Attack), and the final benchmark of 40 words correct per minute using ORF passages from the Dynamic Indicators of Basic Early Literacy Skills (DIBELS). Median split using the Dolch high frequency words proved to be the best method, discriminating responders and non-responders on all measures of growth and outcomes. In

second grade, five methods of determining intervention response were again compared. These methods included split median using the WRMT word-reading gain scores, split median using CBM slope, final normalization defined by a word reading standard score of greater than 90 on the WRMT, slope-discrepancy using CBM-R (1.5 words per week was the benchmark), and dual-discrepancy using CBM-R and normative criteria (1.5 words per week and final CBM normative benchmark). Slope-discrepancy and dual-discrepancy performed the best at second grade. The results found that slope-discrepancy differentiated responders from non-responders on three of five outcome variables and two of five growth variables. The dual-discrepancy method differentiated responders from non-responders on two of five outcome variables and three of five growth variables. Of all the methods examined, the slope-discrepancy and dual-discrepancy methods performed the best.

Extending the work of Fuchs, Fuchs, and Compton (2004), Fuchs, Compton, Fuchs, Bryant, and Davis (2008) conducted a large scale, longitudinal study examining how responsiveness to intervention should be defined. Five methods for determining intervention responsiveness were evaluated including the median split, final normalization, final benchmark, slope-discrepancy, and dual -discrepancy methods. Each method was analyzed for its effectiveness in identifying non-responders. The groups of students identified by each method were analyzed in terms of the sensitivity and specificity of the method and how well the methods identified an appropriate group of non-responders that could be considered reading disabled compared to estimated prevalence rates. Acceptable levels of sensitivity and specificity were set at 0.80. The results showed that final normalization identified an acceptable number of students but had mixed hit rates, sensitivity, and specificity. Final benchmark and median split methods over-identified SLD compared to estimated prevalence rates. The slope-discrepancy and dual-

discrepancy methods over-identified SLD but did result in acceptable hit rates, sensitivity, and specificity. Although none of the RTI methods performed adequately, the final normalization, slope-discrepancy, and dual-discrepancy methods were deemed promising and outperformed ability-achievement discrepancies.

Assessing Student Growth Using Rate-of-Improvement

The measurement of student growth and its application to SLD eligibility decisions is unique to RTI decision-making frameworks and is a central feature of the dual-discrepancy model. Traditional ability-achievement discrepancy models only consider a performance at one moment in time. However, the dual discrepancy approach emphasizes how much a student grows over time in response to instructional efforts in addition to their level of performance. The dual-discrepancy model typically makes use of CBM for monitoring student progress, determining level of performance, and determining growth in response to instructional efforts. CBM is well suited for use with the dual discrepancy framework for the following reasons: it has adequate reliability and validity, it can be repeatedly administered and displayed graphically, and peer-referencing or norm-referencing can be used to determine a student's level of performance (Deno, 1985). CBM also has increased sensitivity to student growth compared to traditional, norm-referenced achievement tests (Marston, Fuchs, & Deno, 1986). CBM data can differentiate students with SLD from low achieving students and typically performing students as well (Shinn & Marston, 1985; Shinn, Ysseldyke, Deno, & Tindal, 1986).

CBM data that are graphed in a time-series manner can be used to measure student growth per some interval of time, usually per week, which is also called rate of improvement (ROI). A review of the literature reveals multiple ways to summarize ROI when graphically displaying CBM data. These methods usually fall into one of two categories: either simple line-

fitting methods or mathematically calculated trend lines (Parker & Tindal, 1992). Research by Shinn, Good, and Stein (1989) compared the accuracy of the split middle and ordinary least squares (OLS) trend lines for estimating future student performance. Student CBM-R data were collected for 36 weeks over the course of one school year. Trend lines were drawn for each method for 10 weeks, 20 weeks, and 30 weeks of data collection and used to predict student performance at 2, 4, and 6 weeks into the future. These predicted performances were then compared to the student's actual performance. OLS trend lines were found to be more precise across both conditions-number of data points and length of prediction.

Norms developed for CBM help provide a framework of reference or benchmark for student performance when using the dual discrepancy framework. Students can be classified according to level of deficiency by comparing their performance to a set of norms. In this application, norms are used to set goals for desired level of performance and as a reference to compare student growth. This can be accomplished by graphing an aimline from the student's current level of performance to the desired level of performance, plotting student data, and comparing the two (Hasbrouck & Tindal, 2006). A similar process can be used incorporating ROI. Using the OLS regression, Fuchs, et al. (1993) established ROIs for a sample of first through sixth grade students under typical instructional conditions over the course of two years using CBM in reading, math, and spelling. This research provides a benchmark of comparison for interpreting ROI at each grade level. Nese et al. (2013) have also investigated ROI and its developmental trajectory across grades. Overall, yearly ROIs tended to be curvilinear for grades 1 to 7, and grade 8 generally showed no growth overall. Longitudinally, ROI shows a steep increase from grades 1 to 3, with a sharp decrease in ROI in grades 4 and 5, and stable growth from grades 5 to 7.

Although this research provides empirically-based benchmarks of comparison for typically developing students when examining ROI, much less is known about what ROI at the individual level is indicative of poor intervention response or disability status. Deno, Fuchs, Marston, and Shin (2001) used an OLS regression to compare growth rates in reading for a sample of first through sixth grade students receiving typical instruction versus a sample of students identified with learning disabilities receiving typical special education programming. They found that students in typical special education programs obtained growth rates of less than half of that of their peers in regular education. This research supports that ROI is sensitive enough to differentiate typically developing readers from students with SLD.

Some research is beginning to examine what ROIs are indicative of lack of intervention response. Christ, Zopluoglu, Monaghan, and Van Norman (2013) investigated decision-making thresholds for ROI that could be used to identify non-responders in an RTI system. Data were collected on a large sample of second and third grade students receiving supplemental, Tier-2 instruction. Simulation data were used to estimate slopes. Data were simulated for weekly and intermittent progress monitoring schedules with durations that ranged from 2 to 20 weeks. Cut points were assigned for the 50th, 20th, and 15th, percentiles based upon standards obtained from the National Center for Response to Intervention. These percentiles correspond with slopes of 1.50, 0.96, and 0.84 respectively. Decision cut points found to have the best accuracy were 1.50, 1.10, and 1.00 respectively. That is, the actual cut points of 0.96 and 0.84 were best predicted by slopes of 1.10 and 1.00 when the various criteria associated with decision accuracy, such as sensitivity and specificity, are balanced. Although this research provides some preliminary guidance on non-responders at Tier-2, it does not address non-response at Tier 3 and what slopes are indicative of SLD.

Passage Equivalency

In addition to not having clear criteria to guide decision making with ROI at Tier-3 and beyond, research on passage equivalency has called into question the use of intra-student measurement with CBM-R and ROI when making high-stakes decisions in RTI frameworks (Wayman, Wallace, Wiley, Ticha, & Espin, 2007). One particular concern is that the alternate forms of CBM-R that are used for progress monitoring can vary significantly in difficulty level. Jenkins, Zumeata, and Dupree (2005) conducted a study examining the passage variability of CBM-R probes as well as its effect on measuring student growth. A sample of students ranging from grades 2 to 10 were administered CBM-R probes. Ten weeks later, they were administered the same four CBM-R probes and an additional four novel CMB-R probes. Five weeks following the second administration, they were administered two randomly selected probes from the set of four previously administered and two additional novel probes. Significant variability across alternate forms of CBM-R was found. There was an average difference of 13.74 words correct between mean performance on all four probes when compared to a specific passage from that same probe set at the first administration. This difference was 15.13 words correct for the second administration. The standard error for the group ranged from 6 to 8 words per minute, depending on which probe was examined. The authors also found that progress monitoring with the same passage significantly reduced this error and did not result in memory effects at ten and five weeks between administrations.

Poncy, Skinner, and Axtell (2005) also conducted a study investigating the variability in ORF, measured by WCPM, as a function of student skill and passage variability. A sample of third graders was administered five probes for four consecutive days. The results showed that 81% of variability in performance could be attributable to student skill, 10% to passage

variability, and 9% due to other sources of error. Standard errors of measurement (SEMs) were calculated for within-student decisions based upon the number of probes administered and were as follows: SEM = 18 WCPM for one probe, SEM = 10 WCPM for 3 probes, SEM = 8 WCPM for five probes, SEM = 7 WCPM for seven probes, and SEM = 6 WCPM for nine probes. Mean student performance was calculated for each probe. When passage variability was controlled for (only probes with +/- 5 WCPM), the amount of variance accounted for by the student increased to 89% and error attributable to passage variability decreased to 1%. SEMs decreased to 12, 7, 5, 5, and 4 WCPM for 1, 3, 5, 7, and 9 probes, respectively.

Francis et al. (2008) examined the effects of alternate forms of a commercially available set of CBM-R progress monitoring probes frequently used in schools on student performance and growth trajectories. Six passages were randomly selected from the second grade probes from the DIBELS. A sample of second grade students was randomly assigned to one of six groups to read each one of the passages in a varied order. All students read three probes in one sitting for the first week and the median score was selected as the students' initial fluency level, as per the recommended benchmarking procedure. These "benchmark" assessments were analyzed to determine if each passage had an equal likelihood of being the median performance, as would be expected if the passages were equivalent. The results showed significant differences in mean performance as a function of the probe selected. While several of the probes were fairly equivalent, the mean performance ranged from 67.9 WCPM for the most difficult passage to 93.9 WCPM for the easiest passage. The order in which the passages were administered also significantly affected student growth trajectories, as students could obtain either a positive or negative trend line depending on the sequence in which the passages were administered.

Ardoin and Christ (2009) further examined the standard errors associated with commercially available CBM-R probes from DIBELS, AIMSweb, and an experimental passage set developed by the authors. DIBELS and AIMSweb use traditional readability formulas for determining equivalent passage difficulty of alternate forms. The experimental passage set used Euclidian distance for determining equivalent passage difficulty for alternate forms. The results found that the experimental passage set significantly outperformed the two commercially-available passages sets in both the standard error of slope (*SEb*) and the standard error of estimate (*SEE*). AIMSweb outperformed DIBELS because in addition to readability formulas, AIMSweb administered the probes to a small sample of students and removed the outlier probes. However, there were still moderate to strong correlations for *SEb* and *SEE* across the different passage sets despite the observed differences. Further investigation found that standard errors were more stable at lower levels of reading fluency and that more fluctuation in performance were observed at higher levels of reading fluency and when interpreting data at the individual level.

The previous studies have established that CBM-R probes, including the commercially available probes frequently used in schools, have significant problems with reliability across the alternate forms used to monitor student progress. Alternate passages can vary by up to approximately 30 WCPM with passage equivalency being a significant source of error not only when estimating level of performance, but also when predicting student growth. This type of error has the potential to interfere with the decision-making process that is central to the dual-discrepancy model. However, the measurement error can be reduced to acceptable levels when additional measures are taken to exclude passages with unacceptably high levels of measurement error.

The Effect of Passage Equivalency on ROI

The variability of the CBM-R probes used to monitor the progress of students has been shown have a significant impact on calculating ROI. Any measurement used for making high-stakes decisions, such as special education eligibility, needs to have documented and acceptable levels of error. As a result, the standard error of individual student ROI is of central importance when identifying students as SLD using the dual discrepancy framework. Christ (2006) conducted a literature review to identify studies that reported the standard error of estimate (*SEE*) for CBM-R in order to estimate the standard error of the slope (*SEb*). Based on those studies, five levels representing the range of *SEEs* reported were selected ($SEE = 10, 12, 14, 16, 18$ WCPM). An additional four levels of *SEE* were selected to examine an optimal range for *SEE* (2, 4, 6, 8 WCPM). Calculations were based upon the assumption that two CBM-R data points were collected per week on Monday and Thursday. A total of 14 progress monitoring durations were examined ranging from 2-15 weeks. The results showed that longer progress monitoring schedules resulted in lower *SEb*. After two weeks of progress monitoring, the median *SEb* was 9.19 WCPM. After five weeks of progress monitoring, the median *SEb* was 2.21 WCPM. After 15 weeks of progress monitoring, the median *SEb* was reduced to .42 WCPM. Additionally, *SEE* was used to determine the impact of measurement conditions on CBM-R. Optimal measurement conditions (quiet location, consistent administrator, consistent setting, standardized directions, consistent probe difficulty) create a lower magnitude of *SEE*. Poorly controlled measurement conditions can result in up to 4 times more error than better controlled measurement conditions. In summary, more frequent progress monitoring with optimal assessment conditions will over time reduce the amount of error in *SEb* to more acceptable levels that are required for high-stakes decision making.

Although incorporating SEM into the interpretation of ROI is important for high-stakes decision making, it is also problematic because measurement error can exceed what can be reasonably expected of students in terms of ROI. Christ and Silberglitt (2007) analyzed 8 years of CBM-R probes administered triannually. Convergent evidence for strong test-retest reliability was found (ranging .88 to .95, median of .93). However, SEM was found to be within the range of 5 to 15 WCPM across grades (median SEM of 10 WCPM). Due to a variety of factors, such as variability and grade level, most likely estimates range from 5 to 9 WCPM. Although this research does not address ROIs derived from individual progress monitoring, it does raise important questions regarding interpreting ROI in relation to growth rates derived from benchmark assessments within a dual-discrepancy framework. Ardoin and Christ (2008) further examined the slopes derived from triannual universal screenings of CBM-R to evaluate the effect of alternate probe sets on level of performance over the course of the year and to examine the impact of alternate probe sets on growth rates over the course of the year as well. Both the consistent probe set and alternate probe set performed well, although using the same probe set for all three screenings was more reliable. Additionally, estimates of growth when comparing single probes across screening periods were very unstable. The use of a median score and a pooled sample provides the most stable estimates of growth from one screening period to the next. The large samples available to determine growth rates and performance levels derived from triannual benchmark assessments reduce the psychometric issues surrounding CBM-R and individually derived ROIs and make for an adequate comparison group for use within a dual-discrepancy model.

Christ et al. (2012) further investigated the effects of the quality of CBM-R data sets, schedules for monitoring student progress, and the method used for determining growth on the

quality of growth estimates. The quality of the data set was empirically defined as follows: very good (*SEE* of 5 WCPM), good (*SEE* of 10 WCPM), poor (*SEE* of 15 WCPM), and very poor (*SEE* of 20 WCPM). Schedules for monitoring student progress of 6, 8, 10, 12, 14, 16, 18, and 20 data points collected once per week were compared. Additionally, several methods for mathematically calculating trend lines were compared including the OLS, the moving median, and the moving mean methods for determining ROI. The results showed that a minimum of 14 data points collected weekly with a very good or good data set is needed to approximate the reliability standards for high stakes decision making (.90). OLS growth estimates were more precise than those produced via moving median and moving mean. The standard error of the growth estimates using OLS for a very good passage set is +/- 0.32 but is +/- 1.32 for a very poor passage set. The implications of this study are that the equivalency of the CBM-R probes used for monitoring student progress will have a substantial impact on the ability to accurately determine student growth, especially since using passages with very poor equivalency results in an amount of error comparable to or greater than the expected student ROI.

In summary, Ardoin, Christ, Morena, Cormier, and Klingbeil (2013) conducted a review of the professional literature through 2010 for guidelines on interpreting CBM-R data when applied to individual student growth (monitoring student progress and ROI). While the empirical support for the use of CBM-R for screening and benchmarking purposes is robust, there has been little research into the use of CBM-R for current practices for the monitoring of student progress as applied within an RTI system. In particular, the authors sought to determine the following: what recommendations have been made regarding decision rules for monitoring progress with CBM-R and what empirical evidence supports these decision rules? The results found that trend line decision rules, particularly the use of OLS, were recommended over data-point decision

rules (e.g., split-median, Tukey method). However, the evidence supporting both methods is poor. The most common recommendation for the number of data points needed to make a decision was seven, although the majority of the studies reviewed suggested that a minimum of 10 data points were needed to guide educational decisions. No studies examining the accuracy of decisions using these rules were located. Recent research has also revealed that many more data points and a longer duration for the monitoring of student progress are needed before low-stakes and high-stakes decisions can be reliably made using individual student ROIs. Despite these concerns, research by Burns et al. (2010) found that the group of students identified as non-responsive using an OLS trend line within a dual-discrepancy model were not significantly altered when the SEM of the CBM-R probes was accounted for.

Using RTI for Determining SLD Eligibility: Current Recommendations, Unanswered Questions, and Future Direction

Research has converged on the dual-discrepancy model being the most technically adequate and valid method for determining intervention response and identifying SLD within an RTI framework. Legislation has also embedded the dual-discrepancy model into the federal regulations for identifying SLD using RTI. The inclusionary criteria included in IDEA 2006 state that a student must demonstrate both inadequate achievement and inadequate response to scientifically-validated intervention when using RTI for SLD identification (IDEA, 2006). Kovalski et al. (2013) have provided perhaps the most comprehensive model for determining SLD eligibility using RTI and the dual-discrepancy model. In this model, national norms derived from triannual screenings with CBM-R data are used as the comparison group when determining dual discrepant status for level of performance and ROI. A gap analysis is recommended for analyzing and summarizing student level of performance and ROI within the

context of SLD eligibility decisions. To conduct a gap analysis of performance level, the student's current level of performance is summarized as a percentile or as a ratio comparing that student's current performance level to the desired level of performance. When using percentiles, a performance at the 10th percentile or lower was recommended. When using ratios, a performance that is 50% or less of the desired level of performance was recommended. Students who demonstrate this theoretically-derived level of deficiency meet the first of the inclusionary criteria referring to inadequate achievement. When conducting a gap analysis for ROI, the benchmark ROI, target student aimline, and the attained ROI of the target student must all be analyzed to determine whether or not the target student's current level of progress is sufficient to meaningfully close the gap between the current and the desired level of performance. To do this, it is recommended that the student's ROI relative to normative standards be examined. Additionally, the student's attained ROI should be compared to the ROI that is needed for the student to successfully close the gap between the current and desired level of performance. The analysis of this trajectory and the likelihood that the student will close the performance gap within a reasonable amount of time with the resources available becomes the basis for determining the second inclusionary criteria of an inadequate intervention response.

Marston, Muyskens, and Canter (2003) provided some empirical support for what level of performance corresponds to SLD eligibility. Analyzing implementation of the Minneapolis Public Schools Problem-Solving Model, they found that a criterion of a 2.0X discrepancy ratio for level of performance using CBM-R resulted in no more students being identified as SLD than previously used practices. This corresponds to the recommendation from Kovaleski et al. (2013) of using a performance level that is 50% of the benchmark standard. However, as Kovaleski et al. acknowledge, there is no research on how deficient a student's ROI should be when

determining SLD identification and thus advocate for analyzing student trajectories using ROI as opposed to an established cut point.

Lack of clear decision rules for what ROI constitutes a disability is problematic when trying to identify a consistent group of students as non-responsive or SLD. According to Barth et al. (2008), decision rules surrounding what ROIs are used to demarcate non-responders have a significant impact on the groups of students identified as non-responsive. In that study, ROIs were derived from four administrations of the Test of Word Reading Efficiency (TOWRE) and 11-13 administrations of reading fluency probes from the Continuous Monitoring of Early Reading Skills (CMERS) and were calculated using a linear growth model. Cut-points examined included 0.5, 1, and 1.5 standard deviations below the mean ROI when compared to a sample of typically achieving students from the schools in which the study was conducted. Cut-points for post-intervention status included the 30th percentile for the TOWRE and 40 WCPM on the CMERS (which corresponds to the 35th percentile from DIBELS). A total of 808 different combinations of the above criteria were examined. Differences in cut-points for both ROI and level of performance were shown to be the most significant factor in determining whether a student was responsive or not.

Burns and Senesac (2005) also examined cut points within a dual-discrepancy model and have offered some preliminary guidance on what levels of growth best correspond to SLD. They compared different levels of slope and its effects on identifying non-responsive students within a dual-discrepancy model. Four levels of slope were compared as the criteria for non-responsiveness: 25th percentile, 33rd percentile, 50th percentile, and one standard deviation below the mean. Data were collected using DIBELS CBM-R probes collected midyear and at the end of the year for at-risk students receiving Tier 2 intervention. End-of-year scores were compared

to the Grey Oral Reading Fluency Test, Fourth Edition (GORT-4) for those students deemed responsive and non-responsive to determine which level of growth best differentiated responders from non-responders. Growth estimates were derived using end-of-year performance minus mid-year performance on CBM-R probes. Growth scores were rank ordered and assigned percentile ranks. The results showed that each of the three percentile groups differentiated dual-discrepant students and not dual-discrepant students, but that the one standard deviation condition did not. Students identified as dual discrepant using the 25th and 33rd percentiles for growth were most consistent with estimated prevalence rates for SLD. However, it should be noted that the procedures in this study included a last-minus-first approach to measuring growth, which is different than the contemporary practice of determining ROI using OLS.

In addition to the lack of empirically-based guidelines, many states have provided inconsistent or insufficient guidance for practitioners using RTI to identify SLD. Flinn (2015) conducted an analysis of state regulatory and guidance documents pertaining to the use of RTI for determining SLD eligibility. In terms of monitoring student progress, the use of direct measurement of skills and establishing a norm group for comparison was the most frequently required practice. The use of CBM, the frequency of the progress monitoring schedule, and using an OLS method for determining ROI are mentioned, but not required practices in some states. In general, most states do not require many of the empirically-based recommendations regarding the monitoring of student progress. In states that mandate RTI, the vast majority require the monitoring of student progress but only recommend the use of CBM. Furthermore, they do not provide guidelines for the frequency and duration of the monitoring of student progress at the advanced tiers nor how many data points are required for a trend line and high-stakes decisions. Finally, most states, including those that mandate RTI, do not provide

guidelines for the magnitude of deficiency required for SLD. As previously mentioned, research has shown that different measures, different frequencies for the schedules of monitoring student progress, and the number of data points used to calculate a trend line can all have a substantial impact on ROI and its technical adequacy as well as significantly altering the decisions made using the data.

Previous research on eligibility decisions using discrepancy models found that a student's level of performance compared to local standards best predicted eligibility for SLD regardless of what model was used or whether the student met the established criteria for that model (Peterson & Shinn, 2002). Preliminary examination of the dual-discrepancy model and eligibility decisions have found similar results. Boneshefski (2017) conducted an analysis examining the impact ROI had on predicting SLD eligibility decisions within an RTI framework. The results showed that level of performance was the most significant factor when making eligibility decisions for SLD. Further analysis using a hierarchical logistic regression analysis found that the addition of ROI added negligible predictive power for classifying students as SLD. However, students determined eligible for special education and those not referred did show statistically significant differences in ROI. Even though research has shown that the addition of ROI adds to the validity of the group of students identified as non-responsive and potentially SLD (Speece & Case, 2001), it seems that in practice the students with the lowest level of performance also have the lowest ROIs, and that multidisciplinary evaluation teams are still identifying a group of students whose academic functioning is most impaired when compared to local standards with growth adding little additional utility in classification. This idea is supported by research conducted by Silberglitt and Hintze (2007) that showed that students whose level of performance using CBM-R was at the bottom of the distribution demonstrated

significantly lower ROIs in general compared to more typically developing readers. Additionally, Stuebing et al. (2014) found a similar relationship between measures of phonological awareness/processing and low achievement in reading. That is, the lowest performers on reading measures were the most impaired in terms of phonological processing. It is possible that students whose reading level is the lowest compared to peers, also show the poorest ROIs, even after intervention, and that level of performance is the main variable of concern.

Vaughn, Linan-Thompson, and Hickman (2003) conducted a study examining the use of RTI as a method for identifying students with learning disabilities in reading with an emphasis on final level of performance without much consideration for ROI. Forty-five second-grade students at risk for reading failure were identified for supplemental intervention using the Texas Primary Reading Inventory (TPRI). These students received 35 minutes of daily, supplemental instruction in groups of three. Pre-established exit criteria were used to determine if students responded adequately to intervention attempts. Exit criteria included a passing score in the TPRI, a median performance of 55 WCPM on a second-grade level passage from the Test of Oral Reading Fluency, and a score of 55 WCPM on second-grade level fluency passages for three consecutive weeks. Students who still did not meet these criteria after 30 weeks of supplemental instruction were deemed non-responders, which would correspond to being eligible for special education. Fewer than 25% of at-risk students did not respond to the intervention. Students who were non-responsive after 30 weeks demonstrated distinctly poorer levels of reading fluency and comprehension than did students who had exited intervention. Given the high correlation between inadequate achievement and inadequate ROI, perhaps this model of RTI and SLD identification may be a simpler, yet just as reliable approach as the dual-

discrepancy model. However, it is unclear how the group of non-responders identified in this study compares to estimated prevalence rates for SLD.

Statement of the Problem

Based on current research and the federally-defined eligibility criteria, a dual-discrepancy approach to the identification of SLD using RTI is the most supportable method at this time. However, research and state guidance thus far has not provided clear guidelines for what ROIs are deficient enough to result in SLD eligibility and require specially designed instruction. As a result, the current recommendation is to use a trajectory analysis of ROI to satisfy the inclusionary criteria pertaining to insufficient growth. The issue of measuring student intervention response is further complicated by the psychometric issues with the ROIs derived from repeated administration of CBM-R probes and the fact that selected cut points for deficiency are highly influential in the group of students identified as SLD. Although previous research has established that measuring both level of performance and growth results in more accurate classification of students as SLD, Boneshefski (2017) did not demonstrate that ROI meaningfully contributes to multidisciplinary evaluation (MDE) team eligibility decisions beyond what is accounted for by level of performance. His results seemed to replicate those obtained by Peterson and Shinn (2002) that suggested low achievement relative to local standards is the main factor in SLD eligibility decision regardless of what eligibility criteria are in place.

The current study aims to add to the research base by examining the following research questions: Do students who are identified as SLD in reading using RTI differ from those students who are determined ineligible for special education and those students receiving supplemental intervention but not referred for evaluation on their ROI and/or level of performance? Does

ROI add to the identification and classification accuracy of students as SLD when using RTI for special education eligibility decisions? The current study will examine this issue and potentially replicate the results obtained by Boneshefski (2017) in one Midwestern state in another Mid-Atlantic state.

Summary

RTI is a research-based alternative to traditional discrepancy models for determining SLD with the potential of reducing the number of students identified and improving the educational outcomes of large numbers of students. In order for RTI to be implemented as a method of determining SLD to its fullest potential, consensus is needed on how to determine inadequate intervention response and how much lack of response is required for identifying students with SLD. The dual-discrepancy model is the most comprehensive and research-based application of using RTI to determine SLD. The greatest area of need when applying this model to school-based disability determinations is the need for consistent, empirically-based guidelines for determining lack of intervention response, particularly for ROI. Future research needs to focus on what ROIs are indicative of a poor-enough intervention response to qualify for SLD.

CHAPTER 3

METHODS AND PROCEDURES

Introduction

Federal law allows for the use of response to intervention (RTI) to make special education eligibility determinations for the exceptionality of a specific learning disability (SLD). However, to date there is no clear consensus as to what constitutes a lack of response or SLD when using this method. The dual-discrepancy model for determining lack of intervention response has the most research support to date (Fuchs & Deshler, 2007). This study examined students identified as SLD using RTI to determine if students identified as SLD using RTI can be effectively differentiated from those students who are determined ineligible for special education and those students receiving supplemental intervention but not referred for evaluation using the dual-discrepancy framework (Fuchs & Fuchs, 1998). It also examined what student attributes best predicted eligibility for special education within a dual-discrepancy framework.

Research Questions and Hypotheses

Research Question 1: Do students who are identified as SLD in reading using RTI differ from those students who are determined ineligible for special education and those students receiving supplemental intervention but not referred for evaluation on their rate of improvement (ROI) and/or level of performance? It was hypothesized that students identified as SLD will have significantly lower levels of performance and lower ROIs than students not identified as SLD and students not referred for evaluation. Peterson and Shinn (2002) found that low achievement relative to local standards best explains which students are identified as SLD, regardless of what eligibility criteria are in place. Research has also shown that using both ROI and level of performance in SLD determination identifies a group of students that are more

impaired on reading and reading-related tasks than other methods of SLD identification (Speece & Case, 2001). Based on this research, it is believed that using both ROI and level of performance will identify a group of students who are significantly impaired compared to their peers and differentiate students determined to be eligible for special education from those determined to be ineligible and those not referred.

Research Question 2: What student attributes best predict eligibility for special education using a dual discrepancy approach? It was hypothesized that ROI and level of performance will predict student categorization as eligible, not eligible, and not referred significantly more than sex, race, free and reduced meal status, and grade. As previously stated, significantly poor achievement compared to local standards best explains the identification of students as SLD (Peterson & Shinn, 2002). Research has also shown that using ROI results in increased sensitivity and specificity when determining SLD (Fuchs et al., 2008) and more accurately discriminates responders from non-responders (Fuchs, Fuchs, & Compton, 2004). Additionally, using CBM data as the basis for special education referrals results in a group of students that better meet established eligibility criteria and are less biased in terms of sex and other characteristics (Marston et al., 1984). Based on this research, it is believed that level of performance and ROI will explain student classification as eligible, ineligible, or not referred more than other student characteristics.

Design

To address research question 1, a one-factor quasi-experimental design was used to determine if students found to be eligible for special education, students referred for evaluation and found to be ineligible for special education, or those students not referred for multidisciplinary evaluation (MDE) to determine special education differ in terms of their ROI

and level of performance using curriculum-based measurement in reading (CBM-R) data. In this design, the students' status in relation to their referral for, or identification of, SLD (eligible, ineligible, not referred) was the independent variable and ROI and level of performance using CMB-R data were the dependent variables. A multivariate analysis of variance (MANOVA) was used to conduct this analysis.

To address research question 2, a multinomial logistic regression was used with ROI; level of performance; race; free and reduced meal status; sex; and grade 2, 3, or 4 as the predictor variables to examine the extent that ROI and level of performance predict group membership as either eligible, not eligible, or not referred.

Population

The population for this study was elementary school students in grades 2 through 4 who were receiving Tier 3 supplemental intervention using Pennsylvania's Multi-Tiered System of Support (MTSS) framework (Pennsylvania Department of Education, 2014).

Sample

A convenience sample was collected from only those schools approved by the commonwealth of Pennsylvania to use RTI for SLD eligibility determinations in reading. The sample was limited to only these schools to provide some level of consistency as to the RTI model used to identify students for special education. The sample included elementary school students from grades 2-4 who received the supplemental interventions and any possible evaluation and identification during the 2015-2016, 2016-2017, and 2017-2018 school years. Students of both sexes and all races were included.

An a priori power analysis based on previous research results was conducted using G*Power software (Faul & Fletcher, 2007) to estimate the needed sample size to carry out the

MANOVA for research question 1. Based on the effect size previously obtained by Bonesheski (2017), a sample size of 66 participants is needed to appropriately conduct the MANOVA. Sample size estimates for an omnibus linear regression using G* Power were 77 participants for a strong effect and 146 participants for a moderate effect. However, Petrucci (2009) recommended the use of at least 10 subjects per independent variable to appropriately conduct a multinomial logistic regression. As a result, a minimum of 60 subjects was required to appropriately conduct the analysis for research question 2. A minimum sample size of 66 participants was found to be sufficient to run the analyses for research questions 1 and 2, although a larger sample as close to 146 participants as possible was more desirable.

Measurement

The research design of the study uses a combination of categorical and continuous variables. What follows are the operational definitions of the categorical variables and the measurement properties of the continuous variables for the independent and the dependent variables as they pertain to each research question.

Research Question 1

The status of the student in relation to special education is a discrete variable with three levels: eligible for special education, ineligible for special education, and not referred. Eligible for special education was defined as students who were referred for MDE for reading concerns and determined eligible under the exceptionality of SLD in basic reading skills or reading fluency. To control for the possibility of students who would perform adequately on CBM-R but still be identified for special education, students who were identified with SLD in reading comprehension only were excluded from the study. Although they represent a small percentage of the overall population (approximately 1% in early elementary grades rising to approximately

9% by fifth grade) and more research is needed on this phenomenon, there are students who possess average reading fluency but below average reading comprehension skills (Meisinger, Bradley, Schwanenflugel, Kuhn, & Morris, 2009). Ineligible for special education was defined as students who were referred for MDE for reading concerns and found ineligible for special education. Not referred was defined as students who received Tier 3 services but were never referred for multidisciplinary evaluation. Although not common, it is possible that the ineligible and not referred conditions included students who were referred and/or received supplemental intervention for comprehension concerns.

The variable of eligible for special education was based on decisions made by MDE teams from the participating districts. Students determined eligible by their participating district as SLD in basic reading skills or reading fluency were counted as eligible. The variable of ineligible for special education was based on decisions by MDE teams from the participating districts as well. Students determined ineligible by their participating district as SLD in reading were counted as such for this study. The variable of not referred for special education was determined by the problem-solving teams of the participating districts. Students who received supplemental Tier 3 intervention but were not referred for evaluation during this time were counted as not referred.

The dependent variables of ROI and level of performance are both continuous variables derived from the repeated administration of CBM-R probes. ROI was calculated using an ordinary least squares (OLS) regression line calculated from repeated administration of CBM-R data over an intervention period using calendar weeks as the measurement interval. CBM-R data included common commercially available probes from DIBELS with standardized administration procedures. CBM-R have been shown to have adequate psychometric properties

as a measure of general reading ability (Deno, 1985). DIBELS probes are produced by using random passages based on a readability formula within a target range. DIBELS probes have an *SEE* of 15.26 and an *SEb* of 0.91 (Ardoyn & Christ, 2009).

A review a research by Ardoyn et al. (2013) suggests that a minimum of 10 data points are needed to obtain reliable ROIs. However, Christ et al. (2012) found that a minimum of 14 data points and a good or very good data set are needed to obtain ROIs needed for reliable high-stakes decision-making. Students who had less than 14 data points in their progress monitoring data were excluded from analysis to ensure reliable ROIs.

Level of performance was determined by taking the median of the last three CBM-R progress monitoring data points during the year of a student's multidisciplinary evaluation for the eligible and ineligible conditions and at the end of an intervention period for students in the not referred condition. Research has shown that using the median of three data points to estimate level of performance helps to account for the error produced by the variability in the alternate forms of the probes and generally results in stable levels of performance (Ardoyn & Christ, 2008).

Research Question 2

The predictor variable of level of performance is a continuous variable based on the administration of CBM-R probes and calculated by taking the median of the final three data points in a subject's data set during the year of multidisciplinary evaluation or at the end of the intervention period. The predictor variable of ROI is a continuous variable and will be calculated using the OLS trend line from the administration of CBM-R probes using calendar weeks as the measurement interval. The predictor variable of race is a discrete variable with two levels: traditionally not overrepresented in special education (White and Asian) and traditionally

overrepresented in special education (Black, Hispanic, and other). Race was determined by demographic information provided from the school the participant attended and coded into traditionally overrepresented/traditionally not overrepresented by the researcher.

Disproportionality of minority students in special education is also a well-documented problem, particularly black and Hispanic students, which is why this is a relevant variable (Hosp & Reschly, 2004; Skiba, Poloni-Staudinger, Gallini, Simmons, & Feggins-Azziz, 2006). The predictor variable of free and reduced meal status is a discrete variable with two levels: receives free and/or reduced meal or does not receive free and/or reduced meal. This is determined by the participant's district using federal and state guidelines for the National School Lunch Program (Benefits.gov, n.d.). Per the requirements of this program, eligible families must complete an application process based on household income to receive the free or reduced meal they may be eligible for because of their income level. As a result, students who may be eligible for a free or reduced meal may not receive one if their family has not completed the appropriate application. For the purposes of this study, only those students identified by their participating district as receiving free and reduced meal will be coded as such. The predictor variable of sex is a discrete variable with two levels: male or female. Sex will be determined by the demographic information provided by the school that the participant attends. Research has shown that student characteristics including sex and socio-economic status are highly influential in team decisions regarding eligibility for special education (Ysseldyke et al., 1982). The predictor variable of grade is a discrete variable and has three levels: grade 2, grade 3, and grade 4. Grade is defined as the grade the student is in during the intervention and evaluation period that corresponds with the data. It will be determined by the participating districts.

The dependent variable of being eligible for special education is a discrete variable and was defined as students who were referred for MDE for reading concerns and determined eligible under the exceptionality of a SLD in basic reading skills or reading fluency. Students identified with SLD in reading comprehension were excluded from the dataset for the previously noted reasons. The dependent variable of being ineligible for special education is a discrete variable and was defined as students who were referred for MDE for reading concerns and found ineligible for SLD. The dependent variable of not referred is a discrete variable and was defined as students who received Tier 3 services but were never referred for MDE. The ineligible and not referred conditions of the dependent variable might have contained students who were referred for MDE and/or receiving supplemental reading intervention due to comprehension concerns, although it would not be common.

Procedures

Schools were contacted by email for participation using a site approval letter. Schools were selected from a list of schools approved by the commonwealth of Pennsylvania to use RTI for SLD eligibility decisions using the contact information provided from this list (Pennsylvania Department of Education, n.d.). Individual school psychologists were contacted by email obtained via school district websites. All participating districts and school psychologists were asked to provide the needed participant data using a Microsoft Excel File containing the following information: a local and confidential identification number for each student, the complete progress-monitoring data set that corresponded to the time of the multidisciplinary evaluation or the most recent intervention period, sex as either M for male or F for female based on the demographic information provided by the district, free or reduced meal as X if the students' receive free or reduced meal or no entry if they do not receive free or reduced meal

based on the demographic information provided by the district, race as determined by the district's demographic information, and grade as either 2, 3, or 4 depending on the students' grade placement determined by the district at the time of the evaluation if in the eligible or ineligible condition or during the intervention period for the not-referred condition. Data as to whether the students were determined eligible for special education, ineligible for special education, and those students not referred for MDE were also included on the spreadsheet. Again, participating districts were asked to only include students determined eligible for SLD in basic reading skills and reading fluency but the students who were in the ineligible and not referred categories could have included students referred for comprehension concerns or receiving supplemental intervention due to comprehension concerns.

Statistical Analysis

The first research question was addressed by conducting a one-way MANOVA. Group categorization as eligible, not eligible, and not referred was the independent variable and ROI and level of performance were the dependent variables. Post hoc analyses and effect sizes were also conducted. A MANOVA has the following assumptions: the dependent variables are interval or ratio data, the independent variables are comprised of two or more independent and categorical groups, there is independence of observations, there are more cases in each cell than there are dependent variables, there is a normal distribution of the dependent variables, there are no outliers, there is a linear relationship between the dependent variables, there are no strong correlations among the dependent variables, and there is homogeneity of the variance-covariance matrices (Laerd Statistics, n.d.).

The second research question was addressed by a multinomial logistic regression with ROI, level of performance, sex, race, free and reduced meal status, and grade as the predictor

variables. Group categorization as eligible for special education, ineligible for special education, and not referred for special education are the dependent variables in this analysis. The dependent variables were the odds ratio for each variable, but of particular interest, the odds ratios for ROI and level of performance. Odds ratios of greater than one are desirable for the variables of ROI and level of performance because the hypothesized relationship is that ROIs and levels of performance will significantly predict special education eligibility. Logistic coefficients and tests of significance using Walds test will determine if the odds ratios for ROI and level of performance are statistically significant. The percentage of correctly classified cases was also examined to see how well the model predicts eligibility (Wright, 1995). The following assumptions must be met for a multinomial logistic regression: the dependent variable must be nominal, there are one or more categorical or continuous independent variables, independence of observations, the dependent variable is mutually exclusive and exhaustive, there is no multicollinearity, there is a linear relationship between the continuous variables and the logit transformation of the dependent variables, and there are no outliers (Laerd Statistics, n.d.).

It was originally planned to address research question 2 with a multinomial logistic regression. However, as indicated in Chapter 4, after the MANOVA was conducted it was determined that a binomial logistic regression would be a preferable statistical procedure. A binomial logistic regression has the same assumptions as a multinomial logistic regression with the only difference being that there are only two levels of the categorical dependent variable. The dependent variables of ineligible for special education and not referred for special education were combined resulting in two categorical outcomes for the dependent variable: eligible for special education and ineligible for special education. When conducting the binomial logistic regression, the dependent variable of eligible was coded as 0 and the dependent variable of

ineligible was coded as 1. The categorical variable independent variables were coded as follows: grade 2 as 1, grade 3 as 2, grade 4 as 0, receives free and reduce meal as 0, does not receive a free or reduced meal as 1, overrepresented in special education as 0, not overrepresented in special education as 1, male as 0, and female as 1.

Summary

This study examined how ROI and level of performance influence SLD eligibility decisions when using a dual-discrepancy model. To do this, a convenience sample of schools approved by the commonwealth of Pennsylvania was recruited to participate in the study. A MANOVA was conducted to see if students in the following groups were significantly different in their level of performance and ROI: referred and eligible for special education, referred and ineligible for special education, and receiving Tier 3 intervention but not referred. Then, a multinomial logistic regression was conducted to determine if ROI and level of performance predict categorization in the previously mentioned groups more than other student characteristics. It was hypothesized that level of performance and ROI will differentiate students and predict special education eligibility more than other student characteristics.

CHAPTER 4

RESULTS

The purpose of this study was to examine several aspects of the dual-discrepancy model and its application to the identification of specific learning disabilities (SLD) in reading. Two research questions were posed. Whether the dual-discrepancy model can differentiate students who are determined eligible for special education from those students who are determined ineligible for special education and those students receiving general education reading intervention and not referred for multidisciplinary evaluation (MDE) was examined by research question 1. It was hypothesized that level of performance and rate of improvement (ROI) would differentiate students in these three groups. In research question two, what student attributes, including level of performance and ROI, best predicated MDE team decisions regarding SLD using the dual-discrepancy model were analyzed. It was hypothesized that level of performance and ROI would predict multidisciplinary team eligibility decisions more than other student characteristics.

Sample

A convenience sample of students who received Tier 3 supplemental reading intervention from the academic years 2015-2016, 2016-2017, and 2017-2018 was obtained from a single school district approved by the state of Pennsylvania to use response to intervention (RTI) for SLD eligibility decisions in reading. Students with less than 14 curriculum-based measurement in reading (CBM-R) data points in their data set were excluded from analysis to ensure reliable ROIs could be calculated (Christ et al., 2012). The final sample included 163 participants. As indicated in Table 1, the sample was comprised of 53.5% females and 46.6% males. As shown

in Table 2, the sample was composed of 40.5% second-grade students, 29.4% third-grade students, and 30.1% fourth-grade students.

Table 1

Number of Student Cases by Sex

Sex	Frequency	Percent
Female	87	53.4
Male	76	46.6
Total	163	100

Table 2

Number of Student Cases by Grade

Grade	Frequency	Percent
2	66	40.5
3	48	29.4
4	49	30.1
Total	163	100

The sample demographics by race were 8.6% Asian, 8.6% Black, 6.1% Hispanic, and 76.7% White as indicated in Table 3. As indicated in Table 4, 53.4% of the sample was ineligible for a free or reduced meal and 46.6% of the sample was eligible for a free or reduced meal. As shown in Table 5, a total of 9.2% of the sample was determined eligible for special education (15 subjects), 4.9% was referred for evaluation and found ineligible for special

education (8 subjects), and 85.9% of the sample was receiving supplemental reading intervention but never referred for MDE (140 subjects).

Table 3

Number of Student Cases by Race

Race	Frequency	Percent
Asian	14	8.6
Black	14	8.6
Hispanic/Latino	10	6.1
White	125	76.7
Total	163	100

Table 4

Number of Student Cases by Free and Reduced Meal Status

Free/Reduced Meal Status	Frequency	Percent
Not Eligible	87	53.4
Eligible	76	46.6
Total	163	100

Results of Analyses

The results of the statistical analyses for each research question are presented below.

Research Question 1

Do students who are identified as SLD in reading using RTI differ from those students who are determined ineligible for special education and those students receiving supplemental

Table 5

Number of Student Cases by Special Education Eligibility Status

Disability Status	Frequency	Percent
Eligible	15	9.2
Ineligible	8	4.9
Not Referred	140	85.9
Total	163	100

intervention but not referred for evaluation on their ROI and/or level of performance? It was hypothesized that these three groups would be different in both level of performance and ROI. A one-way MANOVA was conducted to determine if students who were determined eligible for special education, determined ineligible for special education, and those receiving Tier 3 supplemental intervention but not referred for evaluation were differentiated by their level of performance and ROIs as measured by CBM-R probes.

A MANOVA has the following assumptions: the dependent variables are measured on a continuous scale, there are at least two categorical groups for the independent variable, there is independence of observations, there are more cases in each cell than there are dependent variables, there is a normal distribution of the dependent variables, there are no outliers, there is a linear relationship between the dependent variables, there is multivariate normality, and there is homogeneity of the variance-covariance matrices (Laerd Statistics, n.d.).

The dependent variables of level of performance and ROI are both continuous variables. The research design contains three categorical groups of the independent variable. Each participant's data is only represented in the dataset one time. To determine if there were an

appropriate number of cases in each cells as recommended by Pallant (2016), the number of independent variables was multiplied by the number of dependent variables ($3 \times 2 = 6$). Each cell of the MANOVA contained more than 6 cases. For the dependent variable level of performance, examination of Q-Q plots did not show clustering (See Appendix C), skewness = -.02, and kurtosis = -0.12. The Kolmogorov-Smirnov Test of Normality showed a normal distribution for the dependent variable of level of performance $D(163) = 0.04, p = .20$. Therefore, the assumption of normality was met for the dependent variable of level of performance. No outliers were observed for level of performance as determined by inspection of box and whisker plots (See Appendix D).

The dependent variable of ROI showed a non-normal distribution evidenced by significant results for the Kolmogorov-Smirnov Test of Normality $D(163) = 0.07, p = .044$ and clustering on the detrended Q-Q plots (See Appendix E). Additionally, skewness = 0.16 and kurtosis = 1.32 for the variable of ROI. No significant outliers were observed for the dependent variable of ROI through inspection of the box and whisker plots (See Appendix F). No multivariate outliers were detected using Mahalanobis Distance ($p > .001$). It was decided to continue with the MANOVA despite the violation of this assumption and then verify the results of the MANOVA using a nonparametric alternative.

A lack of strong correlations were observed among the dependent variables ($r = 0.18, p = 0.2$) therefore the assumption of multicollinearity and singularity was met. Box's Test of Equality of Covariance Matrices confirmed that the assumption of homogeneity of the variance-covariance matrices was also met ($p = 0.40$). The homogeneity of variance assumption was confirmed by Levene's Test of Homogeneity of Variance ($p > 0.05$).

Descriptive Statistics for the dependent variables are provided in Tables 6 and 7 below.

Table 6

Descriptive Statistics for Level of Performance and ROI for Total Sample

Variable	Minimum	Maximum	<i>M</i>	<i>SD</i>	Skewness	Kurtosis
Level	11	147	80.39	26.73	-0.02	-0.12
ROI	-0.50	2.47	0.92	0.46	0.16	1.32

Note. ROI= Rate of Improvement. Level of performance is measured in words correct per minute. Rate of improvement is measured in words correct per minute per week.

Table 7

Descriptive Statistics for Level of Performance and ROI by Independent Variable

Dependent Variable	Independent Variable	<i>M</i>	<i>SD</i>	<i>n</i>
Level	Eligible	55.60	30.72	15
	Ineligible	78.88	16.20	8
	Not Referred	83.14	25.50	140
	Total	80.39	26.73	163
ROI	Eligible	0.77	0.31	15
	Ineligible	0.97	0.38	8
	Not Referred	0.93	0.48	140
	Total	0.92	0.46	163

Note. ROI= Rate of Improvement. Mean for level is in words correct per minute. Mean for rate of improvement is in words correct per minute per week.

As indicated in Table 8, results of the overall MANOVA indicated statistically significant differences between the three levels of the independent variables. An ANOVA was conducted

on both dependent variables due to the significant main effect for the MANOVA, $F(4, 318) = 3.93, p = .004$; Wilks' $\Lambda = 9.08$. As indicated in Table 9, there was a statistically significant effect for level of performance but not for ROI, $F(2, 160) = 7.81, p = .001$; partial $\eta^2 = .09$.

Table 8

MANOVA Results for the Effects of Eligibility Status on Level of Performance and ROI

Effect	Wilks' Λ	F	Hypothesis df	Error df	p	η^2
Independent Variable	9.08	3.93	4	318	0.004	.05

Note. ROI= rate of improvement

Table 9

ANOVA Results for the Effects of Eligibility Status on Level of Performance and ROI

Source	Dependent Variable	Type III Sum of Squares	df	Error df	Mean Square	F	p	η^2
Eligibility Status	Level	10291.98	2	160	5145.99	7.81	.001	.09
	ROI	0.38	2	160	0.19	0.89	.412	.01

Note. ROI= rate of improvement

Post hoc analyses were run on the variable level of performance to determine between which levels of the dependent variable significant differences existed. As indicated in Table 10, there was a statistically significant difference between students determined eligible for special education and students not referred for evaluation on the dependent variable of level of performance, $p < .001$; $q = 3.95$. There were no statistically significant differences between

eligible students and ineligible students nor between ineligible students and not referred students on the dependent variable of level of performance.

Table 10

Post Hoc Analyses for Main Effect of Level Using Tukey HSD

Comparison		Mean Difference	Std. Error	<i>q</i>	<i>p</i>	95% Confidence Interval	
						Lower Bound	Upper Bound
Eligible vs.	Ineligible	-23.28	11.24	2.07	0.099	-49.87	3.32
Eligible vs.	Not Referred	-27.54	6.98	3.95	0.000	-44.04	-11.03
Ineligible vs.	Not Referred	-4.26	9.33	0.46	0.892	-17.82	26.34

Because the assumption of normality was violated for the dependent variable of ROI, the non-parametric alternative Kruskal-Wallis Test was conducted to verify the obtained results of the MANOVA. The results were identical to the MANOVA in that there was a main effect for level of performance but not for ROI.

The hypothesis for research question 1 was only partially supported because eligible, ineligible, and not referred students were only different for level of performance and not for ROI. Additionally, students who were eligible only differed from students who were not referred on level of performance. No differences on level of performance were observed between eligible and ineligible students nor between ineligible and not referred students.

Research Question 2

What student attributes best predict eligibility for special education using a dual-discrepancy approach? Originally, a multinomial logistic regression was proposed using sex, grade, race, free and reduced meal status, level of performance, and ROI as the predictor variables and special education eligibility status defined as referred for MDE and determined eligible, referred for MDE and determined ineligible, or receiving supplemental reading intervention and not referred for MDE as the outcome variable. However, since the MANOVA did not differentiate students who were ineligible from students who were not referred, and due to the low number of cases in the ineligible condition, the ineligible and not referred conditions were collapsed into one level of the dependent variable referred to as ineligible. Then, a binomial logistic regression was conducted.

A binomial logistic regression has the following assumptions: the dependent variable must be nominal, there are one or more categorical or continuous independent variables, independence of observations, the dependent variable is mutually exclusive and exhaustive, there is no multicollinearity, there is a linear relationship between the continuous variables and the logit transformation of the dependent variables, and there are no outliers (Laerd Statistics, n.d.).

Examination of the research design shows that the dependent variable is nominal. All cases are represented only one time in the data set and all possible outcomes of the dependent variable (eligible and not referred) are represented in the design. The data set was examined and the independence of observations assumption was also met. The assumption of multicollinearity was met as the variance inflation factor for each variable was less than 10 (See Appendix G). The assumption of linearity was confirmed by the Box-Tidwell (1962) procedure ($p = 0.11$ for the natural log transformation of level of performance by level of performance and $p = 0.23$ for

the natural log transformation of ROI by ROI). There were several outlier cases as determined by examining studentized residuals (See Appendix H). However, these outliers were retained for analysis because they were all in the eligible condition and removing them would have lowered the numbers of this condition which already had a limited number of cases.

The logistic regression model was statistically significant, $\chi^2 = 17.41, p = .015$. The model explained 22% of the variance in students being correctly classified as eligible for special education and correctly classified 90.8% of cases. Sensitivity was 99.3% and specificity was 6.7%. As indicated in Table 11, the only predictor variable that was significant was level of performance. Lower levels of performance using CBM-R probes resulted in an increased likelihood of being found eligible for SLD, Wald $\chi^2 = 5.82, p = .02$. ROI and other student attributes did not meaningfully predict classification as eligible for special education.

The hypothesis for research question 2 was only partially supported because level of performance predicted disability status and ROI did not.

Conclusion

Two research questions pertaining to the dual-discrepancy model were examined in a series of analyses. First, a one-way MANOVA was conducted to determine if students who were not referred for evaluation, students referred for MDE and found ineligible for special education, and students referred for MDE and found eligible for special education were different in terms of their level of performance and ROI using CBM-R probes. The results showed that these three groups are significantly different in their level of performance but not in their ROI. Students found eligible for special education had significantly lower levels of performance than students not referred for evaluation. No significant differences were found between students who were eligible for special education and students referred for evaluation but not found eligible. Nor

Table 11

Binomial Logistic Regression Predicting Special Education Eligibility Based on Sex, Grade, Race, Free and Reduced Meal Status, Level of Performance, and ROI

Independent Variables	<i>B</i>	<i>SE B</i>	Wald χ^2	<i>df</i>	<i>p</i>	Exp(B)	95% CI for EXP(B)	
							Lower	Upper
Sex	-0.43	0.62	0.47	1	0.49	0.65	0.19	2.21
Grade 2			0.33	2	0.85			
Grade 3	0.73	1.28	0.33	1	0.57	2.08	0.17	25.74
Grade 4	0.42	1.04	0.16	1	0.69	1.52	0.20	11.67
Race	0.42	0.70	0.35	1	0.55	1.52	0.38	5.10
Free/Reduced Meal	0.42	0.64	0.44	1	0.51	1.52	0.44	5.30
Level	0.05	0.02	5.82	1	0.02	1.05	1.01	1.09
ROI	0.06	0.77	0.01	1	0.93	1.06	0.23	4.84
Constant	-1.88	1.84	1.04	1	0.31	0.15		

Note. ROI= rate of improvement

were significant differences found between students referred for evaluation and found ineligible and those students receiving Tier 3 support and not referred for evaluation. No significant differences were found between the three groups on ROI. Following this, a binomial logistic regression was performed to determine what student attributes, particularly level of performance and ROI, predict eligibility for special education. Level of performance measured by CBM-R probes was significant. The results showed that lower levels of performance resulted in an

increased likelihood of being determined eligible for special education with ROI and other student attributes not meaningfully predicating eligibility status.

CHAPTER 5

DISCUSSION

The purpose of this study was to determine if students who were identified as having a specific learning disability (SLD) in reading using response to intervention (RTI) were different from other struggling readers receiving supplemental intervention but not referred and/or not identified for special education on their level of performance and rate of improvement (ROI). Additional analyses were conducted to determine if level of performance and ROI predicted eligibility as SLD in reading more than other student characteristics. A one-way MANOVA was conducted with special education eligibility status as the independent variable and level of performance and ROI as the dependent variables. The results showed that students who were SLD in reading had statistically significantly lower levels of performance than students who received Tier 3 services and were not referred for evaluation. Following this, a binomial logistic regression was conducted using sex, grade, race, free and reduced meal status, level of performance, and ROI as the predictor variables. Eligibility status for special education was the outcome variable with two levels: eligible for special education and not eligible for special education. The not eligible condition included students who were referred for multidisciplinary evaluation (MDE) and found ineligible and students who received Tier 3 intervention and were not referred for evaluation. The results showed that lower levels of performance resulted in an increased likelihood of being found eligible for special education.

Summary of Findings

In the sections that follow, the results of the analyses and the accuracy of the hypotheses for each research question are discussed. Additionally, the results of the current study are related to the previous research that formed the basis for the research questions and hypotheses. Results

that are consistent with previous research are noted and potential hypotheses for results that differ from previous research are proposed.

Research Question 1

Do students who are identified as SLD in reading using an RTI framework differ from those students who are determined ineligible for special education and those students receiving supplemental intervention but not referred for evaluation on their ROI and/or level of performance? It was hypothesized that students identified as SLD would have significantly lower levels of performance and lower ROIs than students not identified as SLD and students not referred for evaluation. The results indicated that students who were determined eligible for SLD had significantly lower levels of performance than students who received Tier 3 services and were not referred for evaluation. There were no differences between students who were determined eligible and those who were referred for evaluation and determined ineligible. Nor were there differences between students who were referred for evaluation and determined ineligible and those students who received Tier 3 services and were not referred. There were no differences amongst the three levels of the independent variable on ROI.

The hypothesis for research question one was only partially supported. Level of performance only differentiated students with SLD from students who were not referred for MDE. ROI did not differentiate eligible, ineligible, and not referred students. This is somewhat consistent with the results obtained by Boneshefski (2017) who also found that students with SLD had significantly lower levels of performance than students not referred for evaluation. However, Boneshefski found that students with SLD had lower ROIs than those students not referred for MDE as well. One explanation for this difference could be the limited number of cases in the eligible and ineligible conditions when compared to the not referred condition in the

current study, although Boneshefski's study suffered from the same limitation. Another explanation could be the use of different curriculum-based measurement in reading (CBM-R) probes with different alternate form reliabilities in each study's sample. The probes used by Boneshefski produced a slope with an *SEb* of 0.83 at grade 2 and 0.70 at grade 3. The probes used in the current study have an *SEb* of 0.91. As a result, the ROIs in Boneshefski's study may have had less overall variance which could explain why his study detected a difference in ROIs between students with SLD and those not referred for evaluation.

The results for research question 1 are consistent with previous research conducted by Shinn and Petersen (2002) who found that low levels of performance relative to local standards were the main factor in SLD eligibility decisions regardless of the eligibility criteria that were in place. Similarly, it seems that in RTI systems, the main characteristic of SLD students is significantly lower levels of performance compared to peers, not necessarily lower levels of growth even though this is part of the federal regulations pertaining to the use of RTI for eligibility purposes.

In previous studies of RTI and the dual-discrepancy model the addition of ROI resulted in the identification of a distinct, and a more impaired, group of students in reading (Speece & Case, 2001; Speece et al., 2003). However, the results of the current study did not support this as students identified with SLD were only different from not referred students on level of performance. One explanation for this may be that the research conducted by Speece and Case (2001) and Speece et al. (2003) identified a group of students classified "at risk" that likely included students at a Tier 2 and a Tier 3 level of functioning. It is possible that ROI is sensitive enough to distinguish students at a Tier 3 level of functioning from students at a Tier 2 level of functioning, but not sensitive enough to distinguish students with SLD from other, non-disabled

students receiving Tier 3 support. Speece and Case (2001) used an ROI of one standard deviation below the class average as the cut point for determining dual discrepancy status. Thus, the students who demonstrated a dual discrepancy in this study showed lower ROIs compared to typically developing readers in their class, not necessarily lower ROIs when compared to other at-risk readers.

Previous research has also demonstrated that measurement error is a concern when interpreting ROIs and that the *SEM* of ROIs can interfere with the ability to accurately measure individual student growth using CBM-R (Ardoin & Christ, 2009; Christ, 2006; Christ & Silberglitt, 2007; Christ et al., 2012). This may also be a factor in the current study's findings that ROI does not differentiate students with SLD from other struggling readers. This is particularly relevant since there were low numbers of students in both the eligible and ineligible conditions making those conditions more susceptible to the influences of a measurement with poor reliability. The measurement error associated with ROI is a well-documented flaw of RTI identification approaches. Despite research supporting the dual-discrepancy model as the most technically adequate RTI method for determining SLD, there are still issues with the technical adequacy of the model itself (Fuchs et al., 2008; Fuchs, Fuchs, & Compton, 2004). It is possible that the measurement issues surrounding ROI may be part of these shortcomings of the dual-discrepancy model and that ROI simply is not sensitive enough to reliably differentiate SLD students from other struggling readers in all cases.

In the current study, mean ROIs for the eligible condition (0.77 wcpm/week) were lower than both the referred and ineligible condition (0.97 wcpm/week) and the not referred condition (0.93 wcpm/week) although the comparisons were not significant. However, the standard deviation for the eligible condition was 0.31. This finding illustrates a main criticism with ROI

as a metric and supports the previous point in that the amount of expected variance in the ROIs for the eligible condition can result in ROIs for some students that exceed the mean performance of referred and ineligible and not referred conditions. So, although eligible students on a whole had lower (but not significantly lower) ROIs than referred and ineligible students and not referred students, some eligible students could still have an ROI that exceeds the average ROI of students who were never even considered for special education services.

Consistent with historic research conducted by Shinn and Marston (1985) students who were referred and eligible for special education were different from not referred students on level of performance. However, it is interesting that students who were referred and ineligible were not different from students who were referred and eligible on either dependent variable, particularly level of performance. This begs the question, what is it about these students that was different enough from the eligible students to find them ineligible for services when neither their ROIs or level of performance were significantly different? Sample size may again be the explanation for these results as the referred and ineligible condition had the lowest number of subjects of all three conditions. The results approached, but did not reach significance for level of performance between the eligible and ineligible conditions and the mean performance for the referred and ineligible condition (78.88) was very similar to the not referred condition (83.14) from a practical standpoint. A larger sample for the referred and ineligible condition may result in different findings.

Research Question 2

What student attributes best predict eligibility for special education using a dual discrepancy approach? Originally, a multinomial logistic regression with ROI, level of performance, sex, race, free and reduced meal status, and grade as the independent variables was

proposed. The dependent variable was the students' status in relation to their referral for or identification of SLD (eligible, ineligible, not referred). Since the results of the MANOVA did not support that the ineligible and not referred students were two distinct groups, the analysis was conducted as a binomial logistic regression with the ineligible not referred conditions being merged into one level of the dependent variable called not eligible. It was hypothesized that ROI and level of performance would predict student categorization as eligible and not eligible more than sex, race, free and reduced meal status, and grade. The results indicated that student level of performance on CBM-R probes was predictive of special education eligibility. Lower levels of performance on CBM-R resulted in an increased likelihood of being identified eligible for special education as SLD in reading. ROI and other student attributes did not predict group membership as eligible or ineligible for special education.

The hypothesis for research question two that level of performance and ROI would predict special education eligibility was only partially supported. Lower levels of performance on CBM-R predicted group membership as eligible for special education. However, lower ROIs did not significantly predict eligibility for special education. Similar to the MANOVA, there were a limited number of cases for the eligible level of the dependent variable, which may explain why ROI did not predict eligibility. However, these results are consistent with those obtained by Boneshefski (2017) who also found that ROI did not predict eligibility as SLD when using dual-discrepancy criteria. The results of the current study are also consistent with previous research that shows that referrals and decisions made using CBM-R and dual-discrepancy criteria mitigate the influence of student attributes such as race and socio-economic status on the decision-making process (Marston, Mirkin, & Deno, 1984; Speece & Case, 2001).

As indicated above, research has established that level of performance measured by CBM-R can distinguish typically-developing readers from at-risk readers and students with reading disabilities. However, there does not seem to be evidence that ROI adds any utility to the decision-making process when trying to distinguish disabled readers from other low-level readers. As previously stated, measurement issues with ROI itself could be a possible explanation for this. Another possibility is that the lowest performing students are also most often the students with the lowest ROIs. Research by Silberglitt and Hintze (2007) found that students with the lowest levels of performance also showed the lowest ROIs in general when compared to typically developing readers. In other words, students who start low tend to grow more slowly. This may be more a reflection of reading disabilities themselves and once generally supportive instructional conditions are controlled for, measuring ROI in addition to level of performance may be redundant. However, further research is needed to confirm this hypothesis.

Another explanation for why students identified with SLD using RTI differed on level of performance and not ROI is that despite showing robust growth, MDE teams may still believe that a student's level of performance is too discrepant from grade-level standards of performance or requires too many resources to maintain that level of growth without special education supports, regardless of their ROI. Further examination of the data found that five of the participants in the referred and eligible condition had ROIs comparable to or higher than the mean ROIs of the referred and ineligible and not referred conditions. Of these five, only two had levels of performance that were lower than the mean level of performance for the referred and ineligible condition and the not referred condition. It seems that in several of these situations, it is possible that even though a robust ROI was obtained, the MDE teams may have felt that the

level of performance was still too discrepant to be supported without special education resources. As for the remaining three cases, one possibility is that legal/mediation issues may have resulted in several students being identified SLD despite ROIs and level of performances that were more similar to the ineligible and not referred conditions. In examination of the referred and ineligible condition, two participants had ROIs lower than the mean of the eligible condition. Both of these students' level of performance was below the mean for the referred and ineligible condition but above the mean for the eligible condition. It seems that in the case of students determined ineligible, the dual-discrepancy criteria were more consistently implemented.

Limitations

A limitation of the current study is the sample. The limitations with the current sample stem from two issues that impact both the internal and external validity of the study. Although the total sample size was adequate, there were a low number of subjects in the eligible and referred but ineligible conditions when compared to the not referred condition. Despite meeting the assumptions needed to conduct the analyses, a lower number of cases in these conditions compared to the not referred condition may have had an impact on the outcome of the statistical analysis thus affecting the internal validity of the study. Another shortcoming of the sample used in the current study was that it was a convenience sample obtained from a single school. Recruitment from a larger number of schools and from a wider variety of schools would be a more ideal sample and result in greater external validity.

A design limitation of the current study is the use of students from multiple grades in the same analysis. Research has shown that ROI is the steepest at early elementary grades (first and second grade), begins to taper off in third grade, and levels off by around sixth grade (Nese et al., 2013). A difference in more than half a word per minute in ROI exists between students in

fourth grade compared to second-grade students. As a result, the inclusion of multiple grades may have skewed the analyses conducted when it comes to ROI and affected the internal validity of the study. However, it is worth noting that Bonesheski (2017) only used students from a single grade level in his analyses and also found that ROI did not predict special education eligibility. Additionally, level of performance also has different expectations based on different grade levels and level of performance was different among SLD and not referred students and predicted special education eligibility despite using data from multiple grade levels.

Using calendar weeks as the measurement interval when calculating ROI is another limitation of the current study. Runge, Bennyhoff, Ferchalk, and McCrea (2017) found that different ROIs could be obtained depending on the measurement interval used when calculating ROI using the ordinary least squares (OLS) method. They compared ROIs calculated using the actual dates of the progress monitoring, calendar weeks when the progress monitoring occurred, and school weeks from when the initial fall benchmarking occurred. When interpreting ROI using the OLS method, the measurement interval is assumed to be equal (e.g. per week). Because school holidays and logistic issues result in missed days and weeks of school and uneven intervals for the monitoring of student progress, each of these measurement intervals produced different ROIs. As a result, the ROIs used for this study likely contain a source of additional error.

Another limitation of the current study is the use of MDE team decisions as the independent variable in the first analysis and the dependent variable in the second analysis. The historic subjectivity of MDE team decisions and their influence by extraneous student variables as opposed to objective eligibility criteria is a well-documented problem (McMillon & Speece, 1999; Ysseldyke, Thurlow, Graden, Deno, & Algozzine, 1982). As previously stated, there is no

definitive test for determining SLD. Using these preexisting MDE team decisions as the independent variable in the first analysis and the dependent variable in the second analysis of the study decreases the internal validity of the study because there is no guarantee that the decisions used as the basis for this study represent a group of students that are truly SLD. Since there is no definitive test for identifying SLD, nor are there agreed upon and empirically verified criteria for any model of SLD identification, there is no way to definitively operationalize the dependent variable that results in a group of students that absolutely have SLD free of any kind of error. Although research supports that RTI systems reduce bias in the type of students referred for and found eligible for special education (Marston, Mirkin, & Deno, 1984; Marston, Muyskens, Lau, & Carter, 2003), it does seem that MDE teams using RTI are still biased towards other factors in at least some instances (3 out of 15 in the current study). The current study would also suggest, however, that student characteristics did not bias the MDE team decisions used as the basis of the independent variable in the first analysis and the dependent variable in the second analysis

Implications for Research

Future research should attempt to verify the findings of the current study using a more diverse and larger sample. More cases in the referred and eligible and the referred and ineligible conditions would make for a design with better internal validity. It is important for problem-solving teams using dual-discrepancy criteria to make decisions regarding students' educational programming and their need for specially designed instruction to have research-based guidelines to assist in their decision making. If ROI is truly not sensitive enough to differentiate students with reading disabilities from other at-risk students, then alternative, research-based guidelines for differentiating non-responders needs to be a continued focus of research. A larger sample may also find significant differences between referred and eligible and referred and ineligible

students as well. Replication of the current results with a sample that has more internal and external validity is needed to accurately determine if the dual-discrepancy model can reliably differentiate these three groups of students.

Additional research comparing the dual-discrepancy approach to the exit groups approach may also prove beneficial (Vaugh, Thompson, & Hickman, 2003). If level of performance and ROI are truly related in that lowest performers tend to show the slowest growth, and if level of performance continues to be the driving factor in eligibility decisions even with dual-discrepancy criteria in place, the use of exit groups may be a practical, efficient, and research-based approach to help simplify SLD decisions using RTI. The inclusionary criteria for lack of intervention response is still part of the exit groups model but is operationalized as final performance versus initial performance as opposed to a rate. However, more research on the exit groups methods is needed to establish the sensitivity and specificity of that model and how the group of students identified using exit groups compares to estimated prevalence rates for SLD. Research should also compare the diagnostic accuracy the dual-discrepancy model to the exit groups approach to determine if the exit groups approach is at least as technically adequate as the dual-discrepancy model for the identification of SLD. ROI could still have a place within an exit groups model as a way gauge intervention effectiveness along the way and modify interventions that are clearly not resulting in ambitious student growth. However, the main factor in determining referrals and eligibility for SLD would be level of performance after an appropriate intervention period. In this way, the criteria for growth and level of performance are the same in the end.

More research is also needed on the ROI of students with SLD in general. Much is known about the ROI of typically developing readers at each grade level as well as longitudinally across the elementary years. It may be beneficial to examine the ROI of SLD

students in a similar manner to help guide the referral and eligibility process when using RTI. Deno et al. (2001) conducted some preliminary research into this matter. However, if ROI is to be a central component of SLD eligibility, more research on the ROI of SLD students both within and across grades is needed. Practitioners would benefit from a more precise understanding of the ROI of students with SLD including mean levels of performance at each grade, confidence intervals, and long-term developmental trajectories. This may help improve decision making processes by providing more precise eligibility criteria based on the characteristics of a sample of disabled readers as opposed to subjective comparisons to an aimline or having to independently generate the standard deviation of some reference group.

Some research has questioned if using OLS for calculating ROI is indeed the most technically adequate method available. Vannest, Parker, Davis, Soares, and Smith (2012) found that the use of a nonparametric method for summarizing ROI, called Theil-Sen, may be a promising alternative to the use of OLS. Using the data from 372 published datasets, they found that slopes produced using Theil-Sen were highly intercorrelated with those produced using linear regression methods. Theil-Sen also has the advantage of being more defensible because it does not have to meet the parametric assumptions of linear regression methods, a flaw commonly observed with datasets obtained when using CBM-R probes to monitor student progress. Mercer, Lyons, Johnston, and Millhoff (2015) compared ROIs calculated using OLS to ROIs calculated using several iteratively reweighted least-squares methods referred to as robust estimators. They found that the robust estimators outperformed OLS when outliers were present. Future research should seek to replicate the results of these studies. Future research should replicate the results of the current study using these alternative methods for ROI as well to determine if the results are similar with this type of measurement. Finally, future research

should extend the application of these alternative methods for determining ROI to the dual-discrepancy model and existing criteria for determining non-responders.

Implications for Policy

The construct of SLD has been controversial since its introduction into federal legislation in 1975. Lack of a clear and agreed-upon definition of the construct as well as clearly defined eligibility criteria have made SLD and identification practices related to SLD a source of much confusion and conflict for practitioners and researchers alike (Fletcher, Lyon, Fuchs, & Barnes, 2007). Additionally, many proposed models for the identification of SLD are not empirically reliable. Ability-achievement discrepancies have long been known to be inadequate for the identification of SLD (Hoskyn & Swanson, 2000; Stuebing et al., 2002). More recent models of identification based on intra-individual comparisons and patterns of strengths and weaknesses of cognitive processes related to learning have also been found to be technically inadequate (Fletcher, Stuebing, & Vaughn, 2014; Stuebing, Fletcher, Branum-Martin, & Francis, 2012). Research has shown mixed support for the technical adequacy of RTI models as well (Fuchs et al., 2008; Fuchs, Fuchs, & Compton, 2004; Fuchs et al., 2001).

The shortcomings of using RTI for SLD identification likely have more to do with the construct of SLD than it does with the RTI model. Measuring psychological constructs is difficult since they cannot be directly measured. This matter is further complicated when the construct is poorly defined and there are no consistent or agreed-upon criteria for identification. The current study continues to illustrate the difficulties with reliably identifying a poorly defined psychological construct with no definitive and agreed-upon identification criteria or method for true identification. Research has repeatedly shown that the dual-discrepancy model and using

the additional criterion of growth in response to intervention efforts identifies a group of students that is consistent with the estimated prevalence rates of SLD (Fuchs & Fuchs, 1998; Speece & Case, 2001; Speece et al., 2003). However, these studies had clearly defined and predetermined criteria for insufficient growth. Although federal inclusionary criteria for using RTI to identify SLD mandate that student growth, or lack thereof, be measured and used as part of the eligibility process, Flinn (2015) documented the lack of state-level guidance for using growth as part of the SLD eligibility process for RTI. This is extremely problematic because Barth et al. (2008) showed that the cut points used for level of performance and ROI are very influential on the group of students subsequently identified as SLD. The results of the current study do not offer definitive evidence for using any specific criteria regarding ROI as part of the SLD identification process and would caution policy makers to avoid eligibility criteria with strict guidelines for interpreting ROI. However, without clearly defined and predetermined criteria for using ROI in the context of the dual-discrepancy model, it is possible that ROIs lack of ability to differentiate students with SLD from those students determined ineligible or not referred in the current study is due to a lack of clear decision-making rules available for practitioners and MDE teams.

Given the results of the current study and Boneshefski (2017), it may be tempting and even advisable to explore alternative ways to operationalize growth within RTI systems. However, the longstanding struggles with defining and identifying SLD necessitate that it is time to move on from matters of categorization criteria and focus on creating empirically-based, practical, and efficient guidelines that emphasize student need for specially designed instruction over categorization. Knowing that there will be a certain amount of error no matter what model is chosen, SLD identification practices in schools should focus on identifying a consistently low performing group of students whose academic functioning and progress cannot be supported and

improved with general education services alone. Federal and state regulations should seek to create a system that focuses more on student needs relative to local context and less on the categorization of students. Identification practices should result in a group of students who are similar in terms of their academic functioning and not biased in terms of other student characteristics such as race or socioeconomic status. When approached in this manner, MTSS is a viable and efficient identification and service delivery model for SLD despite the noted limitations with measuring ROI and growth. To support these efforts, Federal and state funding for special education should provide incentives for districts who implement high-quality, empirically-based curricula for students at all tiers and special education funding should be expanded to include the general education intervention supports of RTI systems.

RTI models for pre-referral intervention as well as the identification of SLD have been in existence prior to the passage of IDEA 2004 (Fuchs et al., 2003). Evaluation of these early models found that they resulted in improved outcomes at the student and systems level following their implementation as well as fewer students being identified with SLD (Burns, Appleton, & Stehouwer, 2005). It is worth noting that these models were operating before research had really begun attempting to operationalize and validate growth related to adequate/insufficient intervention response beyond visual inspection methods. Although ROI and RTI models have flaws, there is no reason to believe based on the available research that the lack of absolute criteria for quantifying ROI and student growth will result in the over-identification or the misidentification of students as SLD more than current practices or traditional discrepancy approaches.

In summary, despite some of the shortcomings of the model, RTI is still as valid as other methods for the identification of SLD available to schools and school psychologists with the

added benefit of improved overall school performance. The federal and state regulations surrounding these practices should emphasize how to determine the need for specially designed instruction over criteria for categorization. The MTSS in which RTI is embedded in place an emphasis on high-quality instruction for all students and help schools achieve better overall outcomes. An RTI approach to SLD eligibility by nature seems to address need over-categorization in that a student's response to instruction is the driving factor in determining eligibility, not eligibility criteria focusing in intellectual functioning or other factors. The use of repeated and direct measurement of academic skills also facilitates the identification of a group of students who are in the most need of specially designed instruction by local standards. The use of CBM data to screen, monitor student progress, and make referrals also helps to reduce bias in the referral and identification process. These are perhaps the most important factors to keep in mind when determining which students should be identified for special education services. So, by focusing on need (e.g., intervention response), the use of RTI procedures should result in a classification system with at least adequate technical properties using current recommended practices and many other benefits to schools and the communities they serve.

Implications for Practice

Based on the current and previous studies surrounding ROI, practitioners should be informed yet cautious when interpreting intra-individual ROI for the purposes of making special education referral and eligibility decisions. Although ROI clearly differentiates students who are at-risk for and have SLD from typically developing readers, ROI's ability to differentiate at-risk and disabled readers from each other has not been established. Given the available research, the recommendation by Kovaleski et al. (2013) to use a trajectory analysis seems most appropriate. In this sense, ROI is used to gauge how long and how many resources will need to be committed

to a student to make meaningful progress towards expected levels of functioning. In this way, similar to the results of the current study, level of performance is the main factor when determining referrals and eligibility for special education with ROI helping to determine how realistic it is to support a student with or without special education resources until they reach a more desirable level of performance.

Districts that are currently using RTI to make referrals for special education and special education eligibility decisions may wish to consider using percentiles for ROI to assist in this decision-making process. Legally and ethically, assessing growth when using RTI for SLD eligibility is still a requirement. By using local or national norms, problem-solving and MDE teams can use ROI in a consistent manner to guide decisions and increase the reliability of the groups of students they are identifying. Relying less on visual inspection methods that may be more subjective could also help prevent bias and other student characteristics from becoming more influential in the decision-making process. ROI as a measurement tool has its limitations that practitioners must be aware of but it is still a promising tool and an empirically-based method available for measuring growth when making RTI eligibility decisions.

Additionally, ROI is still a valuable tool for practitioners and problem-solving teams to use for instructional decision-making purposes. Data-based decision making is an important component of MTSS at all tiers and the OLS method for determining ROI still has substantial empirical support (Ardoin et al., 2013; Parker & Tindal, 1992; Shinn et al., 1989). Using guidelines established by Fuchs et al. (1993) for ROI as frame of reference, problem-solving teams have a research-based guide for low-stakes decisions such as determining if a student's intervention response is adequate or if more resources or more intensive intervention efforts are

needed. Although not without its limitations, using ROI in this manner has more empirical support than visual inspection methods.

As a final thought, perhaps RTI and ROI do not need to have precise and rigid eligibility criteria to be a viable alternative to traditional models of SLD identification. In the absence of such criteria, schools and states (e.g., Minneapolis Public Schools and Iowa) have been using RTI for SLD identification purposes with at least no worse results than previous practices (Fuchs et al., 2003, Marston et al., 2003). The majority of the data available suggests improvement not only in special education identification practices but school performance on a whole when RTI is used. The current study supports that despite the shortcomings of ROI in differentiating SLD students from at-risk students and predicting special education eligibility, students identified with SLD using RTI are performing at a consistently poorer academic level than their peers. This suggests that RTI is identifying a consistent group of students in terms of their level of academic functioning and educational need. This would seem to be the most important practical and ethical concern of any eligibility decision.

Summary

The current study examined the dual-discrepancy model and ROI with two research questions. The first research question was: Do students who are identified as SLD in reading using RTI procedures/methods differ from those students who are determined ineligible for special education and those students receiving supplemental intervention but not referred for evaluation on their ROI and/or level of performance? Although it was hypothesized that both level of performance and ROI would differentiate students in these three groups, only students who were eligible for special education were different on level of performance compared to students who were not referred for evaluation. This was consistent with previous research that

found level of performance on CBM-R probes differentiated students with SLD from their non-disabled peers but inconsistent with previous research that found that ROI identified a different, and more impaired, group of students than level of performance alone.

The second research question was: What student attributes best predict eligibility for special education using a dual discrepancy approach? It was hypothesized that level of performance and ROI would predict special education eligibility more than other student attributes. The hypothesis was only partially supported. Level of performance was the only variable that predicted special education eligibility. This was consistent with some previous research that found that ROI did not predict eligibility as SLD but inconsistent with literature that has supported that ROI helps to identify a group of students who are more impaired readers than level of performance alone.

The limitations of the current study include the use of a convenience sample and a limited number of subjects in the eligible and referred but not eligible conditions. Other limitations include the use of data from multiple grade-levels and the lack of a true and accurate test for SLD on which to base the dependent variable on. Future research on ROI should focus on establishing growth rates within and across grade-levels for students with SLD to assist in the eligibility decision-making process. Research should also examine if the exit groups model of RTI performs as adequately as the dual discrepancy model. At the policy level, it is advisable to move towards a need-based system of eligibility for SLD given the long history of difficulties with adequately defining the construct and failures at producing a technically adequate model of identification. RTI is still a viable and empirically-based service delivery model for the identification of SLD. It has been shown in field sites to identify an adequate group of students

as SLD while simultaneously improving school performance. RTI continues to have superior utility for practitioners when compared to other models of SLD identification.

References

- Ardoin, S. P., & Christ, T. J. (2008). Evaluating curriculum-based measurement slope estimates using data from triannual universal screenings. *School Psychology Review*, 27, 109-125. doi:10.17105/spr45-3.310-326
- 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.
<https://naspjournals.org>
- Ardoin, S. P., Christ, T. J., Morena, L. S., Cormier, D. C., & Klingbeil, D. A. (2013). A systematic review and summarization of the recommendations surrounding curriculum-based measurement of oral reading fluency (CBM-R) decision rules. *Journal of School Psychology*, 51, 1-18. doi:10.1016/j.jsp.2012.09.004
- Barth, A. E., Stuebing, K. K., Anthony, J. L., Denton, C. A., Mathes, P. G., Fletcher, J. M., & Francis, D. J. (2008). Agreement among response to intervention criteria for identifying responder status. *Learning and Individual Differences*, 18, 296-307. doi:10.1016/j.lindif.2008.04.004
- Benefits.Gov. (n.d.). Retrieved from <https://www.benefits.gov/benefit/2013>
- Boneshefski, M. J. (2017). *Differential diagnosis of specific learning disability within a response to intervention framework*. (Unpublished doctoral dissertation). Indiana University of Pennsylvania, Indiana, PA.

- Brown-Waesche, J. S., Schatschneider, C., Maner, J. K., Ahmed, Y., & Wagner, R. K. (2011). Examining agreement and longitudinal stability among traditional and RTI-based definitions of reading disability using affected-status agreement statistic. *Journal of Learning Disabilities, 44*, 296-307. doi:10.1177/0022219410392048
- Burns, M. K., 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. doi:10.1177/073428290502300406
- Burns, M. K., Scholin, S. E., Kosciulek, S. K., & Livingston, J. (2010). Reliability of decision-making frameworks for response to intervention for reading. *Journal of Psychoeducational Assessment, 28*, 102-114. doi:10.1177/0734282909342374
- Burns, M. K., & Senesac, B. V. (2005). Comparison of dual discrepancy criteria to assess response to intervention. *Journal of School Psychology, 23*, 393-406. doi:10.1016/j.jsp.2005.09.003
- 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. <https://naspjournals.org>
- Christ, T. J., & Silberglitt, B. (2007). Estimates of standard error of measurement for curriculum-based measures of oral reading fluency. *School Psychology Review, 36*, 130-146. <https://naspjournals.org>
- Christ, T. J., Zopluoglu, C., Long, J. D., & Monaghan, B. D. (2012). Curriculum-based measurement of oral reading: Quality of progress monitoring outcomes. *Exceptional Children, 78*, 356-373. doi:10.1177/001440291207800306

- Christ, T. J., Zopluoglu, C., Monaghan, B. D., & Van Norman, E. R. (2013). Curriculum-based measurement of oral reading: Multi-study evaluation of schedule, duration, and dataset quality on progress monitoring outcomes. *Journal of School Psychology, 50*, 19-57. doi:10.1016/j.jsp.2012.11.001
- Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*, 219-232. doi:10.1177/001440298505200303
- 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, 30*, 507-524. <https://naspjournals.org>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191. doi:10.3758/bf03193146
- Francis, D. J., Santi, K. L., Barr, C., Fletcher, J. M., Varisco, A., & Foorman, B. R. (2008). Form effects on the estimation of students' oral reading fluency using DIBELS. *Journal of School Psychology, 46*, 315-342. <https://doi.org/10.1016/j.jsp.2007.06.003>
- Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2007). *Learning Disabilities: From identification to intervention*. New York, NY: The Guilford Press.
- Fletcher, J. M., Vaugh, S., & Stuebing, K. K (2014). Patterns of cognitive strengths and weaknesses: Identification rates, agreement, and validity for learning disabilities identification. *School Psychology Quarterly, 29*, 21-27. doi:10.1037/spq0000037
- Flinn, C. S. (2015). *Recommended practices for determining rate of improvement: A review of state-level guidance*. (Unpublished doctoral dissertation). Indiana University of Pennsylvania, Indiana, PA.

- Fuchs, D., Compton, D. L., Fuchs, L.S., Bryant, J., & Davis, N. D. (2008). Making “secondary intervention” work in a three-tier responsiveness-to-intervention model: Findings from the first-grade longitudinal reading study of the National Research Center on Learning Disabilities. *Reading and Writing: An Interdisciplinary Journal*, 21, 413-436. doi:10.1007/s11145-007-9083-9
- Fuchs, D., & Deshler, D. (2007). What we need to know about response-to-intervention (and shouldn't be afraid to ask). *Learning Disabilities Practice*, 22, 129- 136. doi:10.1111/j.1540-5826.2007.00237.x
- Fuchs, D., Deshler, D. D., & Reschly, D. J. (2004). National research center on learning disabilities: Multimethod studies of identification and classification issues. *Learning Disability Quarterly*, 27, 189-195. doi:10.2307/1593672
- Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research & Practice*, 18, 157-171. doi:10.1111/1540-5826.00072
- Fuchs, D., Fuchs, L.S., & Compton, D. L. (2004). Identifying reading disabilities by responsiveness-to-instruction: Specifying measures and criteria. *Learning Disability Quarterly*, 27, 216-227. doi:10.2307/1593674
- Fuchs, L. S (2003). Assessing intervention responsiveness: Conceptual and technical issues. *Learning Disabilities Research & Practice*, 18, 172-186. doi:10.1111/1540-5826.00073
- Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, C. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review*, 22, 27-48. <https://naspjournals.org>

- Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research, 13*, 204-219. <https://onlinelibrary.wiley.com/journal/15405826>
- Germann, G., & Tindal, G. (1985). An application of curriculum-based assessment: The use of direct and repeated measurement. *Exceptional Children, 52*, 244-265. doi:10.1177/001440298505200306
- Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5*, 257-288. doi:10.1207/s1532799xssr0503_4
- Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher, 59*, 636-644. doi:10.1598/rt.59.7.3
- Hoskyn, M., & Swanson, H. L. (2000). Cognitive processing of low achievers and children with reading disabilities: A selective meta-analytic review of the published literature. *School Psychology Review, 29*, 102-119. <https://naspjournals.org>
- Hosp J. L., & Reschly, D. J. (2004). Disproportionate representation of minority students in special education: Academic, demographic, and economic predictors. *Exceptional Children, 70*, 185-199. doi:10.1177/001440290407000204
- Houck, S. W. (2012). *Reading statistics and research*. Boston, MA: Pearson Education.
- Individual with Disabilities Education Act, 20 U.S.C. 1400 (2004).
- Jenkins, J. R., Zumeta, R., & Dupree, O. (2005). Measuring gains in reading ability with passage reading fluency. *Learning Disabilities Research & Practice, 20*, 245-253. doi:10.1111/j.1540-5826.2005.00140.x

- Johnson, E., Mellard, D. F., Fuchs, D., & McKnight, M. A. (2006). *Responsiveness to intervention: How to do it*. Lawrence, KS: National Research Center on Learning Disabilities.
- Kovaleski, J. F., VanDerHeyden, A. M., & Shapiro, E. S. (2013). *The RTI approach to evaluating learning disabilities*. New York, NY: The Guilford Press.
- Laerd Statistics (n.d.). *Multinomial logistic regression using SPSS*. Retrieved from <https://statistics.laerd.com/spss-tutorials/multinomial-logistic-regression-using-spss-statistics.php>
- MacMillan, D. L., & Speece, D. L. (1999). Utility of current diagnostic categories for research and practice. In R. Gallimore, L.P. Bernheimer, D. L. MacMillan, D. L. Speece, and S. Vaughn (Eds.), *Developmental perspective on children with high-incidence disabilities* (pp. 111-133). Mahwah, NJ: Erlbaum.
- Marston, D., Fuchs, L. S., & Deno, S. (1986). Measuring pupil progress: A comparison of standardized achievement tests and curriculum-related measures. *Diagnostic*, 11, 77-90. doi:10.1177/073724778601100203
- Marston, D., Muyskens, P., Lau, M., & Canter, A. (2003). Problem-solving model for decision making with high-incidence disabilities: The Minneapolis experience. *Learning Disabilities Research & Practice*, 18, 187-200. doi:10.1111/1540-5826.00074
- Meisinger, E. B., Bradley, B. A., Schwanenflugal, P. J., Kuhn, M. R., & Morris, R. D. (2009). Myth and reality of the word caller: The relation between teacher nominations and prevalence among elementary school children. *School Psychology Quarterly*, 24, 147-159. doi:10.1037/a0017191

- Mercer, S. H., Lyons, A. F., Johnston, L. E., & Millhoff, C. L. (2015). Robust regression for slope estimation in curriculum-based measurement progress monitoring. *Assessment for Effective Intervention, 40*, 176-183. doi:10.1177/1534508414555705
- Nese, J. F. T., Biancarosa, G., Cummings, K., Kennedy, P., Alonzo, J., & Tindal, G. (2013). In search of average growth: Describing within-year oral reading fluency growth across grades 1-8. *Journal of School Psychology, 51*, 625-642. doi:10.1016/j.jsp.2013.05.006
- Pallant, J. (2016). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (6th ed.). New York: NY, McGraw Hill Education.
- Parker, R., & Tindal, G. (1992). Estimating trend in progress monitoring data: A comparison of simple line fitting methods. *School Psychology Review, 21*, 300-312.
<https://naspjournals.org>
- Pennsylvania Department of Education. (n.d.). Retrieved from http://pattan.net-website.s3.amazonaws.com/images/misc/2017/11/16/Approved%20RTI_list_103117.pdf
- Pennsylvania Department of Education. (2014). *Using response to instruction and intervention (RTII) for SLD determination elementary school/ building application for approval*. Retrieved from http://pattan.netwebsite.s3.amazonaws.com/images/instructional/2014/04/24/RtII_SLD_Determin042414.pdf
- Peterson, K. M. H. & Shinn, M. R. (2002). Severe discrepancy models: Which best explains school identification practices for learning disabilities? *School Psychology Review, 31*, 459-476. <https://naspjournals.org>

- Poncy, B. C., Skinner, C. H., & Axtell, P. K. (2005). An investigation of the reliability and standard error of measurement of words read correctly per minute using curriculum-based measurement. *Journal of Psychoeducational Assessment*, 23, 326-338. doi:10.1177/073428290502300403
- Reschly, D. J., & Bergstrom, M. K. (2009). Response to intervention. In T. B. Gutkin & C. R. Reynolds (Eds.), *The handbook of school psychology* (4th edition, pp. 434-460). Hoboken, N.J: John Wiley & Sons.
- Reynolds, C. R., & Shaywitz, S. E. (2009a). Response to intervention: Prevention and remediation, perhaps. diagnosis, no. *Child Development Perspectives*, 3, 44-47. doi:10.1111/j.1750-8606.2008.00075.x
- Reynolds, C. R. & Shaywitz, S. E. (2009b). Response to intervention: Ready or not? or from wait-to-fail to watch-them-fail. *School Psychology Quarterly*, 24, 130-145. doi:10.1037/a0016158
- Runge, T. J., Bennyhoff, C. F., Ferchalk, M. R., & McCrea, A. E. (2017). The role of measurement interval in rate of improvement calculation. *School Psychology Forum: Research in Practice*, 11, 77-90. <http://apps.nasponline.org/resources-and-publications/periodicals/spf-index-list.aspx>
- Shinn, M. R., Good, R. H., Stein, S. (1989). Summarizing trends in student achievement: A comparison of methods. *School Psychology Review*, 18, 356- 370. <https://naspjournals.org>
- Shinn, M. & Marston, D. (1985). Differentiating mildly handicapped, low-achieving, and regular education students: A curriculum-based approach. *Remedial and Special Education*, 6, 31-45. doi:10.1177/074193258500600207

- Shinn, M. R., Ysseldyke, J. E., Deno, S. L., & Tindal, G. A. (1986). A comparison of differences between students labeled learning disabled and low achieving on measures of classroom performance. *Journal of Learning Disabilities*, 9, 545-552. doi:10.1177/002221948601900906
- Silbergitt, B. & Hintze, J. M. (2007). How much growth can we expect: A conditional analysis of R-CBM growth rates by level of performance. *Exceptional Children*, 74(1), 71-84. doi:10.1177/001440290707400104
- Speece, D. L. & Case, L. P. (2001). Classification in context: An alternative approach to Identifying early reading disability. *Journal of Educational Psychology*, 93, 735-749. doi:10.1037/0022-0663.93.4.735
- Speece, D. L., Case, L. P., & Molloy, D. E. (2003). Responsiveness to general instruction as the first gate to learning disabilities identification. *Learning Disabilities Research & Practice*, 18, 147-156. doi:10.1111/1540-5826.00071
- Stoiber, K. C. (2014). A comprehensive framework for multitiered systems of support in school psychology. In P. L. Harrison & A. Thomas (Eds.), *Best practices in school psychology: Data-based and collaborative decision making* (pp. 41-70). Bethesda, MD: NASP Publications.
- Stuebing, K. K., Barth, A. E., Trahan, L. H., Radhika, R. R., Miciak, J., & Fletcher, J. M., (2014). Are child cognitive characteristics strong predictors of responses to intervention? A meta-analysis. *Review of Educational Research*, 20, 1-35. doi:10.3102/0034654314555996

- Stuebing, K. K., Fletcher, J. M., Branum-Martin, L. & Francis, D. J. (2012). Evaluation of the technical adequacy of three methods of identifying specific learning disabilities based on cognitive discrepancies. *School Psychology Review*, 41(1), 3-22. <https://naspjournals.org>
- Stuebing, K. K., Fletcher, J. M., LeDoux, J. M., Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2002). Validity of IQ-discrepancy classification of reading disabilities: A meta-analysis. *American Educational Research Journal*, 39, 469-518. doi:10.3102/0002831203900246
- 9
- Torgesen, J. K. (2007). Using an RTI model to guide early reading instruction: Effects on identification rates for students with learning disabilities. Retrieved from: http://www.fcrr.org/science/pdf/torgesen/Response_intervention_Florida.pdf.
- Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., & Conway, T. (2001). Intensive remedial reading instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities*, 34, 33-58. doi:10.1177/002221940103400104
- U.S. Office of Education. (1977). Assistance to states for education for handicapped children: Procedures for evaluating specific learning disabilities. *Federal Register*, 42, G1082-G1085.
- Vannest, K. J., Parker, R. I., Davis, J. L., Soares, D. A., & Smith, S. L. (2012). The Thiel-Sen slope for high-stakes decisions from progress monitoring. *Behavioral Disorders*, 37, 271-280. doi:10.1177/019874291203700406
- Vaughn, S., Linan-Thompson, S., & Hickman, P. (2003). Response to instruction as a means of identifying students with reading/ learning disabilities. *Exceptional Children*, 69, 391-409. doi:10.1177/001440290306900401

- Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, G. S., Chen, R., Pratt, A., & Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology*, 88, 601-638. doi:10.1037//0022-0663.88.4.601
- Wayman, M. M., Wallace, T., Wiley, H. I., Ticha, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education*, 41, 85-120. doi:10.1177/00224669070410020401
- Wright, R. E. (1995). Logistic regression. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (217-244). Washington, D. C.: American Psychological Association.
- Ysseldyke, J. E., Thurlow, M. L., Graden, J. L., Deno, S. L., & Algozzine, B. (1982). Generalizations from five years of research on assessment and decision making. *Institute for Research on Learning Disabilities*. University of Minnesota: Minneapolis. doi:10.1177/074193258300400109
- Zirkel, P. A., & Thomas, L. B. (2010a). State laws for RTI: An updated snapshot. *Teaching Exceptional Children*, 42, 56- 63. doi:10.1177/004005991004200306
- Zirkel, P. A., & Thomas, L. B. (2010b). State laws and guidelines for implementing RTI. *Teaching Exceptional Children*, 43, 60-73. doi:10.1177/004005991004300107

Appendix A

Institutional Review Board Approval



Indiana University of Pennsylvania
www.iup.edu

Institutional Review Board for the
Protection of Human Subjects
School of Graduate Studies and Research
Stright Hall, Room 113
210 South Tenth Street
Indiana, Pennsylvania 15705-1048

P 724-357-7730
F 724-357-2715
irb-research@iup.edu
www.iup.edu/irb

April 17, 2017

Dear Mr. Drew Hunter:

Your proposed research project, "Using Response-to-Intervention to Determine Eligibility for Specific Learning Disabilities in Reading: What is a Non-Responder?," (Log No. 17-137) has been reviewed by the IRB and is approved. In accordance with 45CFR46.101 and IUP Policy, your project is exempt from continuing review. This approval does not supersede or obviate compliance with any other University requirements, including, but not limited to, enrollment, degree completion deadlines, topic approval, and conduct of university-affiliated activities.

You should read all of this letter, as it contains important information about conducting your study.

Now that your project has been approved by the IRB, there are elements of the Federal Regulations to which you must attend. IUP adheres to these regulations strictly:

1. You must conduct your study exactly as it was approved by the IRB.
2. Any additions or changes in procedures must be approved by the IRB before they are implemented.
3. You must notify the IRB promptly of any events that affect the safety or well-being of subjects.
4. You must notify the IRB promptly of any modifications of your study or other responses that are necessitated by any events reported in items 2 or 3.

The IRB may review or audit your project at random or for cause. In accordance with IUP Policy and Federal Regulation (45CFR46.113), the Board may suspend or terminate your project if your project has not been conducted as approved or if other difficulties are detected

Although your human subjects review process is complete, the School of Graduate Studies and Research requires submission and approval of a Research Topic Approval Form (RTAF) before you can begin your research. If you have not yet submitted your RTAF, the form can be found at <http://www.iup.edu/page.aspx?id=91683> .

While not under the purview of the IRB, researchers are responsible for adhering to US copyright law when using existing scales, survey items, or other works in the conduct of research. Information regarding copyright law and compliance at IUP, including links to sample permission request letters, can be found at <http://www.iup.edu/page.aspx?id=165526>.

I wish you success as you pursue this important endeavor.

Sincerely,

Jennifer Roberts, Ph.D.
Chairperson, Institutional Review Board for the Protection of Human Subjects
Professor of Criminology

JLR:jeb

Cc: Dr. Joseph Kovalski, Dissertation Advisor

Appendix B

Institutional Review Board Modification of Project



Indiana University of Pennsylvania

www.iup.edu

Institutional Review Board for the
Protection of Human Subjects
School of Graduate Studies and Research
Stright Hall, Room 113
210 South Tenth Street
Indiana, Pennsylvania 15705-1048

P 724-357-7730
F 724-357-2715
irb-research@iup.edu
www.iup.edu/irb

May 21, 2018

Dear Mr. Drew Hunter

Your proposed modifications to your previously approved research project, "Using Response-to-Intervention to Determine Eligibility for Specific Learning Disabilities in Reading: What is a Non-Responder?," (Log No. 17-137) been reviewed by the IRB and are approved. In accordance with 45CFR46.101 and IUP Policy, your project is exempt from continuing review in addition to the approval of your request for changes. This approval does not supersede or obviate compliance with any other University requirements, including, but not limited to, enrollment, degree completion deadlines, topic approval, and conduct of university-affiliated activities.

You should read all of this letter, as it contains important information about conducting your study.

Now that your project has been approved by the IRB, there are elements of the Federal Regulations to which you must attend. IUP adheres to these regulations strictly:

1. You must conduct your study exactly as it was approved by the IRB.
2. Any additions or changes in procedures must be approved by the IRB before they are implemented.
3. You must notify the IRB promptly of any events that affect the safety or well-being of subjects.
4. You must notify the IRB promptly of any modifications of your study or other responses that are necessitated by any events

reported in items 2 or 3.

The IRB may review or audit your project at random or for cause. In accordance with IUP Policy and Federal Regulation (45CFR46.113), the Board may suspend or terminate your project if your project has not been conducted as approved or if other difficulties are detected.

While not under the purview of the IRB, researchers are responsible for adhering to US copyright law when using existing scales, survey items, or other works in the conduct of research. Information regarding copyright law and compliance at IUP, including links to sample permission request letters, can be found at <http://www.iup.edu/page.aspx?id=165526>.

IRB to Mr. Drew Hunter,
May 21, 2018

I wish you success as you pursue this
important endeavor. Sincerely,

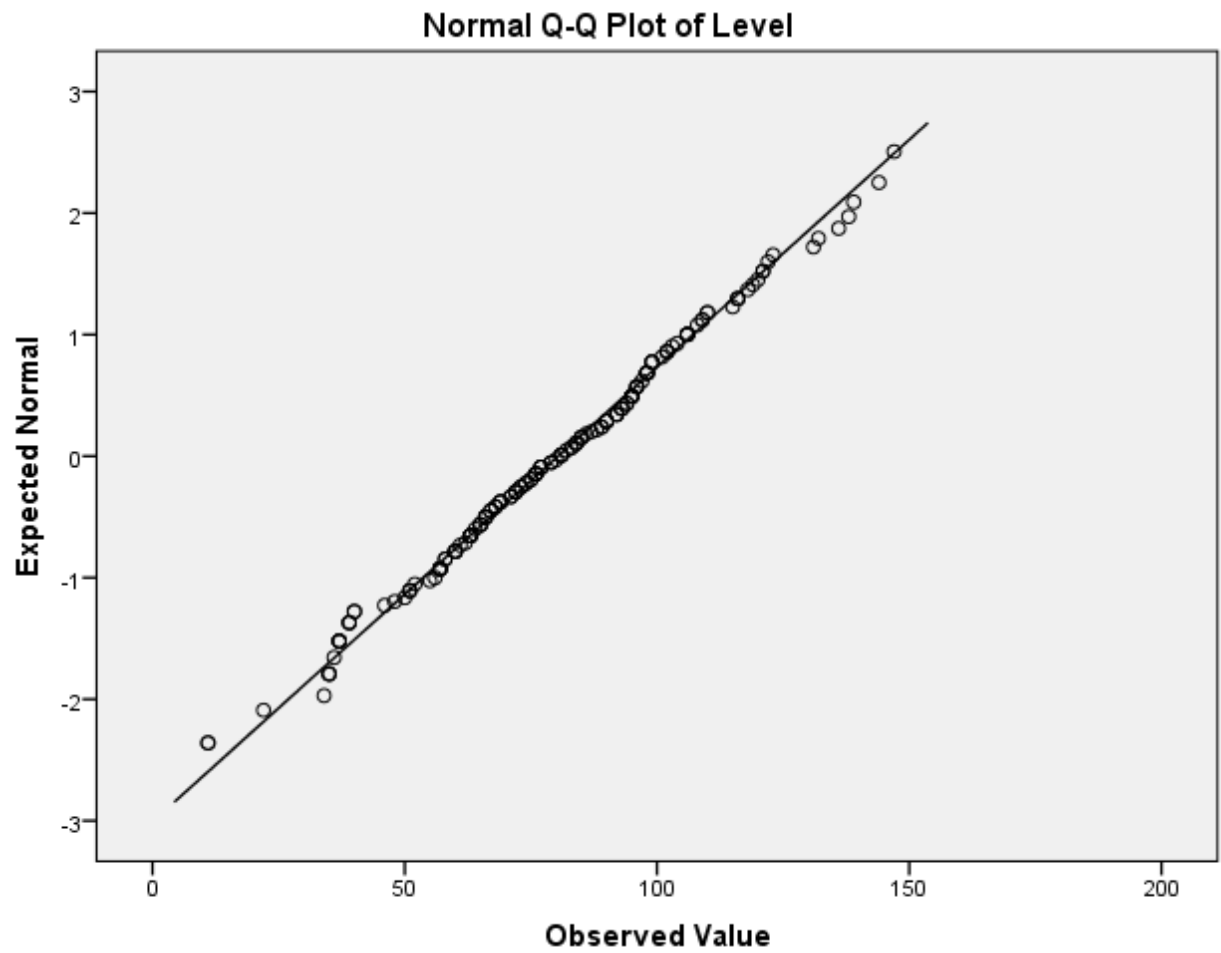
Timothy Runge, Ph.D.
Interim Chairperson, Institutional Review Board for the Protection of Human
Subjects Professor of Educational and School Psychology

TJR:bkj

Cc: Dr. Timothy Runge, Faculty Advisor

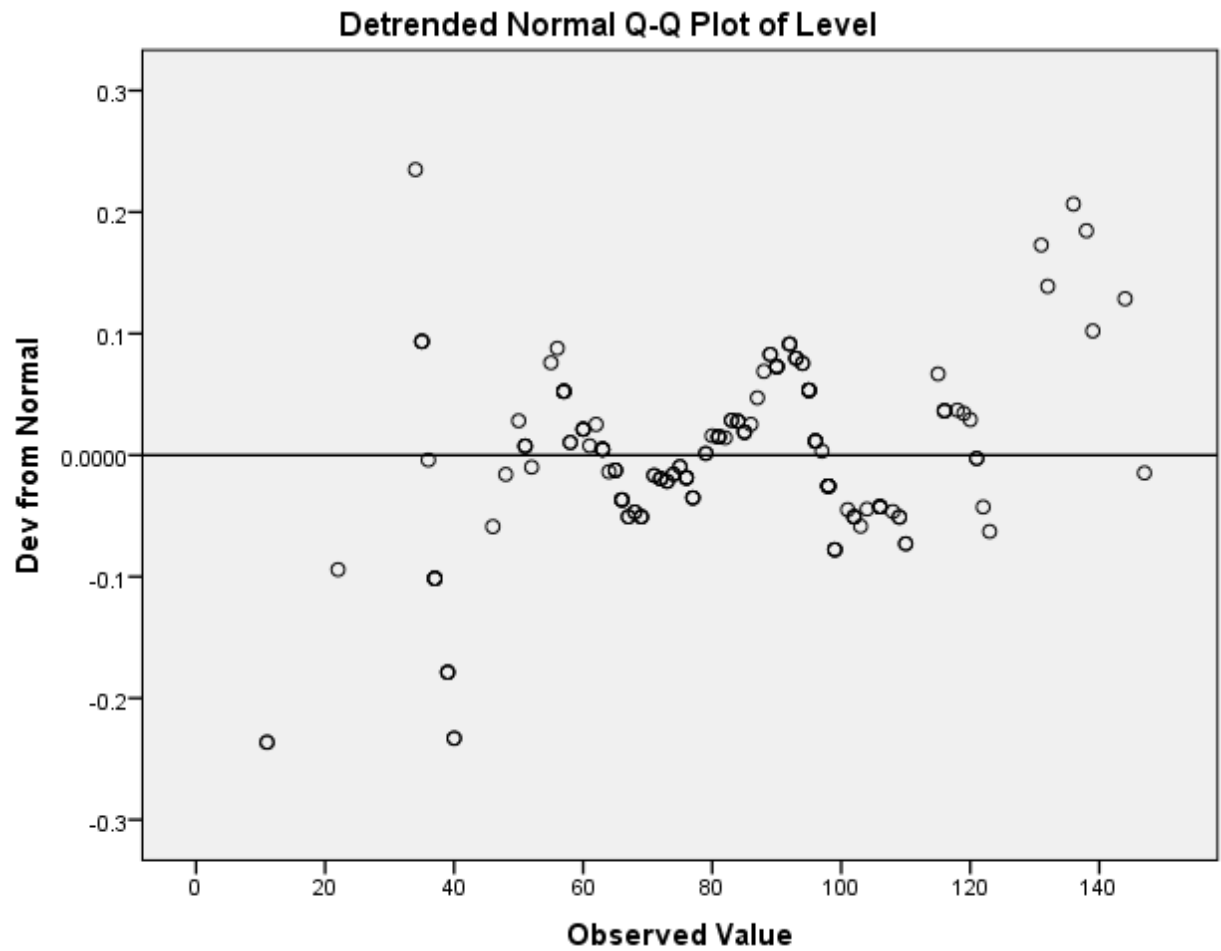
Appendix C

Normal Q-Q Plots for the Dependent Variable of Level of Performance for Research Question 1



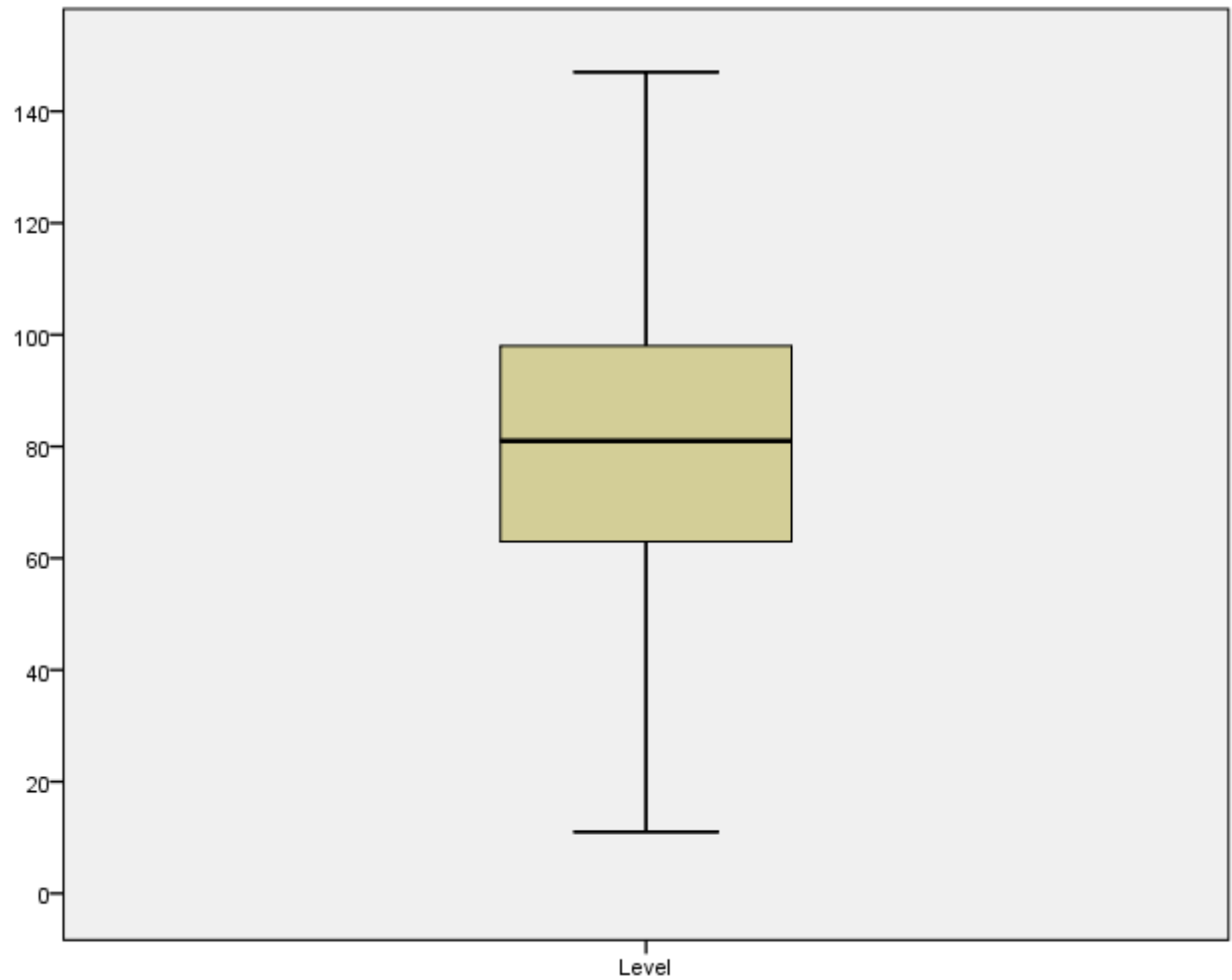
Appendix D

Detrended Q-Q Plots for the Dependent Variable of Level of Performance for Research Question
1



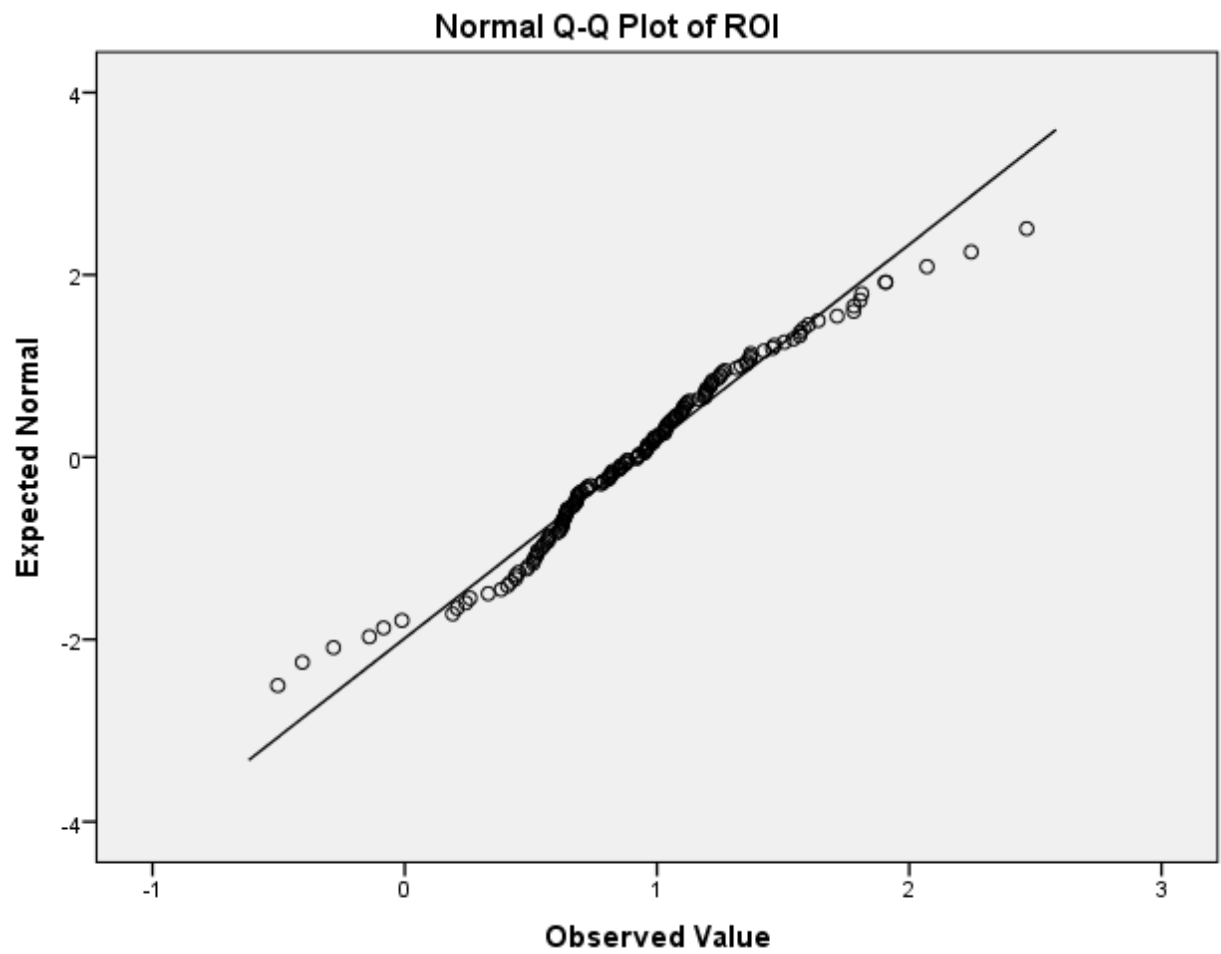
Appendix E

Box and Whisker Plots for the Dependent Variable of Level of Performance for Research Question 1



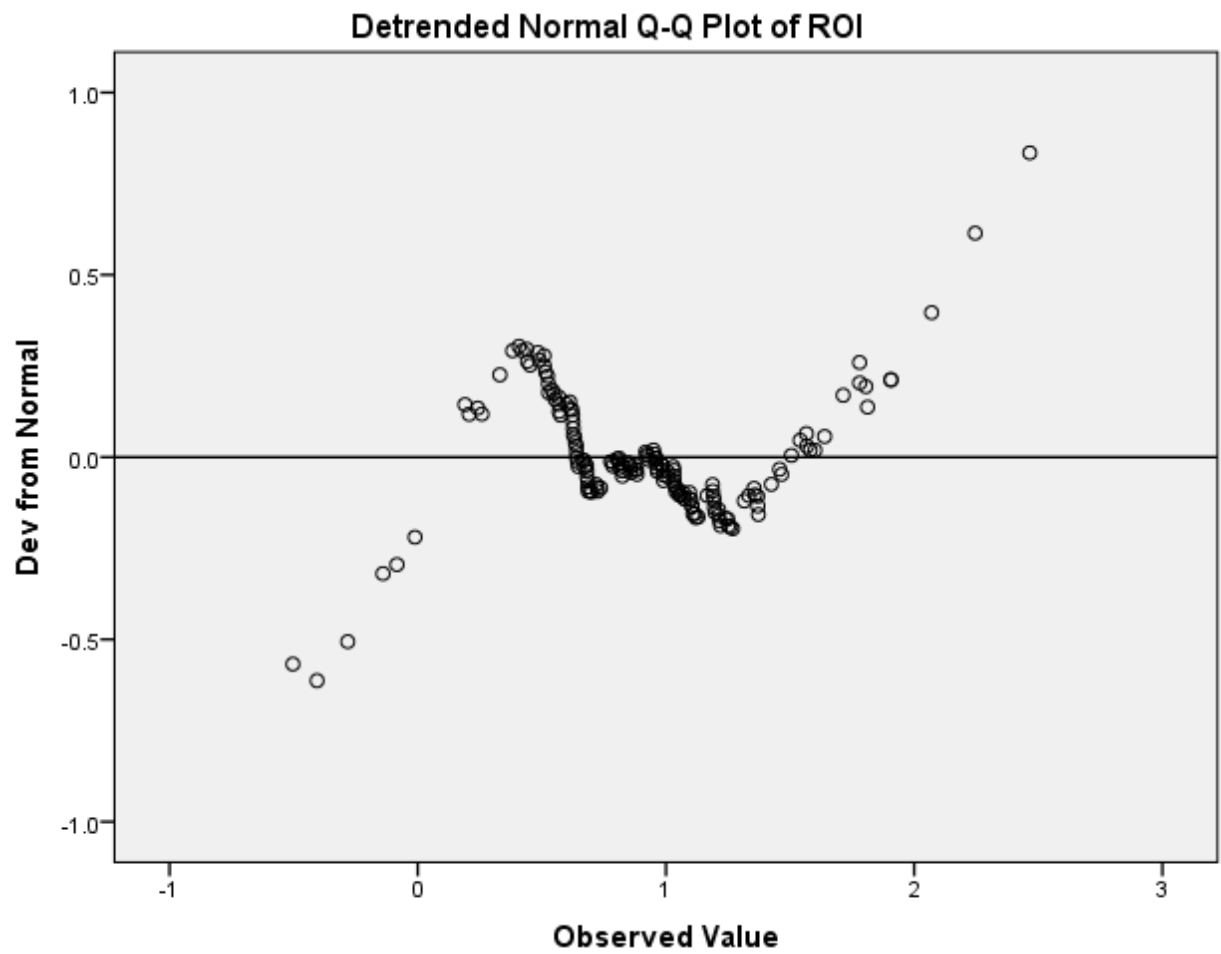
Appendix F

Normal Q-Q Plots for the Dependent Variable of ROI in Research Question 1



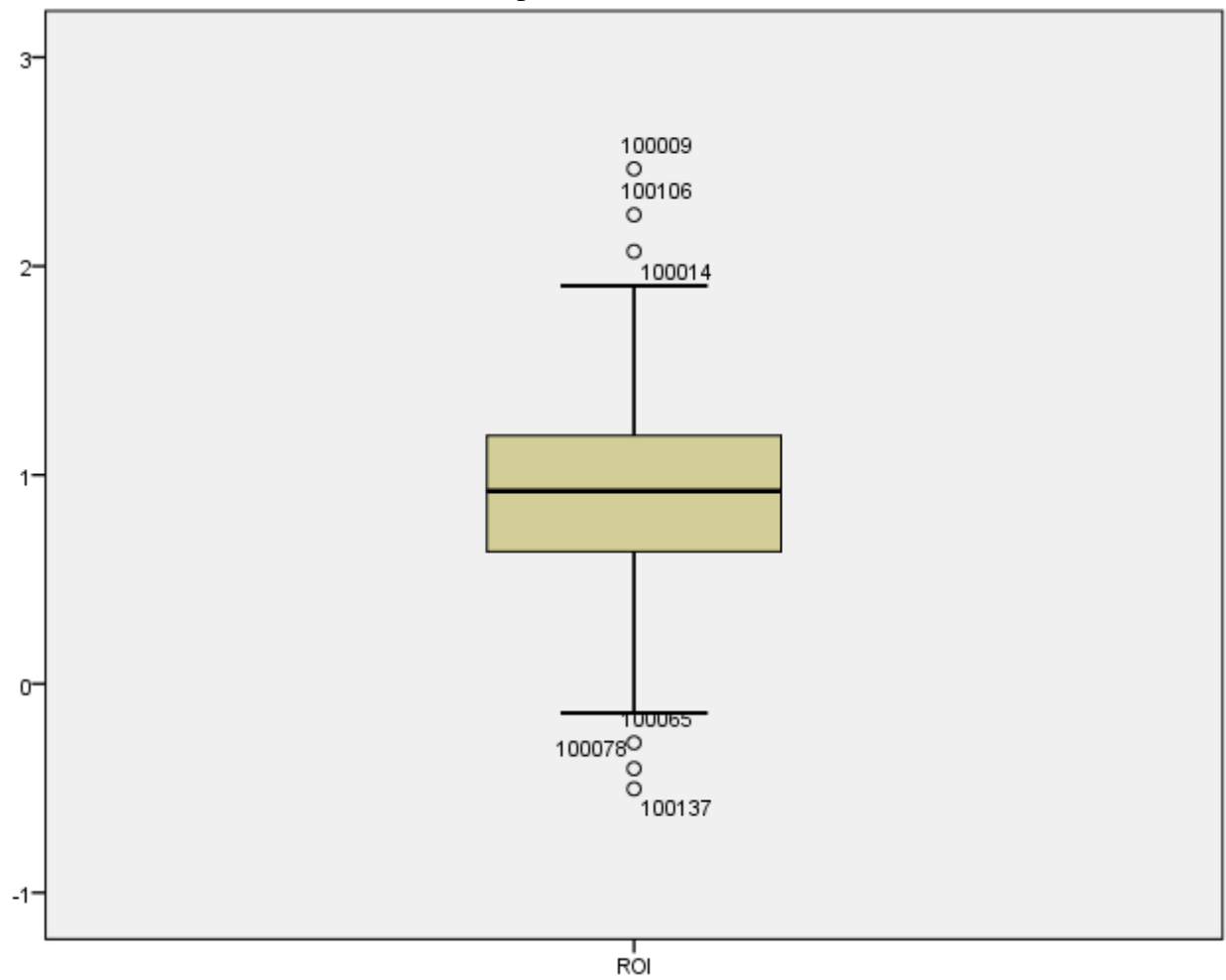
Appendix G

Detrended Q-Q Plots for the Dependent Variable of ROI for Research Question 1



Appendix H

Box and Whisker Plots for the Dependent Variable of ROI for Research Question 1



Appendix I

Variance Inflation Factor for the Independent Variables of Research Question 2

Coefficients^a							
Model	Unstandardized Coefficients		Standardized Coefficients		<i>p</i>	Collinearity Statistics	
	<i>B</i>	<i>SE</i>	Beta	<i>t</i>		Tolerance	VIF
(Constant)	-.52	.02		-30.87	.000		
Sex Categories	.01	.01	.01	1.48	.140	.94	1.07
Eligibility Categories	.53	.00	1	130.41	.000	.9	1.12
Lunch Categories	-.01	.01	-.01	-.99	.326	.91	1.11
Level	.00	.00	.01	1.17	.245	.86	1.16
ROI	.00	.01	.00	.58	.561	.96	1.04

a. Dependent Variable: Log eligibility

Appendix J

Studentized Residuals for Binomial Logistic Regression

Casewise List^b

Case	Selected Status ^a	Observed	Predicted	Predicted	Temporary Variable	
		Eligibility			Resid	ZResid
1	S	E**	.95	I	-.95	-4.27
37	S	E**	.97	I	-.98	-5.81
53	S	E**	.99	I	-.99	-9.72
54	S	E**	.96	I	-.96	-4.73
120	S	E**	.93	I	-.93	-3.77

a. S = Selected, U = Unselected cases, and ** = Misclassified cases.

b. Cases with studentized residuals greater than 2.000 are listed.