RESPONSE TO INTERVENTION: A MIXED-METHODS LONGITUDINAL STUDY OF THE DIFFERENCES IN READING ACHIEVEMENT WITHIN A SCHOOL DISTRICT

A DISSERTATION
SUBMITTED TO THE GRADUATE SCHOOL IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE DOCTOR OF EDUCATION
BY LAURA M. MILLER DISSERTATION ADVISOR: DR. MARILYNN QUICK

BALL STATE UNIVERSITY MUNCIE, INDIANA DECEMBER 2016
RESPONSE TO INTERVENTION: A MIXED-METHODS LONGITUDINAL STUDY OF THE DIFFERENCES IN READING ACHIEVEMENT WITHIN A SCHOOL DISTRICT

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

DOCTOR OF EDUCATION

BY

LAURA M. MILLER

DISSERTATION ADVISOR: DR. MARILYNN QUICK

APPROVED BY:

________________________________________ Date
Committee Chairperson

________________________________________ Date
Committee Member

________________________________________ Date
Committee Member

________________________________________ Date
Committee Member

________________________________________ Date
Dean of Graduate School

BALL STATE UNIVERSITY
MUNCIE, INDIANA
DECEMBER 2016
Reading is essential to success in our society and closing the achievement gap is of critical concern. Although research is ample on the components of Response to Intervention (RTI), minimal research is available on actual implementation. For these reasons, this mixed-methods longitudinal study examined the implications of RTI on student outcomes based on pre/post-tests analysis, identification for a specific learning disability, and educator perception. A nonequivalent group design was used to examine the differences on reading achievement for students who participated in intervention in kindergarten or first grade as compared to students who scored slightly above them and did not receive intervention. Cohorts of students from three elementary schools ($n = 639$) were measured through third grade. In addition, differences by gender, socio-economic status, and program structures were analyzed. To deepen the understanding of the results, data were triangulated through educator focus group interviews.

The findings from this study showed that students who received intervention substantially closed the gap between themselves and non-intervention students. With regards to poverty and gender, this study illustrates that these social injustices can be ameliorated. Evidence from this study, while not extensive, suggests that RTI contributed to a reduction in special education
placement. These outcomes support the converging evidence endorsing the RTI framework.

Given these results, further field studies need to be undertaken to allow for additional generalization. In times of increased academic demands from society, improved student achievement through program effectiveness is more critical than ever.
DEDICATION

I would like to dedicate my dissertation work to my family for their never-ending support.

Oh, and yes, to my family and friends, no more degrees. I promise!
ACKNOWLEDGEMENTS

To my loving husband, Craig: Thank you standing by me through this journey with the utmost encouragement and support, including leaving me alone when needed. Your support has allowed me to pursue this endeavor.

To my mother and father, Martha and Wayne: Thank you for a lifetime of support – emotionally, ethically, and spiritually - through your unending love.

To my sister and brother, Ann and John: Thank you for listening, telling me everyday stories to help me decompress, being there for me, and nudging me on when I needed it most.

To my nieces, Julia and Kayla: Thank you for allowing me to miss some of our time together and providing me much needed hugs along the way.

To my friends: Thank you for your support, saving my sanity, and pushing me through these last years. I hope each of you know you are family to me.

To Dr. Marilynn Quick, Doctoral Chairperson: Thank you for guiding me through the dissertation maze. In many ways, you are the most critical person that I need to thank. Without your support, encouragement, insights, and feedback, I may not have continued to forge on and accomplish this monumental endeavor. On a personal note, I want you to know how much I appreciate your understanding of my personal circumstances and the friendship you showed me as I was working through this process and life. You have been an invaluable source of knowledge, guidance, and support.
To Dr. Serena Salloum, Dr. Terry Wiedmer, and Dr. John Ellis, my dissertation committee:
Thank you for your support through this process. The extra things you have done did not go unnoticed. I am truly grateful for the contributions and insight to improve my work.

To Dr. Kianre Eouanzouli, Statistician: Words cannot express my gratitude for your patience and the support you showed through the data analysis. You helped me deepen my understanding of the data. Thank you for your assistance.

To the Educational Leadership faculty at Ball State University: Thank you for setting high standards and instilling a desire in me to meet or exceed your expectations.

Finally, to my school colleagues: I am especially appreciative of all of the administrators and teachers who gave of their time to participate in the study and who offered words of encouragement along the way. I do need to specifically mention one person, Megan White. Your assistance in compiling the data was remarkable. I could not find a better group of educators or people to share a large portion of my life with on a daily basis.

To all who have walked with me, I humbly thank you.
# TABLE OF CONTENTS

CHAPTER 1 ........................................................................................................................................ 17

Statement of the Problem .............................................................................................................. 18

Purpose of the Study ..................................................................................................................... 21

Significance of the Study ............................................................................................................... 22

Research Questions ...................................................................................................................... 23

Assumptions .................................................................................................................................. 24

Delimitations .................................................................................................................................. 24

Definition of Terms ....................................................................................................................... 25

Summary ......................................................................................................................................... 29

CHAPTER 2 ......................................................................................................................................... 31

Purpose of the Study ..................................................................................................................... 31

Research Questions ...................................................................................................................... 32

Specific Learning Disability ......................................................................................................... 32

Discrepancy Model Issues ........................................................................................................... 34

Atheoretical ..................................................................................................................................... 35

“Wait-to-fail” Model .................................................................................................................... 36

Quality of Instruction ..................................................................................................................... 37
Data-based Decision Making.................................................................64
Relationship with Special Education.......................................................65
RTI Approaches.....................................................................................66
  Standard treatment protocol..............................................................66
  Problem-solving protocol................................................................67
  Hybrid protocol..................................................................................69
Indiana RTI: A Response to Instruction ..................................................70
Non-responsiveness..............................................................................72
Considerations for SLD Identification.....................................................73
Summary...............................................................................................76
CHAPTER 3............................................................................................79
Purpose of the Study.............................................................................79
Research Questions..............................................................................80
Research Design..................................................................................80
Quantitative Research Design ...............................................................85
Qualitative Research Design .................................................................86
  Trustworthiness................................................................................87
Population and Sample.........................................................................88
Environment.........................................................................................91
Treatment/Intervention Protocols.........................................................93
Focus Group Participants....................................................................94
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instrumentation</td>
<td>95</td>
</tr>
<tr>
<td>mClass Reading 3D</td>
<td>96</td>
</tr>
<tr>
<td>DIBELS</td>
<td>97</td>
</tr>
<tr>
<td>TRC</td>
<td>99</td>
</tr>
<tr>
<td>Scholastic Reading Inventory</td>
<td>99</td>
</tr>
<tr>
<td>Indiana State Assessment: ISTEP+ and IREAD-3</td>
<td>100</td>
</tr>
<tr>
<td>Qualitative Instrument: Guiding Interview Questions</td>
<td>102</td>
</tr>
<tr>
<td>Data Collection</td>
<td>103</td>
</tr>
<tr>
<td>Quantitative Data Collection</td>
<td>103</td>
</tr>
<tr>
<td>Qualitative Data Collection</td>
<td>104</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>106</td>
</tr>
<tr>
<td>Quantitative Data Analysis</td>
<td>106</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>108</td>
</tr>
<tr>
<td>Researcher's Perspective</td>
<td>109</td>
</tr>
<tr>
<td>Limitations of the Study</td>
<td>110</td>
</tr>
<tr>
<td>Summary</td>
<td>110</td>
</tr>
<tr>
<td>CHAPTER 4</td>
<td>112</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>112</td>
</tr>
<tr>
<td>Research Questions</td>
<td>113</td>
</tr>
<tr>
<td>Participant Demographics</td>
<td>114</td>
</tr>
<tr>
<td>Student Cohort Demographics</td>
<td>114</td>
</tr>
</tbody>
</table>
Gender by nested group ................................................................. 119
Socio-economic by nested group .................................................. 121
Special education by nested group ................................................ 122

Results to the Research Questions .................................................. 124

Research Question 1: What are the long-range (3 or 4 years) differences on reading achievement for students who participated in RTI Tier 2 reading intervention as compared to students who scored slightly above the intervention group and did not receive treatment? .................................................. 124

Research Question 2: What differences occur among the RTI Tier 2 groups as compared to those students just outside the intervention group in regards to special education identification by the end of third grade? ................................................................. 133

Research Question 3: What differences are seen in students' reading achievement among and within the various nested groups (e.g. school, implementation year, achievement measures)? ................. 138

Research Question 4: From the perspective of the focus group, what understandings result from the quantitative data analysis of the academic outcomes between and among the nested groups? ............... 145

Perspective of practitioners on long-range differences on reading achievement ............... 146
Perspective of practitioners on special education outcomes ........................................ 148
Perspective of practitioners on differences in nested groups ....................................... 149
Perspective of practitioners on logistics and future needs ............................................ 151

Quality of Tier 1 instruction .................................................................. 152
Size of the intervention group ............................................................... 152
Organization of the program to meet individual student's needs ................................. 153
Staffing aspects .................................................................................. 153
Future needs for stability........................................................................................................155

Summary .................................................................................................................................156

CHAPTER 5 ...............................................................................................................................157

Overview of the Problem..........................................................................................................157

Purpose of the Study................................................................................................................158

Research Questions .................................................................................................................159

Review of the Research Methods ..........................................................................................159

Limitations of the study. ..........................................................................................................161

Major Findings ........................................................................................................................164

Findings Related to the Literature .........................................................................................164

Research Question 1: What are the long-range (3 or 4 years) differences on reading achievement for students who participated in RTI Tier 2 reading intervention as compared to students who scored slightly above the intervention group and did not receive treatment? .................................................165

Conceptual Framework .........................................................................................................170

Section summary ....................................................................................................................171

Research Question 2: What differences occur among the RTI Tier 2 groups as compared to those students just outside the intervention group in regards to special education identification by the end of third grade? ............................................................................................................................172

Conceptual Framework .........................................................................................................175

Section summary ....................................................................................................................176

Research Question 3: What differences are seen in students' reading achievement among and within the various nested groups (e.g. school, implementation year, achievement measures)? .................................................177
Research Question 4: From the perspective of the focus group, what understandings result from the quantitative data analysis of the academic outcomes between and among the nested groups? ....183

- Quality instructional delivery in Tier 1 ............................................................................. 185
- Intervention group size ...................................................................................................... 186
- Organization of the program to meet student needs ......................................................... 187
- Various staffing aspects .................................................................................................... 188
- Matching the staff member to the student Tier 2 intervention groups ............................ 188
- Ongoing professional development for Tier 2 instructors ................................................ 189
- Leadership ........................................................................................................................ 190
- Summary perception ........................................................................................................ 190

Conclusions ............................................................................................................................. 191

Implications for Action ........................................................................................................ 193

Recommendations for Further Research .......................................................................... 194

Recommendations for District Practitioners ..................................................................... 196

References ............................................................................................................................ 198

Appendix A ............................................................................................................................ 232

Appendix B ............................................................................................................................ 233

Appendix C ............................................................................................................................ 234

Appendix D ............................................................................................................................ 235

Appendix E ............................................................................................................................ 237

Appendix F ............................................................................................................................ 239
LIST OF FIGURES AND TABLES

Distribution of Children Served for a Specific Learning Disability by Race/Ethnicity .......................... 38
RTI Data Decision Cycle. ........................................................................................................................ 50
RTI Data-Based Decision Making Model ............................................................................................... 51
Instructional Data-driven Decision Making Process for RTI. ................................................................. 51
Pyramid RTI Framework .......................................................................................................................... 55
RTI Component Summary ........................................................................................................................ 60
Brief Overview of Key Problem-solving Protocol Studies ...................................................................... 69
Cylinder Model of RtI Framework ..................................................................................................... 71
Visual Model for Exploratory Sequential Mixed-Methods ........................................................................ 84
Design Notation ....................................................................................................................................... 86
Cohort Groups .......................................................................................................................................... 90
Sample Nested Groups ............................................................................................................................. 91
Distinctions in RTI, Both Indiana and the Study District, and Traditional Approaches ..................... 92
Focus Group Participants ......................................................................................................................... 95
District Comprehensive Assessment Plan ............................................................................................... 95
Cohort Frequency ..................................................................................................................................... 117
Intervention and Non-intervention Demographics ................................................................................ 118
Cross tab: Receive intervention by cohort with gender .......................................................................... 120
Gender Distribution by Cycle within the Intervention Subgroup ........................................................... 120
Male Gender by School by Subgroup

Cross tab: Receive intervention by cohort with SES

Cross tab: Receive Intervention by School with Socio-economic Status

Cross tab: Receive Intervention by Cohort with Including in Special Education

Cross-tab: Receive Intervention by School with Including in Special Education

Graphical Representation of the General Linear Model Results of the Difference in Gains

between the Intervention and Non-intervention Subgroups

Pre/Post-Test Descriptive Levels by Subgroup

Mean Based on the Assessment Score for Each Pre- and Post-test by Subgroup

Mean Based on Failure [Pass (1) and Fail (0)] for Each Pre- and Post-test by Subgroup

Mean for the Rate of Passing by Inclusion in Intervention

Odds Ratio Comparing Pre- and Post-Tests

Cross-tab: Chi-square and Phi Coefficient for the cross-tab of receiving intervention with inclusion in Special education by school, implementation year, or intervention grade

Graph of Estimated Marginal Means of Gain for the Two-way ANOVA of the Effect of Group Inclusion and Identification for Special Education

Two-way ANOVA Interaction Effect for Various Variables

Graph of the District Special Education Percent

Graph of the School Special Education Percent

Cross tab: Pre-test and Post-test by Intervention by School

One-way ANOVA Post Hoc Test: Year of Implementation

Mean Gain from TERA-3 NCE to SRI NCE

Odds Ratio for the Failure Rate by Year of Implementation
CHAPTER 1
INTRODUCTION

The 21st Century information-based society is dependent on a literate population with the ability to read as a key component (Center for Educational Research and Innovation, 1992; EnGauge, 2003). Not only is the ability to decode words important, but reading for depth of knowledge is essential in today's world. Learning from, interpreting, and analyzing complex text-based materials are critical skills. To ensure a literate population, schools must develop strong readers through effective teaching in the core reading program and implementation of effective and timely intervention strategies for students who struggle.

For students who struggle in reading, research indicates that intentional and intensive interventions should begin no later than first grade (i.e. six or seven years of age) (Bursuck & Blanks, 2010). Prior to the Response to intervention (RTI) method, the discrepancy model required a severe discrepancy between achievement and intellectual ability. Under the reign of the discrepancy model, students were on average nine-years-old before receiving supplemental support (Lyon, 2002). The discrepancy model delayed support and was unsatisfactory for meeting students' needs.

In a world where "reading is the gateway skill that makes all other learning possible...[and] necessary for survival, how can we send our kids out into the world [under-literate]?” (Obama, 2005). Change was necessary. RTI represents the much needed change in the process of supporting struggling students and merits investigation (Fuchs & Fuchs, 2006;

This first chapter presents the background of the study, details the problem, and illustrates the significance of the research of RTI. The chapter concludes by informing the reader of the delimitations of the study and defining several key terms used.

Statement of the Problem

In the Information Age, the intellectual capital of citizens is the driving economic and social force of a society (EnGauge, 2003); therefore, society requires a population with strong literacy skills (i.e. reading) to support the necessary level of education to drive the economy (CERI, 1992). By 2020, it is estimated that roughly 65% of jobs created will require some form of post-secondary education, as compared to 1973 when 72% of jobs required only a high school diploma (Carnevale & Desrochers, 2003; Carnevale, Smith, & Strohl, 2010; Carnevale, Smith, & Strohl, 2013). In a speech by then Senator Obama (2005), he stated, "literacy is the most basic currency of the knowledge economy we're living in today."

Reading is a gateway skill to learning and fundamental to economic security in a knowledge-based society; nevertheless, statistics do not show a positive trend towards higher levels of literacy. Studies reveal that 25% of secondary students were unable to comprehend the material in their textbooks (NCTE, 2006; NCES, 2009; NCES, 2011). In addition, between 1992 and 2003, literacy scores actually dropped and the reading achievement of certain racial, ethnic, and socio-economic groups continued to be significantly lower than their counterparts (NCTE, 2006; NCES, 2009; NCES, 2011). As stated in NCTE Principles of Adolescent Literacy Reform (2006), “our nation cannot afford an under-literate workforce” (p. 4). Individuals must be able to read complex text, pluck embedded ideas from the text, understand subtle relationships within
the text, grasp complicated information, negotiate sophisticated structures, maneuver within elaborate and unconventional styles, comprehend context vocabulary, and distinguish implicit, sometime ambiguous, purposes from the text (ACT, 2006; NCTE, 2006). Consequently, people's ability to learn through reading plays a significant role in their financial security throughout life.

In a society where literacy is critical for survival, these unfortunate statistics are in part the result of the discrepancy model (Fuch, 2006; Hintze, 2008; Martinez, Nellis, & Prendergast, 2006). The discrepancy model was the traditional method of identifying learning disabilities and required a student to demonstrate underachievement (i.e. typically two years below grade level) in comparison to the student's ability (i.e. cognitive I.Q.). In 2004, 42% of students identified for special education fell in the category of Specific Learning Disability (SLD) \(^1\) (USDOE, 2013). This represented 5.7% of the total school-age population. Under the reign of the discrepancy model, neglect has led to long-term issues and in some populations students were over identified (Coffey & Obringer, 2000). According to the Commission on Excellence in Special Education (PCESE, 2002), too little emphasis was placed on prevention and intervention when a student began to struggle academically. A discrepancy model was not a sound option for America’s youth; a new method was warranted (IDEA, 2004; NCLB, 2001; National Reading Panel, 2000).

A wealth of research, along with common sense, places high merits on early intervention to prevent reading problems and aid in identification of a disability (Coyne, Zipoli, & Ruby, 2006; Gaffney, 1994; Showers, Joyce, Scanlon, & Schnaubelt, 1998; Slavin, 1994). Because as many as 50% of students with a disability fail to graduate from high school, early intervention is

\(^{1}\) For the purpose of this research the current term, Specific Learning Disability was used in all situations including earlier research articles on Learning Disabilities.
essential. Using the discrepancy model, many students were identified with an SLD prior to the exposure to preventative measures (Fuchs, Mock, Morgan, & Young, 2003). The discrepancy model, commonly called the "waiting-to-fail" model, was found to be immersed in challenges. This model hindered early intervention and as a result increased the likelihood of overidentification for special education (Fisher & Frey, 2010). This was counter to the research.

The potential of early intervention to reduce the achievement gap and diminish overidentification caused a shift in the identification process (Bursuck & Blanks, 2010). President George W. Bush signed the reauthorization of the Individuals with Disabilities Education Act (IDEA) on December 3, 2004 (IDEA, 2004). The law changed the SLD identification process from a discrepancy model to a response model, which determined how the child responded to high-quality research-based interventions [34 CFR 300.309] [20 U.S.C. 1121e-3; 1401(30); 1414(b)(6)]. Since approximately 80% of individuals identified with an SLD have a disability that falls under the category of reading, the law had the most impact on reading instruction (Lyon, 1996).

One model that meets the legal criteria for IDEA, but has limited field research, is Response to Intervention (RTI). At the core of the RTI process are three main facets: (a) early detection of children who are struggling, (b) execution of research-based intervention measures at the first sign of difficulty, and (c) quality instruction for all students (Fuchs & Fuchs, 2005). RTI is a tiered instruction model. An intended purpose of the RTI framework is to narrow the achievement gap prior to a child failing and consequently improve the referral process (Bursuck & Blanks, 2010; Fuchs & Fuchs, 2005; Martinez, Nellis, & Prendergast, 2006). Research has focused on the individual components of RTI (e.g. early intervention, quality of instruction), not on the application of the RTI framework. In the Information Age, field studies on the
effectiveness of RTI to improve academic achievement and the identification process for SLD are necessary due to the critical nature of reading.

In theory, the overall RTI process should be effective since the conceptual framework provides a strong underpinning (Heller, Hotlzman, & Messick, 1982; Fuchs, 1995; NASDE, 2006); however, minimal studies have been conducted on the effectiveness, efficiencies, outcomes, and best practices of RTI framework within a real-life field studies. Due to the lack of empirical data, essential guidance has not been provided to practitioners on the following: (a) the most suitable RTI approach, (b) the appropriate intervention strategies, (c) the staff training necessary, or (d) the technically adequate measurement tools to guide decisions regarding responsiveness to the intervention (Bradley, Danielson, & Doolittle, 2005; Fuchs & Fuchs, 2006; Hilton, 2007; Hale, Beringer, Brach, 2013). Denton (2012) reiterated the importance of examining a multitude of questions concerning the outcomes of utilization of RTI in field studies by stating, "Scientific evidence grows incrementally over time, and small numbers of studies rarely result in clear-cut 'answers' that can be easily applied. It is the convergence of findings that best informs practice" (p. 240).

Purpose of the Study

The purpose of this study was to examine the outcomes of the RTI process, such as student achievement in reading and/or placement in special education. In addition, the data provided evidence to the district to guide program decisions. The study investigated whether students identified as lower achieving readers in kindergarten or first grade, who received deliberate intervention in reading within the RTI program, had higher levels of reading achievement by the end of third grade than students who did not receive the RTI reading intervention. The independent variables in this study were inclusion in the intervention or the
non-intervention group, gender, socio-economic status, implementation year, cohort, school, and entry grade. Academic achievement in reading was the dependent variable and was measured by reading assessment results and placement in special education.

**Significance of the Study**

The significance of this research fell into two areas: scholarly research and educational practice and policy. The significance included: (1) additional insight through a field study into the implementation of the RTI framework for both researchers and educational practitioners; (2) potential data for closing the gap, which is of critical concern to educators, necessary for student success, and mandated by NCLB & IDEA; and (3) an evaluation of the RTI implementation for the district. Examining student growth in reading through the study of the implementation of the RTI framework will be beneficial in a number of ways.

In respect to the scholarly research, this quasi-experimental study aids in filling the gap by focusing longitudinally on cohorts of students’ academic achievement in relationship to placement within the RTI tiers. Even though substantial information exists on the components of RTI framework (Flippo, 2001; Speece & Walker, 2007), minimal research is available on the actual implementation of the RTI framework in its entirety within schools (Allington, 2005; Bradley, Danielson, & Doolittle, 2005; Hale et al., 2013; Jimerson, Burns, & VanDerHeyden, 2007). One study documented variability in the implementation of RTI practices across 41 schools (Mellard, McKnight, & Woods, 2009). In this study, variability was found in program equity, efficiency, and the viable nature within the school. Scholars cite research on individual components of the RTI framework, but have suggested additional research be conducted on the implementation of the model to fill the gap between research and practice (Denton, 2012; Fuchs & Fuchs, 2006; Hale et al., 2013; Hill, King, Lemons, & Partane, 2012; Hintze, 2008; Mask &

In times of high-stakes accountability and increased academic demands from society, improved student achievement through program effectiveness is more critical than ever. This research analyzed the influence on student achievement to measure program effectiveness and provided data for potential program improvements to the practitioners. Educational leaders from school districts, state agencies, and federal organizations would be interested parties because of the stakes involved in meeting the federal requirements of ESSA (formerly NCLB) and IDEA. This study provided specific insight into the implementation of the RTI framework regarding long-term reading achievement and special education placement. The intent of this mixed-methods longitudinal study was to add to the academic body of knowledge related to variables leading to quality implementation of the RTI framework.

**Research Questions**

Given this is a mixed-methods study, this research encompassed a primary or central question with sub-questions to guide the exploration of the data collection (Creswell, 2003). This overarching question guided this study:

1. What are the long-range (3 or 4 years) differences on reading achievement for students who participated in RTI Tier 2 reading intervention as compared to students who scored slightly above the intervention group and did not receive treatment?

The following sub-questions guided the research and data analysis for this study:

2. What differences occur among the RTI Tier 2 groups as compared to those students just outside the intervention group in regards to special education identification by the end of third grade?

3.
4. What differences are seen in students' reading achievement among and within the various nested groups (e.g. school, implementation year, achievement measures)?

5. From the perspective of the focus group, what understandings result from the quantitative data analysis of the academic outcomes between and among the nested groups?

**Assumptions**

The following assumptions were cited for the research study:

1. The interventionist (teachers and instructional assistants) participating in the study adhered to the guidelines of the intervention programs. Only when RTI is implemented with high levels of procedural integrity and fidelity can data be used for decision-making regarding student programming.

2. The validity and reliability studies of each assessment applied to this study for benchmark and progress monitoring cut scores held accurate within the school setting.

3. Assessments were consistently and accurately administered.

4. The teachers responsible for teaching the core instruction provided high-quality instruction.

5. The selected sites for the study were not contaminated by the interactions with the researcher.

6. During the focus group discussions, participants answered honestly about the program implementation.

**Delimitations**

Delimitations speak to how the scope of the study was narrowed (Creswell, 1998). The following were the delimitations placed on this research:
1. The research looked at three elementary schools within a school district. The selection of the schools was made to support the research purpose of improving the RTI process within the district.

2. The district studied was located in a rural setting, which may hinder generalization to other settings.

3. Interventions were delivered by trained paraprofessionals.

4. The data included the participants’ scores from 2008-2009 to 2013-2014 school years, which was the length of implementation of RTI process at the school district.

5. The focus group was limited to educators who were school employees from 2008 through 2014 because these participants have a working knowledge of the entire process.

6. Finally, the study focused on the implementation of the RTI framework in reading, as this was the only content area in which the full model was implemented.

**Definition of Terms**

The following definitions provide a common language for the researcher, fellow educators, and readers based on the context of the study:

*Benchmark:* Key end-of-the-year goals for a grade level which are typically assessed three times per year (e.g. fall, winter, spring) so as to measure and monitor progress towards the goal; assessment used to measure a student's progress towards a standard level of achievement (RTI Action Network, 2014).

*Curriculum based measure (CBM):* A standardized, progress monitoring tool that provides equivalent, alternative forms to measure discrete skills. The validity and reliability of CBM are well documented (Deno, 1985; Fuchs & Deno, 1994) and have been prove to be a valid
predictor of identification of students at risk of academic failure (Deno 2003; Good, Kaminski, Simmons, & Kameenui, 2001). In addition, CBMs are less vulnerable to bias associated with sub-groups (Vaughn & Fuchs, 2003).

**Discrepancy Model:** The method used to determine if a student has a learning disability prior to reauthorization of IDEA in 2004. The model is “the discrepancy between a student’s cognitive IQ and achievement test scores” (Greer, 2005, p. 44).

**DIBELS (Dynamic Indicators of Basic Early Literacy Skills):** DIBELS, a form of CBM, is a set of one-minute measures that assess the attainment of early literacy skills from kindergarten through sixth grade. DIBELS is a research-based, statistically reliable and valid assessment that can be used for screening, diagnostic, and progress monitoring essential reading skills (Good & Kaminski, 2013; Kame'enui, 2002). DIBELS measures provide indication of skill level in the areas of phonemic awareness, alphabetic principle, accuracy and fluency with connected text, reading comprehension, and vocabulary.

**Lexile:** “A unit of measurement used when determining the difficulty of text and the reading level of readers. A Lexile is equivalent to 1/1000th of the difference between the comprehensibility of basal primers (the midpoint of first-grade text) and the comprehensibility of an electronic encyclopedia (the midpoint of workplace text)” (MetaMetrics, 2014).

**Nested groups:** Existing groups that have implicit hierarchy, such as assigned school or year of enrollment (Tabachnick & Fidell, 2001).

**Problem-solving protocol:** Similar to a typical problem-solving approach, the problem-solving protocol for RTI identifies the problem, analyzes the options, implements the plan and evaluates the plan. The cyclic model is applied to each student individually, unlike the standard protocol model. It is heavily dependent on the integrity and fidelity of the implementation
Progress monitoring: A short, skill specific repeated measure of performance designed to monitor growth in a specific skill over time. Progress monitoring is based on formative assessment and is a means to improve not only learning, but overall instruction (Batsche et al., 2005; Fuchs, 2003; Fuchs & Fuchs, 2006; Hallahan & Mercer, 2002).

Research-based curricula: Research-based curriculum incorporates general principles of research-based design. In the area of reading, research-based curriculum is composed of “5 Big Ideas”, which include phonemic awareness, phonics, fluency, vocabulary, and reading comprehension (FCCR, 2006; National Reading Panel, 2000).

Response to Intervention (RTI): Response to intervention is a data-driven, systemic approach for identifying and responding to the academic and/or behavioral needs of students, which provides guidance for instruction and intervention thus providing a continuum of services to all students through high-quality instruction (Brown-Chidsey & Steege, 2005; Martinez, Nellis, & Prendergast, 2006).

Specific learning disability (SLD): A Specific learning disability is a presumed neurologically manifested disorder that affects persons of average intelligence ability to either interpret what they see and hear or to link information from different parts of the brain. Due to a disorder, or deficit, in one or more of the basic psychological processes (oral expression, listening comprehension, written expression, basic reading skills, reading comprehension, mathematics calculation, or mathematic reasoning), the person exhibits significant and unexplained under-achievement in a specific area. The deficit is not due to intellectual disability, emotional disturbance, sensory impairment, cultural differences, or a lack of opportunity to learn. SLD may be caused by environmental influences (familial) or family genetics (inheritable). The
person who is identified as having an SLD has both psychological processing strengths and weaknesses (Learning Disabilities Association of America [LDA], 2010; IDEA, 2004 [§300.541(a)]; Horowitz, 2014; Thurlow, Moen, Liu, Scullin, Hausmann, & Shyyan, 2009).

*Standards treatment protocol:* The standards treatment protocol is similar to those implemented in the medical field. The same empirically validated, fixed-duration intervention is used with all students exhibiting the learning deficit. It is a well-specified and often scripted (CRI, 2014; Nelson-Walker et al., 2013; RTI Action Network, 2014).

*Universal screening:* A brief assessment delivered to all students periodically (typically three or four times per year) that demonstrates diagnostic accuracy for predicting which students will develop learning problems. The assessment must also be statistically valid and reliable (RTI Action Network, 2014).

*Reading:* Reading is the act of making sense of print text. Several decades of scientific research indicate the ingredients necessary to teach all children to read accurately, rapidly, and with comprehension by the end of third grade include phonemic awareness, phonics, fluency, vocabulary, and comprehension (Learning Point Associates, 2004; National Institute of Child Health and Human Development [NICHD], 2000).

*Phonemic Awareness:* As the foundation to understanding print, it is the ability to hear the individual sounds and segment the sounds within words (FCCR, 2006; NPR, 2000).

*Phonics:* Phonics is the ability to use the association of sounds with letters to form words. Phonics allows individuals to decode and encode words (FCCR, 2006; NPR, 2000).

*Fluency:* Essentially, fluency is the ability to read written works fluidly and with appropriate pacing for the context (FCCR, 2006). Three measures of fluency include the words-per-minute which a person can read with comprehension, intonation, and voice (FCCR, 2006).
**Vocabulary:** Vocabulary is the understanding of word meaning, which along with background knowledge, allows the reader access to reading materials (FCCR, 2006).

**Comprehension:** Comprehension is one’s understanding and interpretation of the text (FCCR, 2006). It is the ability to express meaning from text) (FCCR, 2006; NPR, 2000).

**Summary**

Chapter 1 examined the outcome of the RTI process on student achievement. Beyond the merits of individual student achievement, one must remember that schools are held accountable for adequate yearly progress under NCLB and most recently the Every Student Succeeds Act (ESSA, 2015), and the skills required for the 21st century are escalating. For these reasons, this research analyzed the RTI method as a means to enhance education for struggling students through student achievement data, RTI tier placement, and architectural components of the program. This research expands the knowledge-base on the effectiveness of the RTI framework.

Reading is essential to success in our society, and closing the gap is of critical concern to educators, necessary for student success, and mandated by ESSA, NCLB & IDEA. Dr. Susan M. Burns, Director of Committee on the Prevention of Reading Difficulties in Young Children National Research Council/National Academy of Sciences stated:

If [a child] doesn’t learn to read well enough to comprehend what he is reading, if he doesn’t learn to read effortlessly enough to render reading pleasurable, if he doesn’t learn to read fluently enough to read broadly and reflectively across all the content areas, his chances for a fulfilling life, by whatever measure – academic success, financial success, the ability to find interesting work, personal autonomy, self-esteem – are practically nil. (Griffin, Burns, & Snow, 1998, p. 233).
Looking ahead, Chapter 2 presents the history behind the surge of RTI implementation across the country and a review of the literature surrounding the underlying conceptual framework and current implementation of RTI. Chapter 3 reveals the mixed-methods design used to carry out the longitudinal and quasi-experimental study. A follow-up explanatory study used focus group data to further the understanding of the data analysis. Chapter 4 expounds on an analysis of the quantitative and qualitative data. Finally, Chapter 5 connects the previous chapters together in a summary that reviews the results, interprets the findings, and provides recommendations for further study.
CHAPTER 2

LITERATURE REVIEW

This chapter presents the foundation for the change to the Response to Intervention (RTI) model for the identification of a Specific Learning Disability (SLD) along with the gaps in said research. The chapter goes on to further explain the RTI framework, including the RTI framework, the conceptual and theoretical foundation, approaches a school may take when implementing RTI, and the critical features of the RTI process.

Purpose of the Study

The purpose of this study was to examine the outcomes of the RTI process, such as student achievement in reading and/or placement in special education. In addition, the data provided evidence to the district to guide program decisions. The study investigated whether students identified as lower achieving readers in kindergarten or first grade, who received deliberate intervention in reading within the RTI program, had higher levels of reading achievement by the end of third grade than students who did not receive the RTI reading intervention. The independent variables in this study were inclusion in the intervention or the non-intervention group, gender, socio-economic status, implementation year, cohort, school, and entry grade. Academic achievement in reading was the dependent variable and was measured by reading assessment results and placement in special education.
Research Questions

Given this is a mixed-methods study, this research encompassed a primary or central question with sub-questions to guide the exploration of the data collection (Creswell, 2003).

This overarching question guided this study:

1. What are the long-range (3 or 4 years) differences on reading achievement for students who participated in RTI Tier 2 reading intervention as compared to students who scored slightly above the intervention group and did not receive treatment?

The following sub-questions guided the research and data analysis for this study:

2. What differences occur among the RTI Tier 2 groups as compared to those students just outside the intervention group in regards to special education identification by the end of third grade?

3. What differences are seen in students' reading achievement among and within the various nested groups (e.g. school, implementation year, achievement measures)?

4. From the perspective of the focus group, what understandings result from the quantitative data analysis of the academic outcomes between and among the nested groups?

Specific Learning Disability

A person who has a Specific Learning Disability (SLD), initially called Learning Disability,\(^2\) exhibits a disorder or deficit in one or more of the basic psychological processes (LDA, 2010). The term was first introduced by Kirk (1962) to characterize children who have delayed development or disorder in speech, language, reading, writing, or mathematics. In the 1970’s Congress included SLD in the Education for All Handicapped Children Act (PL 94-142)

\(^{2}\) For the purpose of this research the current term, Specific Learning Disability was used in all situations including earlier research articles on Learning Disabilities.
as a unique group of children who demonstrated unexpected learning difficulties (Fuch, Mock, Morgan, & Young, 2003). These processes are defined by Individuals with Disabilities Education Improvement Act of 2004 (IDEA) (2004) to include oral expression, listening comprehension, written expression, basic reading skills, reading comprehension, mathematics calculation, or mathematic reasoning [§300.541(a)]. SLD is an umbrella term for a range of presumed neurologically manifested conditions. SLD is a chronic condition with the majority of students (70 - 80%) identified with impaired reading (Fletch, Lyon, Fuchs, & Barnes, 2007; Foorman, Fletcher, & Francis, 1997; Hadler & Fierson, 2011). In the area of language arts, an SLD may selectively hinder the demonstration of verbal and/or nonverbal abilities (Thurlow, Moen, Liu, Scullin, Hausmann, & Shyyan, 2009).

The U.S. Department of Education (1977, p. G1082), defined SLD as "a severe discrepancy between achievement and intellectual ability." This means a person who is identified as having a learning disability has both psychological processing strengths and weaknesses as identified by a multidisciplinary team. Low achievement alone is not a suitable diagnostic indicator (LDA, 2010). For classroom purposes, this would mean a child who exhibits low achievement or who meets minimal academic standards may or may not be identified as having SLD. While the method of identification has been altered, a survey by Learning Disabilities Association of America (2010) concluded that the preponderance of the research community believed the definition of SLD should remain the same; the challenge did not fall to the definition, but rather the identification process.

Accuracy and timeliness of support were common SLD identification stumbling blocks. The number of children identified had escalated since the establishment of an SLD as a disability, particularly in some minority populations (IRA, 2007; Walmsely & Allington, 2007).
In addition to the escalation in identification, misidentification, and missed identification have plagued student identification for a Specific Learning Disability (Vaughn, Linan-Thompson, & Hickman, 2003). The Office of Special Education Programs (OSEP) conducted a multi-year investigation into the need for early and accurate identification of children with disabilities called the LD Initiative. This investigation found (a) strong converging evidence supporting the concept of SLD, (b) the ability-achievement discrepancy model insufficient for accurate identification, (c) the Intelligence Quotient (I.Q.) test to be not necessary for most identifications, and (d) early intervention with quality research-based instruction to be the most promising method for accurate and timely identification (Bradley, 2005). Considering the issues listed above and findings from studies such as LD Initiative, IDEA (2004) removed the discrepancy model as the primary model for initial determination of a Specific Learning Disability.

**Discrepancy Model Issues**

Paramount to the identification issues was the ability-achievement discrepancy model. The discrepancy model was the most common process used in the identification of an SLD. The model was used despite the fact that the definition of SLD does not require a severe discrepancy between achievement and ability, numerous measurement problems existed, theoretical components were counter to the research, and the discrepancy score had been found to be unreliable (Fletcher, Lyon, Barnes, Stuebing, Francis, Olson, Shaywitz & Shaywitz, 2002; Kavale, 2002; Reynolds, 1984).

The discrepancy model is based on a difference between a student’s ability, typically measured by I.Q., and the student’s academic achievement with no regard to the quality of the instruction or a requirement for early intervention to support the student (Fuchs & Fuchs; 2006; Mesmer & Mesmer, 2008; Vaughn & Fuchs, 2003; Fuchs, Mock, Morgan & Young, 2003).
According to the IRIS Center at Vanderbilt University (The IRIS Center for Training Enhancements, 2006), this aspect of the model is based on a normal curve of expectations. Typically, the model uses a criterion of at least two standard deviations, or at least a 30 point difference, between a student’s outcome on an intelligence test such as the Wechsler Intelligence Scale for Children (WISC-IV) and academic achievement scores on an assessment such as the Woodcock-Johnson Achievement Test. In the case of a substantial difference, a student may be identified as having a specific learning disability. The discrepancy model has several flaws mentioned previously, but its strengths include the wealth of research behind the typical assessments used to measure ability and the relative ease of implementation requiring minimal teacher time (CEC, 2008).

Atheoretical

The discrepancy model has been criticized as atheoretical (Fuchs & Fuchs, 2006; Lyon, 1987; Willson, 1987) because it was rooted in the idea that the difference between two variables could separate the presence or absence of an SLD amongst the array of uncontrolled variables (i.e. teaching quality, instruction quality). The lack of a foundation for the discrepancy model caused identification methods to grow in diversity across the nation. Variations were prevalent in (a) the methods of computation, which included, but were not limited to standard scores minus achievement data, regression calculation, and simple noncompliance by some school-level personnel to any form of computational method, (b) the standard deviation required for SLD identification, (c) the tests used to determine ability and achievement, and (d) the sophistication of school psychologists’ abilities to correctly evaluate ability-achievement difference scores (Fuchs & Fuchs, 2006; Kavale, 2002).
A challenge built within the definition of SLD is that it is primarily exclusionary by nature (Kavale, 2002). The exclusionary nature of the SLD definition along with the lack of procedural guidelines for identification permitted individual states, districts, schools, and even individual teachers to be free to choose their methodology. This led to widespread arbitrariness in SLD identification (Kavale, 2002; Vaughn & Fuchs, 2003), or as Fuchs & Fuchs (2006) stated an SLD identification was “whatever teachers and parents want it to be.” The arbitrary identification and structures of the discrepancy model created a “wait-to-fail” model that did not consider the quality of instruction (Fuchs & Fuchs, 2006; Fuchs, Mock, Morgan, & Young, 2003; Mesmer & Mesmer, 2008; Vaughn & Fuchs, 2003).

“Wait-to-fail” Model

The discrepancy model necessitated a statistically significant gap between expected and actual performance causing a “wait-to-fail” model (Berkley, et al., 2009; Fletcher, Coulter, Reichly, & Vaugh, 2004; Meyer, 2000). Under the discrepancy model, the critics would say students were not supported in a timely manner. Educators and families were frustrated as students did not receive support services, and the delays were potentially harmful (Fletcher et al., 2004). Studies have found that many students did not receive support until third or fourth grade when the difference between achievement and ability demonstrated a sufficient gap (Fuchs, Mock, Morgan, & Young, 2003; Lyon, et al., 2001; Velluntino, Scanlon, & Lyon, 2000). Years of programming under the discrepancy model demonstrated that the “gift of time” for the issues to resolve themselves was not beneficial to student academic growth, self image, or interest in learning (Fisher, 2010; Fletcher et al. 2004). IDEA references response to intervention as a method to ensure that students receive early intervention when struggling academically, rather than utilizing a “wait-to-fail” structure.
Quality of Instruction

The discrepancy model does not consider or evaluate the quality of the instructional delivery (Fuchs & Fuchs, 2006; Fuchs, Mock, Morgan, & Young, 2003) and for this reason questions arose as to the number of false positives yielded from poor teaching resulting in low achievement (Fletcher et al., 2004; Fuchs & Fuchs, 2006). Studies reveal that many students’ struggles were based in the lack of quality instruction and that the impact of the quality of instruction is often underestimated (Donovan & Cross, 2002; Foorman, Fletcher, & Francis, 1997; Fuchs & Fuchs, 2006; Fuchs et al., 2003). The lack of consideration for the quality of instruction caused misidentification of students with SLD.

Disproportionality

In addition to the atheoretical nature of the discrepancy model, the identification data following years of implementation of the ability-achievement model has been shown to be disproportional in SLD identification among subgroups (Coffey & Obringer, 2000; Fuchs & Fuchs, 2006). Reschly (2002) compiled evidence from the U.S. Department of Education data demonstrating that other areas of special education did not have a disproportionality of subgroups as compared to the general population. Along with other researchers, Fuchs & Fuchs (2006) reached a consensus that the inconsistency in the definition of the ability-achievement method has unfairly withheld identification and over identified other groups of students. Studies show that race, gender, and socio-economic status impact SLD identification (Coffey & Obringer, 2000; Bocian, Beebe, MacMillan, & Gresham, 1999). As indicated in Table 1, 36% of all students served in 2011-2012 under IDEA, Part B were identified under the SLD identification category. According to national data, Hispanic, African-American, and American Indian/Alaska Native students are overrepresented; white students are slightly underrepresented;
and Asian students are substantially underrepresented.

Table 1

<table>
<thead>
<tr>
<th>Race/Ethnicity</th>
<th>Percentage of the Special Education Student Served for SLD</th>
<th>Percentage of Total American Public School Student Population</th>
<th>Percentage of SLD Served</th>
</tr>
</thead>
<tbody>
<tr>
<td>All races</td>
<td>36.0%</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>White</td>
<td>32.5%</td>
<td>51.7%</td>
<td>48.6%</td>
</tr>
<tr>
<td>Black</td>
<td>38.8%</td>
<td>15.8%</td>
<td>20.2%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>43.6%</td>
<td>23.7%</td>
<td>25.7%</td>
</tr>
<tr>
<td>Asian</td>
<td>23.1%</td>
<td>4.3%</td>
<td>1.5%</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>41.7%</td>
<td>1.1%</td>
<td>1.6%</td>
</tr>
<tr>
<td>Two or more races</td>
<td>30.7%</td>
<td>2.6%</td>
<td>2.1%</td>
</tr>
</tbody>
</table>

*Note: These data included children age 3 to 21 years old served under IDEA, Part B in 2011-2012. From National Center for Education Statistics. (2013). Digest of Education Statistics, Table 204-50.*

In regards to gender discrepancies, males are overrepresented (66% of the SLD population, while only 51% of the student population); females are underrepresented (34% of SLD population, while 49% of student population) (Cortiella & Horowitz, 2014; Reschly, 2002). Research indicates an equal number of boys and girls share the common characteristics of SLD (Cortiella & Horowitz, 2014; Reschly, 2002) yielding a discrepancy of 15% for males and females. As noted in the NCLD report (2014), the 2010 U.S Census Bureau survey revealed the frequency of 2.8% SLD identification for school age children (6-17 years old) living in poverty; whereas, those living above the poverty line yielded a 2.0% SLD identification. It is noteworthy that children living in poverty are at greater risk of poor health when considering the connection between socio-economic status and disabilities (Cortiella & Horowitz, 2014). Likewise, students living in foster care or who are homeless comprise another underrepresented population, most likely due to their mobility and limited capacity for family support in the educational system (Cortiella & Horowitz, 2014; Reschly, 2002). Finally, the discrepancy model is insensitive to students who are in the upper and lower ranges of I.Q. (Stuebing, Fletcher, LeDoux, Lyon, Shaywitz & Shaywitz, 2002). Students with I.Q.s in the upper range tend to be over-identified using the discrepancy model and those in the lower range are under-identified. Overall the
discrepancy model has been shown to not represent the various sub-groups equitably.

**Special Education Cost**

Beyond the disproportionality of subgroups identified and atheoretical nature using the ability-achievement discrepancy model, economics played a role in the decision to move to RTI as a decision-making model. A student with a disability costs on average two to three times more than a student without a disability (Fuchs & Fuchs, 2006). Using the discrepancy model under the guidance of the Education of All Handicapped Children Act of 1975, the population of children identified as having an SLD rose from less than 2% to over 6% (Fuchs & Fuchs, 2006). The conceptual framework of the RTI framework had been recognized as having the potential to intervene and therefore reduce the number of students identified as having an SLD. This reduction in identification would, of course, also reduce the cost of special education services. This potential was recognized when from 2004 to 2011, the percentage of American public school students identified as having a specific learning disability has decreased from 5.7% to 4.8% (or 2.8 million to 2.4 million) (Cortiella & Horowitz, 2014; U.S. Department of Education, 2013). This represents a decline of 0.9% or 400,000 students and a substantial cost savings.

The ability-achievement discrepancy, or “wait-to-fail”, model requires time lags to demonstrate significant gaps and therefore does not intervene in a timely manner; nor does it consider the quality of the instruction. It has been demonstrated that the discrepancy model does not serve students well. The weaknesses of the discrepancy model contributed to two federal acts (No Child Left Behind Act of 2001 and IDEA) that support intervention models that improve instructional support for struggling students. Response to Intervention (RTI) promotes the use of evidence-based instruction and early intervention and has emerged as an alternative approach to identification of a Specific Learning Disability.
Shift in the Identification Process

A shift in the identification process for SLD was warranted, due to the discrepancy models:

- the negative impact of a “wait-to-fail” model,
- over-identification of some subgroups,
- lack of regard for the quality of the instruction,
- the absence of a theoretical foundation,
- insensitivity of developmental differences in cognition and achievement,
- difficulty distinguishing between children with SLD and low achievers, and
- measurement problems including unclear direction as to which IQ measure or score should be used (Aaron, 1997; Berninger-Abbott, 1994; Bocian, et al., 1999; Fletcher et al., 1994, 2002; Foorman, Francis, & Fletcher, 1995; Fuchs & Fuchs, 2003; Fuchs & Fuchs, 2006; Fuchs, Fuchs, Mathes, Lipsey & Roberts, 2002; Gudnerson & Siegal, 2001; LDA, 2010; Kavale, 2002; Peterson & Shinn, 2002; VanDerHeyden et al., 2007; Stanovick & Siegel, 1994; Vaugh et al., 2003).

In a study by MacMillan, Gresham, and Bocian (1998), the results of 61 SLD placements by school districts indicated that when applying the research diagnostic criteria to the students identified with an SLD, only 29 of the students met the research diagnostic criteria for SLD classification (less than 47.5%). This was interpreted as a disregard for the “exclusionary criteria” specified in the authoritative definition of SLD. It had become the sociologic sponge of education (Lyon, 1987). The SLD identification process had become unfair, inconsistent in its logic, and in some situations mystifying. The shift from the ability-achievement discrepancy method to RTI framework for Specific Learning Disabilities arose due to these criticisms (Fuchs
& Fuchs, 2006). RTI was developed in response to the flaws, primarily the lack of attention to quality instruction and the absence of early interventions, in identification of students with SLD in the discrepancy method (Francis et al., 2005; O’Malley, Francis, Foorman, Fletcher, & Swanks, 2002; Stanovick 2005; Mesmer & Mesmer, 2008). RTI addresses the learning opportunities provided at all levels of instruction and provides early intervention at the first sign of a struggle (Lyon et al, 2001; Martinez, Nellis, & Prendergast, 2006; Vaughn & Fuchs, 2003), unlike its forerunner the ability-achievement discrepancy model.

**Response to Intervention (RTI)**

NCLB established an accountability requirement that public schools provide more choices for parents, increase local control and flexibility, and emphasize scientifically-based instructional strategies to meet Adequate Yearly Progress (AYP) for the general student body along with disaggregated subgroups (e.g. special education). Identification of specific learning disabilities was altered by the IDEA. This reauthorization did not alter the definition of SLD, but instead changed the procedures by which local education agencies would identify students with a specific learning disability. The legislation allowed states to use evidence-based alternative methods of identification as they saw fit for students of their state. RTI is a popular, widely accepted method that meets these legal requirements (CEC, 2008; Fuchs, 2006; IDEA, 2004; Shapiro, n.d.; Vaughn, Wanzek, Woodruff, & Linan-Thompson, 2007). The ability-achievement method is a "waiting-to-fail" model, while the RTI framework is a prevention/intervention model.

RTI has no defined format, but by design incorporates assessment and intervention at various stages of instructional intensity within the multi-tier system (Shapiro et al., 2011). The key elements of the RTI framework include: (a) periodic use of a universal screener to aid in

...
sorting students and identifying areas of need, (b) systematic use of progress monitoring to guide instruction at each level, (c) focus on research-based instruction at the core level and all intervention levels, and (d) at the heart of RTI, early intervention when progress is not being attained (Greenwood & Kim, 2012). The principles of early intervention, prevention, and a rigorous problem-solving model uphold the ideals and requirements of ESSA, NCLB, and IDEA.

RTI is described as a method of determining if a child has a learning disability based on a lack of response to a research-based intervention (OSEP, 2004: Section 614(b) (6)). Response is derived from performance compared to peers, a normative cut score on an assessment, or a combination of both. Theoretically, a lack of responsiveness relative to one’s peers may denote a learning disability. In contrast to the discrepancy model, the RTI framework provides layers of intervention, considers the quality of the learning opportunity in all layers, and evaluates the child's pattern of strengths and weaknesses prior to special education referral (U.S. Department of Education, 2006).

The goal of RTI is to support students' achievement. The underlying assumption behind RTI is that all students can learn and reach high levels of achievement if the system provides appropriate instruction and allows the student the “gift of time.” RTI builds on Tomlinson’s work in differentiated instruction (DI) and Wiggins and McTighe’s work with understanding by design (UbD) (Fisher, 2010). Research shows that early intervention provides better support, thus preventing the development of significant academic deficiencies and long-term academic gains to the struggling reader (Flippo, 2001; Kovaleski, Prasse, 2004; Lopez, 2009; Pikulski, 1994; Vaugh et al., 2007; Ziolkowska, 2007). Intervention, as well as core instruction, should address a range of critical reading elements (i.e. phonological awareness, phonics, word/vocabulary development, listening/reading comprehension, and reading fluency) to meet
the student's individual needs (Alanis, Munter, & Villamil; Florida Center for Reading Research, 2006). By differentiating for a struggling student through increased levels of early interventions, the RTI framework provides the scaffolding and additional time necessary for student success.

Hintze's work (2008) chronicled the benefits of using the RTI framework for the identification of SLD. He also posed a series of questions that remained unanswered. His work is supported by a range of researchers (Brown-Chidsey & Steege, 2005; Danielson & Doolittle, 2007; Denton, 2012; Flippo, 2001; Fuchs, 2003; Fuchs & Fuchs, 2006; Glover, DiPerna, & Vaughn, 2007; Gresham, 2002; Hale et al., 2013; Kratchwill, Clements, & Kalymon, 2007; Mask & McGill, 2010; Shapiro, Zigmond, Wallace, & Marston, 2011; Slavin, 1994; Speece & Walker, 2007; Vaughn, Wanzek, Woodruff, & Linan-Thompson, 2007; Wanzek & Vaughn, 2007). The first, and extremely advantageous, benefit is early identification of learning struggles. As early as the fall of kindergarten, students are identified as non-responders to core instruction through benchmark assessments and receive targeted instructional support in Tier 2. This factor greatly decreases the probability of a student slipping through the cracks in the system. However, unanswered empirical questions remain in the areas of timing, accuracy, and the relationship between the screener and the core instruction. Examples might include:

(1) How accurate are screeners when kindergarteners developmentally change so rapidly?

(2) To what extent do the core curriculum and the screening tool yield false positives?

(3) How does the progress of lower achieving students with weak academic foundation compare with or without intervention?

(4) To what extent does the delivery of an intense, explicit intervention alter the likelihood of a student exhibiting a disability?
The second benefit of identification through the RTI framework is the potential for a reduction in the screening bias of teacher-based referrals in the discrepancy model (Vaughn & Fuchs, 2003). Reduction in bias through the systemic approach built into the RTI framework has the means to address the misidentification, under identification, and over identification of various subgroups of the student population. Solid data do not exist at this time to emphatically guide practitioners in the determination of disability status, evidence needed to designate an underachieving reader, and appropriate rates of improvement. These questions are left to the LEA to ascertain based on program specific available research (i.e. DIBELS, AIMS Web), not the RTI framework as a whole.

The third benefit of the RTI framework is the relationship among the identification process, assessment, and instruction. RTI aligns the identification process with the instruction through the use of the assessment. Students’ reaction to the instruction at both an individual and group level should guide decisions, instead of autonomous processing skill assessments previously used in the discrepancy model. While promising, a key question yet to be dissected is do the schools have the resources, skills, and knowledge to adequately assess student progress?

Lastly, a positive feature of the RTI framework is that is rests on the logic of hypothesis testing (Hintze, 2008). RTI addresses a problem by proposing a treatment, testing the treatment, and monitoring the outcome. If the impact does not occur as desired after several treatment attempts, the student would likely be identified as having an SLD.

These four benefits promote the promise of the RTI framework for identification of SLD and improvement of overall instruction. However, many questions about the RTI framework exist and merit continued research to further operationalize procedures, establish a sound basis for making decisions, and determine SLD identification criteria.
Principles of RTI

In the National Association of State Directors of Special Education report, *Response to intervention: Policy considerations and implementation*, RTI was defined as "the practice of providing high-quality instruction and intervention matched to student needs and using learning rate over time and level of performance to make important educational decisions" (Batsche et al., p.1). It listed four underlying principles that guide the implementation of RTI. The principles are (a) all children can learn if effectively instructed, (b) core instruction and intervention programs should be research-based, (c) early intervention is critical to improved achievement for students, and (d) student data should drive decision-making.

**All children can learn.** To help all children learn, Leverett (2006) stated that a variety of strategies were essential: (a) varied instructional approaches, (b) early childhood education, (c) family involvement, and (d) student engagement. First to deepen understanding, instruction should be differentiated to match students’ learning styles or intelligences; needs to be leveled to provide a range of support; and must include rich, vivid resources to build background knowledge (Leverette, 2006). All people possess all intelligences, but each person has dominant styles (Gardner, 1983). Traditional instruction tends to lean on two core intelligences: linguistic and logical-mathematical intelligences. While some research concludes that schools should focus on the core intelligences (Armstrong, 2013); more research has demonstrated that teaching in multiple ways has been shown to improve learning (Guignon, 2010; Hattie, 2011). Each person learns at different rates, therefore, another means of differentiating may be additional time provided through leveled instruction. Finally, background knowledge has been found to be a compelling factor in student achievement (e.g. correlation of .66 between a person's background knowledge and the extent to which they learn new information) (Marzano, 2004); one that must be considered. In brief, varying strategies and time allotment along with building background
knowledge are necessary to help all children learn.

Beyond differentiation, a high-quality, early childhood education program (i.e. preschool and kindergarten) is vital to develop the early literacy and math skills that provide the foundation for learning (Leverett, 2006). Student's families must be consistently exposed to first-rate learning resources and family events. As well as, students must be engaged through verbal, written, or kinesthetic interaction (Koutsoftas, Harmon, & Gray, 2009). Life experiences, intelligences, learning style, and engagement all affect a student's achievement. RTI embraces these key ingredients through high-quality, leveled instruction.

**Research-based Instruction.** One of the key reasons IDEA (2004) modified the identification process is that an effective core instruction provides the foundation for all learning. As the foundation, the core instruction is critical in special education placement decisions. Measuring a student’s response to instruction can only be deemed reliable if the core instruction is presumed adequate (Stecher, Fuchs, & Fuchs, 2005).

The quality of the core instruction should be measured. Many researchers contend that if the Tier 1 core instruction is not meeting the needs of 80-85% of students, the core instruction should be addressed (Fuchs & Fuchs, 2006; NASDSE, 2006; Shapiro, 2011). The core instruction should be scientifically valid and delivered with fidelity to uphold the integrity of the program (Lyon et al., 2001; Vaughn & Fuchs, 2003).

Reading is “an active and complex process that involves: understanding written text; developing and interpreting meaning; and using meaning as appropriate to the type of text, purpose, and situation” (National Assessment of Educational Progress, 2010, p iv). In the late 1990s, Congress appointed the National Reading Panel to evaluate the status of research-based knowledge related to approaches that teach children to read. The panel reviewed over 100,000
studies on reading dating from 1966 to 2000 and determined that scientifically-based reading instruction is composed of “5 Big Ideas.” The "5 Big Ideas" include phonemic awareness (the ability to hear the individual sounds and segment the sounds within words), phonics (the ability to use the association of sounds with letters to form words), fluency (the ability to read written works fluidly and with appropriate pacing for the context), vocabulary (the understanding of word meaning), and reading comprehension (the ability to express meaning from text) (FCCR, 2006; NPR, 2000). Phonemic awareness and phonics form the foundation for which vocabulary can be developed. Fluency and comprehension are intertwined as comprehension is dependent upon fluency. A gap in one area poses challenges to one's ability to read for meaning. "Strong reading comprehension skills are central not only to academic and professional success, but also to a productive social and civic life" (Shanahan et al., 2010, p. 5). To provide a comprehensive program, an effective reading curriculum weaves the “5 Big Ideas” together (Shapiro, Zigmond, Wallace, & Marston, 2011).

**Early intervention.** Research indicates that early intervention increases the achievement rate and provides a stronger likelihood of academic success (Flippo, 2001; Lopez, 2009; Vaugh et al., 2007). Diamond and Mandel (1996) found that 75% of students who had a delay in intervention continued to have reading difficulties. On the other hand, 20 minutes of direct instruction three or four times per week would increase a student’s phonemic awareness skills and improve a student’s achievement in reading. To achieve results, the intervention should be explicit, intense, and monitored (Fuchs & Fuchs, 2006; Gersten, et al., 2009; Gunn, Biglan, Smolkowski, & Ary, 2000; NASDSE, 2006; Vaughn et al., 2006). The intervention services may need to increase in intensity by adding time, creating smaller groups, or altering the intervention program. Any of these provided more individualization through active learning,
more engagement, and rigor (Fuchs & Fuchs, 2006; Lyon, et al., 2001; Sharpiro, 2013; Strangeman, Hitchcock, Hall, & Meo, 2006). Student progress should be monitored closely to make appropriate adjustments. The key to RTI is the use of student data to guide decisions. Through the use of benchmark and progress monitoring data student placement and program adjustments are made (Fuchs & Fuchs, 2006; Fuchs & Vaughn, 2005; Hale et al, 2013).

**Data-based, decision-making.** RTI is a data-based, decision-making process. Allington (2005) asserted that more than the reading program itself, the teacher’s expertise in reading development and the teacher’s professional accountability are key factors in student achievement. Mesner and Griffith (2005) stated that assessment data allowed teachers to shape reading instruction to meet the unique needs of children and therefore leads to academic achievement. Expertise and accountability may be guided by data. In a data-based model, information is gathered, reviewed, and analyzed in order to produce useful knowledge (NCREL, 2004). The student’s assessment data are represented graphically to provide evidence of progress towards a goal. The baseline assessment and the end goal establish the aim line for each student. The data provide a visual representation of the strengths and weaknesses in the targeted area (NCREL, 2006). Through triangulation of the data (e.g. assessment, demographics, climate), hypotheses regarding instruction are developed. The focus of the hypotheses leads to realistic, effective goals. Data-based, decision-making has been shown to create effective and efficient school improvement processes.

These principles guide the RTI process, but they do not define the architectural landscape required for full implementation of the RTI program. Currently, there are not specific steps or defined guidelines for the RTI process. There is, however, a general consensus among researchers and practitioners as to the components required within the RTI framework (Jimerson,
Burns, & VanDerHeyden, 2007). The components of the RTI framework include a universal screener for all students, research-based tiered instruction and intervention, progress monitoring the students' responsiveness to the instruction, and a team decision-making process (Elliot & Morrison, 2008; Fuchs, 2006; Jimerson et al., 2007; Linin-Thompson et al., 2006; Lyons et al., 2001; Stanovick, 2005; Shapiro, 2011).

The research base for reading instruction has a strong focus on the use of data to prevent reading difficulties by providing intervention (Denton, 2012; National Reading Panel, 2000). Previous research indicates that the intervention should provide (a) explicit instruction, (b) corrective and positive feedback, (c) copious amounts of engaging practice in reading and responding to text, (d) and active student involvement (Denton, 2012; Gersten et al., 2009; National Reading Panel, 2000). One longitudinal study examined the relative effectiveness of intervention during kindergarten and first grade on reading achievement by third grade (O’Conner, Harty, & Fulmer, 2005). All students \( n = 31 \) received intervention. Students who continued to struggle received two layers of intervention. The study's findings revealed that all students who had received Tier 2 intervention and 40% of the students who received both Tier 2 and Tier 3 intervention performed on average in the third grade on measures of word reading and oral reading fluency. In sum, the intervention-initiative substantially reduced the number of children who performed below the desired standard. Other studies show similar benefits of intervention (Fuchs & Fuchs, 2006; Nelson-Walker, et al., 2013). Kerins, Trotter, and Schoenbrodt (2010) compared two groups (i.e. \( n = 9 \) no intervention and \( n = 11 \) additional 16.5 hours of intense phonological awareness and phonics instruction), but upon initial analysis did not notice discernible noteworthy differences in the students' skills. After further investigation, they recognized the core program emphasized phonemic awareness and phonics (i.e. National
Reading Panel recommendation). The findings affirm the panel's stance on the critical nature of
the two skills; however, it also emphasized the need to use data to determine individual student's
needs and align the program to those needs. Based on the work of Hill et al. (2012), these
studies bring to light the need for alignment of the instructional levels, as well as, quality core
instruction. The researchers suggested more studies be developed which analyze empirical data
for a range of parameters in various settings (Fuchs & Fuchs, 2006; Kerins, et al., 2010; Nelson-

**RTI Conceptual Framework**

RTI is not a specific method, but rather a framework that aims to shift educational
resources toward prevention and intervention to meet the individual academic needs of children.
The conceptual framework of RTI is grounded in a cyclic decision-making approach (Figure 1).
Instructional decisions (Figure ,1 step 2) are delivered through research-based instruction at all
levels (Figure 1, step 3). These decisions are guided by curriculum-aligned assessments (Figure 1
steps 1 and 4) (Fuchs & Fuchs, 2006; Heller, Holtzman, & Messick, 1982; Linan-Thompson,
Vaughn, Prater, & Cirino, 2006; Reschly, 2002).

![Figure 1. RTI Data Decision Cycle. From Marion County Public School, Ocala, Florida. Reprinted with permission.](image)
The above figure shows the fluid natures of instruction, whereas Figure 2 demonstrates the inclusive nature of all students in Tier 1 core instruction. In addition, this figure demonstrates that the foundation of the instructional model is the use of universal assessments.

Figure 2. RTI Data-Based Decision Making Model. Adapted from "RTI Implementer Series: Module 3: Multi-Level Prevention System—Training Manual," 2012, National Center on Response to Intervention.

These decisions are anchored in the analysis of individual and group data (Figure 3) to assure quality instruction is the foundation for all students (i.e. Tier 1) and appropriate instructional decisions for each child during interventions (i.e. Tier 2 and 3).

Figure 3. Instructional Data-driven Decision Making Process for RTI. Adapted from RTI Action Network, 2014, National Center for Learning Disabilities.
Integrating the commonalities across the three figures, it is evident that instructional decisions occur at all levels of RTI implementation and all levels of instruction. Educators use screening and progress monitoring data to make decisions about instruction, movement of students within the multi-level prevention system, and disability identification.

Each conceptual component of the RTI framework has a dearth of research substantiating its effectiveness, but as a unit RTI has limited empirical data (Bradley, Danielson, & Doolittle, 2005; Fuchs & Fuchs, 2006; Hale, et al., 2013; Ridgeway, Price, Simpson, & Rose, 2011; Speece & Walker, 2007; Sugai & Horner, 2009). The National Center on Response to Intervention (NCRTI) maintains there are four conceptual components to the RTI framework: (a) school-wide, multi-level instructional system for the prevention of academic failure through early intervention; (b) universal screening to begin the process of identification of low achieving students; (c) progress monitoring throughout the levels of instruction; and (d) data-based decision making for instruction, intervention, and potentially the identification of a learning disability (2010). The overarching goal of RTI is to maximize student achievement.

**School-wide Implementation**

School-wide implementation of the RTI framework (NCRTI, 2010) is imperative given the evidence that organizational structures are required for effective outcomes (O'Connor, Harty, & Fulmer, 2005; Fuchs & Fuchs, 2006; Vaughn et al., 2009). RTI is a proactive approach to aid in the prevention of learning difficulties (Kratchwill, Clements, & Kalymon, 2007). Achieving a proactive stance requires schools to alter some of the structures, staffing, and management (Greenwood et al., 2008) to allow students to flow within and among the tiers of service as necessary. Studies on the impact of teachers, leadership, and a professional learning environment indicate the need for a school-wide implementation of the RTI framework to
achieve the desired outcomes.

Research indicates that when facilitated by a highly-qualified teacher, the core instruction that is built on research-based curricula provides the majority of students with a successful learning opportunity (Speece & Case, 2001; Vaughn, Linan-Thompson, & Hickman-Davis, 2003). Research confirms that the percent of well-qualified teachers in a school is the most significant predictor of student achievement in reading (r between .61 and .80, p<.001) (Darling-Hammond, 1999). Teacher quality is significantly and positively correlated with student outcomes.

Beyond high quality teachers, leadership is integral in the execution of a school-wide RTI framework due to the need for strong support (e.g. time, fiscal resources, training), a clearly articulated vision, and the development of a purposeful community. With dynamic leadership, school communities are more willing to take risks, feel empowered to make decisions, and provide valuable feedback (Fullan, 2001; Mazano, Waters, & McNulty, 2005; Silins, Mulford, & Zarins, 2002). Strong collaborative district and building leadership have been shown to be a vital asset not only in the initial implementation, but also in the prospect that any new instructional practice will be implemented fully and maintain integrity over time (Mazano, Waters, & McNulty, 2005; NASDSE, 2006). A meta-analysis found the correlation between the leadership of the principal and the academic achievement of the students to be .25, second only to teachers (Leithwood, Louis, Anderson, & Wahlstrom, 2004; Mazano, Waters, & McNulty, 2005). District, school, and classroom leadership are necessary to fully implement a change with the magnitude of a school-wide, multi-level RTI framework.
In addition, professional development for school leaders, teachers, and paraprofessionals was crucial to implementation (Bailey, 2010; Hauerwas & Gressling, 2008; Nicholson, 2009). Teachers initially need professional development in the areas of decision-making process, analysis of data, and RTI process (Nelson-Walker et al., 2013, Tileston, 2009). In addition, ongoing professional development to improve the delivery of the core instruction is imperative. One study showed a large significant effect size based on differences between means (Hedge's $g^3 = 1.31$) in the overall quality of instruction for the treatment group following intense professional development (Nelson-Walker et al., 2013). In addition, the characteristics of a professional learning environment (e.g. communication, collaboration, and data-driven decision-making) were shown to be essential. Fuchs & Fuchs (2006) also emphasized the district role in providing the resources (e.g. human capacity and materials).

**Multi-tiered Instructional System**

Fuchs and Fuchs (1998) synthesized the Heller et al. (1982) framework into the RTI framework. In the RTI framework, students move to progressively higher tiers of intervention, if they are not successful in the research-based core instruction, until the student achieves sufficient progress toward a predetermined goal. The research does not favor a specific number of tiers, but three or four instructional levels are most common (Berkely et al., 2009; Fuchs 2006; Bradley, Danielson, & Doolittle, 2005; Tilly, 2008). The conceptual framework for RTI may be illustrated by a pyramid of instructional opportunities (Figure 4).

---

3 Hedge's $g$ is a measure of effect size. It is similar to Cohen’s $d$. Both of these measures consist of the difference between means divided by the standard deviation. However, Hedge’s $g$ uses the version of the standard deviation formula that divides by $N-1$. Hedge’s $g$, like Cohen’s $d$, are scale free. Specific guidelines for a large effect size do not exist, but Cohen suggests that an effect size of 0.8 is a large effect has been adopted (Lane, 2014).
The National Reading Panel cited an array of research studies to demonstrate the value of early intervention in reading and to ensure high-quality reading instruction (Snow, Burns, & Griffin, 1998). A second phase of the National Reading Panel analyzed the research on types of instruction necessary to learn to read. The panel’s analysis of the research confirmed that systematic reading instruction should include the following elements: phonemic awareness, phonics, fluency, vocabulary, and comprehension. The RTI framework focuses on both systematic reading instruction and intervention through a multi-tiered approach.

**Tier 1: Core instruction.** Tier 1 core classroom instruction provides a quality, standards-aligned, research-based curricula delivered by a highly qualified teacher (Fuchs & Fuchs, 2005; Gersten & Dimino, 2012). This core instruction is delivered to all students and provides a foundation for learning (Denton, Vaughn, & Fletcher, 2003; Sharpiro et al. 2011). In the core instruction, readers should be exposed to a range of genres and classic and contemporary literature along with a variety of informational texts for the purpose of extracting deep meaning and critical analysis of the text. Focusing specifically on RTI in reading, quality evidence-based reading instruction is sufficient for most students to learn to read at a standard level (Foorman,
According to the National Reading Panel (2000) and a multitude of researchers, effective Tier 1 instruction should include explicit instruction in phonemic awareness, phonics, and automatic recognition of high-frequency irregular words; an emphasis on vocabulary; the development of background knowledge; explicit instruction in meaning making reading strategies; and a plethora of opportunities to read and respond to text. In addition, the instruction should be differentiated to meet the individual needs of the students (Conner, Piasta, Fishman, Glasney, Schatschneider, Crowe, 2009; Gersten et al., 2009; National Reading Panel, 2000; National Reading Council, 1998). In Tier 1, growth rate of all students is measured to regulate if: (1) the mean rate of growth across all students is sufficient to demonstrate mastery of the content as compared to other classrooms in the school, district, state, and/or nation, and (2) the general education curriculum is sufficient for most students, to begin the identification process, and intervene for students whose level of performance is below the aggregate mean (Fuchs & Fuchs, 1998).

In essence, Tier 1 is the basis for a guaranteed and viable curriculum for all students (see Table 2: RTI Components Summary). The guaranteed and viable curriculum delivered in a quality Tier 1 instruction program should meet the needs of 80-85% of students (Fuch & Fuchs, 2006; NASDSE, 2006; Sharpiro, 2011) as measured by the benchmark assessments. If this occurs it may be presumed that the core instructional program is effective (Greenwood & Kim, 2012; NASDSE, 2006). RTI Tier 1 core instruction alleviates the question of poor or inadequate instruction as the cause of a student’s inability to achieve academically (Graner, Faggella-Luby, & Frischmann, 2005).
**Tier 2 and Tier 3: Targeted and intensive intervention.** If the universal screener shows a lack of academic progress for an individual student, an intervention is added. These intervention levels are sometimes known as secondary and tertiary instruction following the precursor from the medical field, or more commonly in the education setting as Tier 2 or Tier 3. These interventions are above and beyond the core instruction and are focused on the area(s) of weakness for the individual student. Students receive interventions geared to supplement the general education curriculum with the attempt to meet their individual needs. The instruction within Tier 2 and Tier 3 should align to the core instruction, Tier 1, so as to reinforce and complement the core instruction (Bean, 2009; Fuchs, Fuchs, & Compton, 2012). An empirically-validated intervention, one that has been demonstrated as effective for most students, is used during the intervention (Batsche, Elliott, Graden, Grimes, & Prasse, 2005; Strangeman et al., 2006).

Student progress is monitored frequently, and appropriate adjustments are made to the intervention based on the data (Fuchs & Fuchs, 1998; Fuchs & Fuchs, 2006; Fuchs & Vaughn, 2005). Adjustments to the intervention may include altering the instructional program, modifying the length or timing of the intervention, or changing the instructor or group arrangements (Fuchs, Fuchs, & Speece, 2002; Fuchs, Mock, Morgan, & Young, 2003). As a student moves up through the tiers, the intervention intensity increases. Tier 3, many times the final tier, is for those students who are non-responders to the Tier 2 intervention. In some models, Tier 3 includes special education services; while other models extract special education from RTI (Coyne, Kame'enui, & Simmons, 2004; Fletcher, Coulter, Reschly, & Vaughn, 2004, NCRTI, 2010). If Tier 3 intervention does not demonstrate sufficient growth, it may be assumed that the lack of response to the corrective actions indicates special education services may be
warranted (Fuchs & Vaughn, 2005). The intensity and duration of the intervention is increased as a student moves to higher tiers, but student placement in the tiers is fluid and based on ongoing student data.

Converging evidence has supported the effectiveness of interventions for reading difficulties (Denton, 2012) with reviews and meta-analyses revealing a larger effect size in the early stages of reading instruction than those interventions provided after second grade. Wanzek and Vaughn (2007) found a larger effect from intervention provided to kindergarten and first grade students than students in third through fifth grade. Older students can be remediated through intervention, but the approach needs to be more intense (Torgensen, 2004). Finally, Marston (2005) concluded from his analysis of several experimental studies that 50 to 75% of students receiving Tier 2 intervention responded to the instruction (effect size for Tier 2 as compared to the control group was 0.55).

Several studies indicated that a paraprofessional providing the intervention can be effective provided: (a) the interventionist is carefully selected, (b) group size is kept small, (c) a highly structured reading intervention program is implemented, and (d) the paraprofessional is guided by a teacher or coach (Denton, 2012, Hauerwas & Gressling, 2008; Tileston, 2009).

**Group size.** Most group size studies (e.g. Prime Time; intervention) focus on the effectiveness of very small groups (i.e. 1:1 up to 1:5). Wanzek and Vaughn’s (2007) synthesis of the research on K-3 students found that greater academic impact was seen in small group settings. Schwartz, Schmitt, and Lose (2012) studied student achievement in relationship to teacher-student ratio (i.e. 1:1, 1:2, 1:3 and 1:5). Results showed significant gains in all measures for all groups and trend analysis showed a drop in gains as group size increased. While the more intense Tier 3 groups tend to fit this research, RTI Tier 2 groups tend to be between 5-8 students.
Denton (2012) concluded that the group size should be selected on an individual basis through the RTI problem-solving team.

**Timing.** Denton (2012) found insufficient research guidance on the ideal time to begin Tier 2 intervention. O’Conner et al., (2005) reported a large number of students who received intervention in kindergarten required subsequent interventions in later grades. However, Coyne, Kame'enui, Simmons, and Harn (2004) reported that kindergarten students who responded strongly to intervention did not require further intervention in first grade. Some researchers have found kindergarten to be a window of opportunity, while others found the challenges of identification in kindergarten precarious. Currently, results are mixed in regards to effectiveness.

**Dosage.** The total time in intervention, or dosage, is comprised of the duration (i.e. # weeks), the frequency (i.e. # sessions/week), and the minutes per session. With mixed experimental results, research provides limited insight into the dosage necessary to provide effective, long-range academic growth (Denton et al., 2011). Studies have reported dosage recommendations between 12 - 55 hours with a duration from 8 - 22 weeks, a frequency of 3 - 5 days/week, and a length from 20 - 40 minutes/session (Al Otaiba, Schatschneider, & Silverman, 2005; Fuchs & Fuchs, 2005; Fuchs, Compton, Fuchs, Bryant, and Davis, 2008; Elbaum, Vaughn, Hughes, & Moody, 2000; Gersten et al, 2009; Wanzek and Vaughn, 2007). Subsequently, several studies comparing various frequency and duration aspects have found minimal to no statistical difference between groups (Denton et al, 2011; Hatcher et al., 2006; Wanzek & Vaughn, 2008). Counter to this, Al Otaiba, Schatschneider, and Silverman (2005), showed large effect sizes for word reading and comprehension kindergarten intervention groups.
Currently, the *What Works Clearinghouse* recommends a Tier 2 dosage of 3 - 5 times/week for 20 - 40 minutes/session for a reasonable time. A reasonable time is dependent on the student's grade level and specific needs, but intervention duration should be at least five weeks. Common practice would recommend 8-9 week long cycles (Center for RTI, 2014; National Center for RTI, 2010; RTI Action Network, 2014).

The table below summarizes the multi-tier components of RTI. It illustrates the need for additional studies to refine protocols and provide sound decision-making structures for practitioners.

### Table 2

**RTI Component Summary**

<table>
<thead>
<tr>
<th>Tier</th>
<th>Core Instruction and Design</th>
<th>Population served</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier 1</td>
<td>Universal screener or Benchmark assessment (3 times per year)</td>
<td>100% of students</td>
</tr>
<tr>
<td></td>
<td>Scientifically-based core instruction</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Dosage:</strong> 90 minute, uninterrupted instruction for reading, plus language arts instruction</td>
<td></td>
</tr>
<tr>
<td>Tier 2</td>
<td><strong>Dosage:</strong> beyond Tier 1; 20-30 minutes per session, 3 - 5 days per week for 8 - 22 weeks</td>
<td>~15-20% of students</td>
</tr>
<tr>
<td></td>
<td><strong>Group:</strong> 3-7 students; base on program protocol</td>
<td>Students who were not achieving at expected levels for long range academic success</td>
</tr>
<tr>
<td></td>
<td><strong>Program:</strong> explicit, systematic instruction based on area(s) of weakness determined through benchmark assessment and teacher's analysis of student's areas of need</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Progress monitored:</strong> weekly to monthly</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Cycle outcome:</strong> Students may (1) be removed from Tier 2 and receive Tier 1 instruction only; (2) continue receiving Tier 2 support; or (3) receive additional support through Tier 3.</td>
<td></td>
</tr>
<tr>
<td>Tier 3</td>
<td><strong>Dosage:</strong> additional time beyond Tier 2; 30 minutes daily for 8 - 22 weeks</td>
<td>~5-10% of students</td>
</tr>
<tr>
<td></td>
<td><strong>Group:</strong> 1-3 students; based on program and student need</td>
<td>Students who were non-responsive to Tier 2</td>
</tr>
<tr>
<td></td>
<td><strong>Program:</strong> explicit, systemic intensive instruction based on area(s) of weakness as determined by the universal screener, progress monitoring, essential skills screener, and the teacher's analysis of the student's areas of need</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Progress monitored:</strong> two time per week or weekly</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Cycle outcome:</strong> may or may not include special education or English Language Learners</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* The table is a composite of the current guidance available. From Center for RTI, National Center for RTI, RTI 4 Success, and RTI Action Network.

---

4 The *What Works Clearinghouse* (WWC) is an initiative of the Institute for Education Sciences at the U.S. Department of Education. It is administered by the National Center for Education Evaluation. WWC provides credible and reliable evidence of the effectiveness of a practice, program, or policy for educational decision making. WWC analyzes research that meets a high research standard to determine the level of program effectiveness.
**RTI Assessments**

With an RTI goal of identifying and providing support to students who are at high academic risk, researchers have ascertained that two forms of assessment are critical to the prevention/intervention nature of RTI (Greenwood & Kim, 2012; NCRTI 2010). Universal screener and progress monitoring assessments have been shown to result in earlier identification of academic or behavioral struggles and lead to early intervention for children. Early intervention provides intense instruction at critical stages of learning and results in lower referral rates for special education services (Wallace, Espin, McMaster, Deno, & Foegen, 2007).

**Universal screener.** In the RTI framework, a benchmark assessment is used to universally screen all students. Typically educators would give all students the benchmark assessment at periodic intervals, three or four times per year (Fuchs & Fuchs, 2005). The goal of a benchmark assessment is to categorize student performance based on a predetermined cut score into risk levels (Stecker, Lembke, and Foegen, 2008). Schools utilize the assessment data along with classroom data to identify students at risk of academic failure. Through the analysis of data, students’ weaknesses are identified and matched to interventions when problems first appear (Yell, Shriner, & Katsiyannis, 2006), thus avoiding the wait issue related to the discrepancy model. The school then provides an evidence-based intervention and monitors student progress. The universal screener is used as a preventative approach, which is grounded in research from the medical field on early screening tests.

**Progress monitoring.** Following the use of a universal screener to identify areas of need, a measure is used to monitor responsiveness to the intervention. These assessments are delivered to students who participate in the intervention. Measurement may be based on performance at the end of the intervention, growth over the course of the intervention, or a blend of both as there is no standardized method (Fuchs, 2003). The most common and the most
recommended measure is the use of a progress monitoring probe, or curriculum-based measure (Batsche et al., 2005; Fuchs, 2003; Fuchs & Fuchs, 2006). Progress monitoring probe is used to examine student growth in a specific area throughout the intervention. It is a repeated measure of performance that is used to inform instruction. The probes are short, skill specific measures designed to monitor growth in a specific skill over time. Progress monitoring probes are administered to students at least monthly and up to twice per week depending on the level of intensity of the instruction and the monitoring goal. Progress monitoring assesses the rate of individual student’s improvement, identifies those students who are not making ample improvement under the current intervention, and permits a team of educators to determine the appropriate level of intervention on an ongoing basis (NCRTI, 2010).

Progress monitoring is an instructional technique based on formative assessment, a recognized form of assessment to guide instruction. The Research Center at the University of Minnesota found that progress monitoring is a means to improve not only learning, but overall instruction (Hallahan & Mercer, 2002). It permits teachers to gauge student growth by recording performance on series of short assessments over time. Progress monitoring can guide the teacher's instructional decisions-making based on student performance and academic trajectory (Duran, Hughes, & Bradley, 2011). Traditional wisdom would require at least 3-4 weeks of data and a minimum of four data points to guide decisions (Stecker, Lembke, & Foegen, 2008). Decisions may then be based on the overall data trend or the most recent set of data points, and should be weighed against the aim line. Progress monitoring is a seamless method to track student growth from one skill to another or from year to year (Wallace, Espin, McMaster, Deno, & Foegen, 2007).
Curriculum-based measures, or CBM, are one form of progress monitoring, which has a recognized body of evidence supporting their validity and reliability (Deno, 1985; Fuchs & Deno, 1994; Fuchs, Fuchs, & Maxwell, 1988; Jenkins & Jewell, 1993; Lembke & Busch, 2004; Shinn, Knutson, Collins, Good, & Tilly, 1992). CBMs are less vulnerable to bias associated with sub-groups (Stecker, Fuchs, & Fuchs, 2005; Vaughn & Fuchs, 2003). CBM are brief probes that take minimal time away from instruction and provide discrete skill measures that are sensitive to small learning changes (Deno, 1985; Fuchs & Deno, 1994). CBM probe results are typically plotted to demonstrate the trajectory of academic growth, or lack of growth, on discrete tasks. In addition, alternative forms of a probe produce consistent scores. DIBELS, the Dynamic Indicators of Basic Early Literacy Skills, is a commonly used reading benchmark and progress monitoring tool (Good & Kaminski, 2003).

Scores from a CBM represent global proficiency in the academic area that provides a different perspective than the traditional classroom assessment. Traditional assessments measure mastery on a segment of the curriculum. The global proficiency aspect permits teachers to gauge the students’ progress towards long-range academic goals (Stecker, Lembke, & Foegen, 2008). One such example was an eight year study of 1,362 students who demonstrated a significant positive and predictive relationship between oral reading fluency probes and scores on state testing (McGlinchey & Kixson, 2004). Progress monitoring, through the use of CBM has been shown to make teachers more mindful of the impact of their instruction on student outcomes and to improve overall student outcomes (Fuchs, Deno, & Mirkin, 1984; Stecker, Fuchs, & Fuchs, 2005). This increased awareness and use of data has led to improved instruction and increases in student achievement (Fuchs et al., 1984; Stecker & Fuchs, 2000). CBM have been shown to improve all layers of instruction, not just a presumption of better SLD identification.
Benchmark assessments are a universal screener provided to all students. The universal screener is used to identify struggling students along with evaluation of the quality of the core instructional practices (Graner et al., 2005). Progress monitoring measures are short, skill specific measures designed to monitor growth over time. The data from both the benchmark and the progress monitoring are used to guide a student’s instruction, move a student between tiers, and guide the school's curricular decisions.

**Data-based Decision Making**

The need for data-driven instruction has been well documented (CEC, 2008; Fuchs & Fuchs, 2006; Greenwood, et al., 2003; Mellard & McKnight, 2006, Walmsely & Vaughn, 2007). In the decision making process, a team of educators uses the students' responses from the progress monitoring data as the basis for educational decisions (Brown-Chidsey & Steege, 2005; NASDSE, 2005). Progress monitoring measures (e.g. oral reading fluency) are used since research has found them to be a better tool for making decisions (e.g. less time consuming, option for more frequent administration, and increased instructional utility) about a student’s progress than the traditional standardized assessments (Fuchs et al., 2001; Good et al, 2001). Decisions may include additional tiers of intervention, removal from the intervention with monitoring in the core instruction, or possible referral for eligibility for special education (Fuchs & Fuchs, 2006; Graner, Faggella-Luby, & Fruschmann, 2005; Greenwood et al., 2003; Kovaleski & Prasse, 2004; Kratchwill, Clements, & Kalymon, 2007).

Research would indicate that fidelity, or treatment integrity, to the delivery of an evidence-based program is crucial to maximize program effectiveness (Foodman & Moats, 2004; Fuchs & Fuchs, 2006; Gresham et al., 2000; Instructional Research Group, 2006; NASDSE, 2006; Telzrow, McNamara, & Hollinger, 2000). Although methodologies and protocols are
available, fidelity is an area of concern in the implementation of the RTI framework (Gresham, MacMillan, Boebe-Frankenberger, & Bocian, 2000). Fidelity requires that the delivery of the instruction, including assessments, follows the structures intended and is validated by research (Gresham et al., 2000). Fidelity of implementation is a prescription the instructor must follow to assure the program maintains similar outcomes as the research indicates. As a team of educators use the student performance data to inform instructional decisions, assurances to program implementation fidelity are integral (Bianco, 2010). Within the RTI framework, fidelity to high quality instruction is the foundation that satisfies one of IDEA’s legal requirements (IDEA, 2004).

One final component recommendation by Fuchs, Mock, Morgan, and Young (2003) would be to include cognitive assessments for students who are non-responders in Tier 2. While not wishing to return to the discrepancy model, their belief is that without this additional component identification SLD identification may falter.

Finally, RTI’s multiple instructional tiers, use of assessment data for decision making, and review of the integrity of the curriculum address each of Heller et al.’s criterion (1982). While the assessments are an essential component of the RTI framework, the instruction is the driving force of the RTI framework (Shapiro, n.d.). The universal screener and the progress monitoring tools are gatekeepers to guiding the decision making process.

**Relationship with Special Education**

Special education is associated with RTI, but RTI is not a special education program (CEC, 2008; NCRTI, 2010). Rather, special education is an important component of a comprehensive RTI framework that incorporates core instruction and intervention levels. IDEA 2004 does permit the use of RTI as a method of identification for children with learning
RTI: DIFFERENCES IN READING ACHIEVEMENT

...disabilities through the use of the rate of improvement demonstrated over time to help make the decision (CEC, 2008; Fuchs & Fuchs, 2006). In part, RTI is designed to increase the accuracy of identification of children with a learning disability (Vaughn & Fuchs, 2003).

**RTI Approaches**

The research delineates two main approaches for the delivery of the RTI framework (Fuchs, Mock, Morgan, & Young, 2003). Both methods are cyclic as indicated in the decision-making model (Figure 1). The first is the standard treatment protocol that followed an if-then structure where all students receive the same research-based intervention following strict protocols. The second approach, the problem-solving model is more fluid in that each child’s individual needs are analyzed and an appropriate program is designed by a group of educators. Common components of both protocols are modeling, practice with immediate error correction, and the inclusion of motivational strategies (VanDerHeyden et al., 2006).

The standard treatment protocol has been studied more thoroughly and due to the scientific evidence supporting the standard treatment, experts (e.g. Fuchs & Fuchs) would lean towards the standard treatment protocol (Strangeman et al., 2006). While the well-defined procedures provide ample opportunity for replication, most implementations using the standard treatment protocol have been completed through a research project. The financial and personnel support provided by the research project present challenges for schools to replicate and little guidance has been provided to put the process into operation in a standard school setting (VanDerHeyden, n.d.). Currently more practitioners use the problem-solving method (Fuchs, Mock, Morgan, & Young, 2003; Strangeman et al., 2006) or a hybrid model.

**Standard treatment protocol.** RTI standard treatment protocol is similar to those implemented in the medical field (Nelson-Walker et al., 2013). The same validated, fixed-duration intervention is used with all students exhibiting the learning deficit (Fuchs et al., 2003).
This form of intervention is more typically implemented outside of the classroom by specialists or paraprofessionals who are well trained on the specific treatment process (Fuchs et al., 2003). Research indicates improved student achievement and special education placement, as represented by these studies:

- Marston’s (2005) analysis of several standard treatment protocol research studies showed significant impact on student growth and a decrease in placement.
- Torgeson et al. (2001) found a marked decrease in special education referrals from a program with high integrity, consistent fidelity, and a well-trained facilitator.
- VanDerHeyden, Witt, and Gilbertson's (2007) results indicated a decrease in special education referrals, an increase in the accuracy of the referrals (i.e. from 50% to 69.5%) that held true regardless of race or gender, and improvement in the disproportionality between males and females (i.e. from 1.52:1 to 1.35:1)
- Vaughn, Linan-Thompson, and Hickman (2003) showed improved reading scores with 53% of the students exiting the program. Of those who exited the program, most maintained adequate progress in the general education classroom.
- Vellutino, Scanion, Zhang, and Schatschneider's (2008) results indicated that 84% of students who received treatment in kindergarten and/or first grade met grade-level expectations in reading by the end of third grade.

These studies have demonstrated the power of standard treatment protocol to enhance learning (Fuchs et al., 2003; Marston, 2005; Torgesen et al., 2001).

**Problem-solving protocol.** The problem-solving protocol has been used for almost three decades and has been often credited to work done by the Heartland Area Education Agency in
Johnston, Iowa (Shapiro, Zigmond, Wallace, & Marston, 2011). This approach calls for individually tailored interventions. It typically calls for teachers to implement the intervention within the classroom (Fuchs et al., 2003). The protocol follows a 5-step process which includes: (1) problem identification through benchmark assessments, (2) problem analysis using multiple data sources showing a convergence of evidence, (3) selection of an intervention by a team of educators, (4) implementation of the intervention by the classroom teacher, and (5) monitoring of the student’s response through progress monitoring by the teacher and analysis of the results by the team (Fuchs & Fuchs, 2006; Fuchs, Mock, Morgan, & Young, 2003).

The team of educators may vary, but typically would include the classroom teacher and peers. Peers may include fellow teachers, special education teachers, the school psychologist, administrators, and other reading specialists (Batsche et al., 2005). The collaborative approach to problem solving builds the capacity of the individual members. Prior to the intervention, the team should discuss questions regarding the specific gaps in the student’s learning, potential underlying causes, available options to address the problem, and details of the implementation of the plan including progress monitoring plans and a time frame. Following the intervention cycle and progress monitoring, the team would reconvene and discuss the student’s growth data along with the fidelity to the treatment. While the problem-solving protocol has less data than standard treatment protocol (Fuchs et al., 2003; Kovalaski, Gicklinng, Morrow, & Swank, 1999; Marston, 2005; Tilly, 2003), studies have demonstrated positive effects on student achievement (Table 3). As with the previous discrepancy model fidelity is crucial to the implementation, which can be a challenge in the problem-solving protocol (Telzrow, McNamara, & Hollinger, 2000).
Table 3

Brief Overview of Key Problem-solving Protocol Studies

<table>
<thead>
<tr>
<th>Problem-solving Protocol Study</th>
<th>Research Design Procedures and Participant(s)</th>
<th>Measured Outcomes(s)</th>
<th>Author(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minnesota Problem-Solving Model (MPSM)</td>
<td>Historical Contrast Design (HCD); included collaborative consultation, N=121 students</td>
<td>Placement rates remained largely constant. Referral rates of minority students improved (baseline odds-ratio was 2.7 while RTI method odds-ratio ranged from 1.9 to 2.1, which is closer to the expected rate of 1.0).</td>
<td>Shapiro, Zigmond, Wallace, &amp; Marston (2011)</td>
</tr>
<tr>
<td>Heartland Area Education Agency, Iowa</td>
<td>3-tier RTI framework which excluded the use of cut scores on standardized tests N=121 schools</td>
<td>Placement rate remained constant during the study (25% baseline and study groups). However, placement rate decreased at all grade levels (based on a pre/post implementation analysis)</td>
<td>Fuchs, Mock, Morgan, &amp; Young (2003) Marston (2005)</td>
</tr>
<tr>
<td>Ohio Intervention-Based Assessment (IBA)</td>
<td>Correlation and Descriptive N = 277 schools (90% elementary) N = 291 student goals</td>
<td>Survey results indicated an overall improvement in student outcomes (academic or behavioral) as indicated by a 4/5 rating. Correlation between student outcomes and the problem-solving RTI method were fairly modest.</td>
<td>Telzrow, McNamara, &amp; Hollinger (2000)</td>
</tr>
<tr>
<td>St. Croix River Education District Model</td>
<td>Descriptive and QED with HCD; included collaborative consultation</td>
<td>Special education rates decreased (4.5% to 2.5%) over a 10 year period. Student reading outcomes improved from 35% to 70% meeting CBM cut scores.</td>
<td>Bollman, Silberglitt, &amp; Gibbons (2007)</td>
</tr>
<tr>
<td>Tiers of Reading Intervention (TRI)</td>
<td>HCD, 2 schools, p = 22 Facilitated by university researchers outside of the classroom</td>
<td>Compared to a historical contract group from the same school, the students receiving intervention showed improved reading skills. Placement rates decreased from 15% to 8% in 4 years.</td>
<td>O'Connor, Harty, &amp; Fulmer (2005)</td>
</tr>
</tbody>
</table>

Note. Examples of RTI research studies.

**Hybrid protocol.** A hybrid protocol is a blend of the problem-solving and standard protocols. By eliciting a pre-determined set of research-based intervention programs, the hybrid protocol utilizes the prescriptive nature of the standard-protocol approach. At the same time, the hybrid protocol emphasizes the diagnostic nature of the problem-solving approach by hand selecting the program based on the individual student needs rather than a one-size fits all approach. The intent of the mix is to highlight the advantages of each protocol. This blend can provide a moderate level of quality control and at the same time increase the sensitivity to individual differences (Fuchs, Mock, Morgan, & Young, 2003). The Idaho results-based model
(RBM) combined both for the purpose of providing strategic interventions (standard protocol) delivered by the teacher during Tier 2 and intensive interventions (problem-solving protocol) delivered outside the classroom during Tier 3 (Callender, 2007). Results of a five-year study indicated the RBM students progressed significantly more in reading than did the non-RBM students. As for placement in special education, over a three year period schools utilizing RBM saw a decrease of 3% and the non-RBM schools experienced a 1% increase.

More research is necessary to aid practitioners in designing a protocol that best meets the students' academic needs as well as the school's resource capacity. Regardless of the approach used, understanding the Tier 1 core reading program is critical to informing decisions based on the data and longitudinal research is required to determine whether the effects of intervention continue to impact children’s reading ability as they advance through school (Kerins, Trotter, & Schoenbrodt, 2010).

**Indiana RTI: A Response to Instruction**

In 2010, Indiana provided official guidance for the implementation through *Response to Instruction: Foundations for Implementation* (IDOE, 2010). Prior to 2010, some training had been provided. As of the 2010 guidance, all Indiana school corporations were required to move to the RTI framework (IDOE, 2010). Overall the Indiana RTI framework follows the traditional conceptual framework and principles expressed in the literature with an emphasis on the instructional aspect.

---

5 For the purpose of this study, RTI represents Response to Intervention even though Indiana has modified the name as Response to Instruction. The emphasis on instruction was intended to focus on the continual change and adaption of instruction to correspond to student’s needs. But for the purpose of utilitarian use of this study, I used the more traditional, national name for RTI (Response to Instruction, 2010).
The visual representation of the Indiana RTI framework (Figure 5) shows a fluid and overlapping model in which struggling students receive additional services beyond the core instructional time to support their learning. Indiana leadership modified the model slightly to emphasize this fluid, overlapping model. Instead of traditional name of Response to Intervention (RTI), the Indiana model is called Response to Instruction (RtI) to emphasize quality instruction at all levels of instruction. RtI is intended to provide a continuum of services for all learners. The Indiana definition includes more than the struggling students. The model is all-inclusive and intended to provide a continuum of services to ensure all students meet or exceed proficiency standards. It specifically states that it is not a special education initiative, but instead a systemic process that aligns all school improvement goals. This guidance influenced the district’s instructional protocols for core instruction along with supplemental programming.

![Cylinder Model of RtI Framework](image)

*Figure 5. Cylinder Model of RtI Framework.*

*Note:* The cylinder model of RtI framework demonstrates the inclusionary nature of the RtI framework and the emphasis on additional instructional time and resources as a student moves through the tiers. From IDOE. Response to Instruction (RtI): Foundations for implementation (2010).
Non-responsiveness

Within the RTI framework, non-responsiveness must be defined. Determining adequate or inadequate response to the intervention is critical for practitioners’ decision-making (Fuchs, 2003). Researchers have not uncovered a single operational definition of adequate response (Fuchs & Fuchs, 2006; Hintze, 2008; Ridgeway, Price, Simpson, & Rose, 2012). In general, students who make minimal gains after receiving intervention are described as non-responsive and may be eligible for special education services (Fuchs & Fuchs, 2006; Linan-Thompson, Vaughn, Prater, & Cirino, 2006; NASDSE, 2006). Common approaches include idiographic (i.e. slope) and nomothetic (i.e. risk and/or final status) with some implementations combining the two methods (Christ & Hintze, 2007). By combining the learning rate and the performance level, decisions are made based on the student's response to the intervention, a pre-established goal, and the student's current academic risk (Callendar, 2007).

Different criteria are likely to result in different student identification for adequate or inadequate response (Barth et al., 2008; Fuchs et al., 2008). Non-responsiveness has been based on end of treatment results (i.e. cut points, standard measures, or criterion-referenced benchmarks) and response over time (i.e. CBM data results on the level of performance and slope of the growth line) (Demo, 1985; Fuchs & Fuchs, 2006; Good, Simmons, & Kame'enui, 2001; Speece, & Case, 2001; Torgenson et al., 2001). The operational and conceptual differences between the problem-solving or standard treatment protocols create the inherent difference in the meaning of non-responsiveness and the pending identification outcomes (Fuchs & Fuchs, 2006). The problem-solving protocol determines non-responsiveness relative to the classroom or grade-level population, where standard protocol bases non-responsiveness on the results of the program research. The more systematic standard treatment protocol does identify
children in need of services (i.e. true positives). However, the numbers of false negatives (i.e. children who are deemed not to need services but are unable to achieve with only Tier 1 instruction) mitigate the true positives (Fuchs & Fuchs, 2006). On the other hand, the less systematic problem-solving protocol is less likely to identify false negatives, but more likely to identify false positives. These identification errors beg the question: Is it more advantageous to err on the side of over identification (false positive of a problem-solving protocol) or under identification (false negative of a standard treatment protocol)? Determination of the protocol selected for use may be identified on the basis of the primary intent of identification or prevention.

The mixed results are somewhat based on the range of strategies employed to establish responsiveness. Strategies differ in (a) determination of achievement level (i.e. growth versus cut score), (b) the comparison used (i.e. predetermined cut score or comparison to peer performance), (c) measures used (i.e. oral reading fluency or comprehension), (d) assessment protocols (i.e. timed or untimed, summative or progress monitoring, oral, written, or a combination, or individual verses group assessment delivery), and (e) the specifics of determining adequate or inadequate response (i.e. as a specific cut score, benchmark, or slope criteria) (Fuchs et al., 2008). The research in this area is limited and the results are diverse. Researchers suggest additional studies should be conducted to refine and define the level of non-responsiveness for practitioners (Fuchs & Vaughn, 2005).

**Considerations for SLD Identification**

While many questions related to the RTI framework remain unanswered (Al Otaiba, Folson et al., 2011), considerations for SLD identification nuances should be weighed. The Learning Disability Association of America’s study (Hale et al., 2010) strongly supported the
use of RTI to address the intervention needs of students, while not supporting the use of RTI as the sole means of identification of SLD. In fact, Brown-Chidsey & Steege (2005) advocated for mandated RTI to advance academic achievement. While RTI is set in theory, it currently has a limited empirical base regarding educational effectiveness. Many researchers would say the scientific data is not sufficient to make RTI the sole identification method for SLD (Al Otaiba, Folson, et al., 2011; Berninger & Holdnack, 2008; Decker, 2008; Feifer, 2008; Fletcher-Janzen & Reynolds, 2008; Fuchs & Deshler, 2007; Fuchs & Fuchs, 2006). This is in part due to the nature of the definition of SLD by which non-responsiveness determines inclusion into a group (Kavale, 2002). Further studies regarding the accuracy of a non-responsiveness SLD determination are merited before RTI should be the exclusive method for identification. The LDA (2010) analyzed survey data from 58 expert panel participants from the fields of education, psychology, medicine, and law. An overwhelming majority found that neither the ability-achievement discrepancy model nor the response to intervention model alone should identify a student as having a specific learning disability. As with the discrepancy model, the RTI framework does not provide a means to differentiate diagnosis of SLD and other disorders, nor does it sort out low achievers from SLD identification (LDA, 2010). However, the experts did deduce that the RTI framework could be used to prevent learning problems through early intervention. As for identification of SLD, experts in the field recommend a comprehensive evaluation including cognitive and neuropsychological processes in addition RTI (CEC, 2008).

The Learning Disability Association of America study found an inconsistency in the belief of the expert panel (43% strongly agree or agree; 29% neutral, and 28% disagree or strongly disagree on Item 10 of the survey) in regards to the use of the RTI framework prior to a comprehensive evaluation of a child for SLD. The LDA (2010) noted the disagreement in
ratings appeared to be reflected by a procedural component of the RTI framework, specifically the use of increasing intervention intensity and a lack of consistency in implementation across schools, districts, and states. An array of studies (Berninger, 2006; Berninger & Holdnack, 2008; Decker, 2008; Elliot, 2008; Feifer, 2008, Feifer & Della Toffalo, 2007; Flecher-Janzen, 2008; Fletcher-Janzen & Reynolds, 2008; Hale, 2006; Kaufman, 2008; LDA, 2010; Reynolds & Shaywitz, 2009; Torppa et al., 2007), clearly supported a comprehensive evaluation in conjunction with the RTI framework.

As with the discrepancy model, the RTI approach for identification of SLD has some issues (Barth et al., 2008; Berninger & Holdnack, 2008; Decker, 2008; Feifer, 2008; Feifer & Della Toffalo, 2007; Fletcher-Janzen & Reynolds, 2008; Ruchs & Deshler, 2007; Fuchs & Fuchs, 2006; Fuchs, Fuchs, & Compton, 2004; Hale, 2006; LDA., 2010; Speece, 2005; Torppa et al., 2007). These issues include (a) a lack of one model for the delivery of RTI (i.e. standard protocol, problem-solving, or a hybrid method), (b) inconsistency in measurements within the model, (c) no standard curriculum or instructional methods, and (d) the current lack of empirical support for the RTI process in correlation to SLD identification. Even with the challenges mentioned, the RTI framework is an improvement over the ability-achievement discrepancy model due two key factors: the elimination of the time lag for support through intervention and monitoring the use of quality instructional strategies at all levels of service (Fuchs & Fuchs, 2007). RTI may be valuable as a first step in narrowing the student population to help aid in identification of SLD, but without research data on the issues mentioned previously, experts recommend additional steps for SLD identification.
Summary

As mentioned previously, identification of SLD rose to over 6% of the population using the ability-achievement discrepancy model. Prior to the 2004 reauthorization of IDEA, identification of students with a specific learning disability was determined primarily by a model which required a severe discrepancy between a student's achievement (i.e., standardized assessments, grades) and the student's intellectual ability (i.e., IQ). The discrepancy model has been called the "waiting-to-fail" approach because it has been known to allow students to struggle and fail in school prior to a determination of eligibility for support services (Downer, Rimm-Kaufman, & Pianta, 2007). Denton (2012) stated that withholding intervention may significantly diminish the probability that an acutely impaired reader will learn to read adequately. Additional flaws in the discrepancy model include issues with the conceptual framework of the model; measurement precision, consistency, and accuracy in identifying struggling students (Vaughn & Fuchs, 2003); and a tendency to discriminate against select populations.

NCLB (2001) and ESSA (2015) set high expectations for student achievement. As with many professions, the use of evidence to inform practice has increased in the education field (Nelson-Walker et al., 2013). Teachers must embrace research-based instructional strategies (i.e. Response to Intervention and evidence-based core instructional programs) so as to not lose instructional time on ineffective practices that are not meeting individual student's needs (Rock, Thead, & Gable, 2006). At the heart of RTI is a desire to provide all students with the greatest opportunity to succeed academically. The theory behind RTI is that through assessments teachers can recognize students who are struggling and address the problems before they become pervasive enough to merit referral for special education services (CEC, 2008; Samuels, 2008;
Mask, McGill, 2010). According to NCRTI (2010), the goals of RTI are to utilize all resources available to negate the consequences of learning struggles and to bolster the process for appropriate disability identification. Denton et al. (2006) stated that RTI increases in frequency, duration, and intensity of the intervention to address the student's individual needs and is theoretically grounded in research-based practices including differentiation, early intervention, and evidence-based instruction.

Due to its emphasis on student outcomes and intentional connection between identification and instruction, RTI holds considerable promise to identify students with SLDs earlier and more reliably, reduce the number of students who are referred inappropriately to special education, and reduce the over and under identification for special education services for students in various subgroups (Batsche et al., 2005; Donavan & Cross, 2002; National Joint Committee on Learning Disabilities, 2005; Kavale et al., 2005; Vaughn & Fuchs, 2003). Key benefits of the RTI framework are (a) early identification of learning challenges, (b) potential for a reduction in the screening bias of teacher-based referrals, (c) the relationship between the identification process, assessment, and instruction, and (d) a foundation based on hypothesis testing (Hintze, 2008). As mentioned earlier, while these benefits are currently theoretically accurate, the data is not sufficient to provide definitive guidance to practitioners.

Shapiro, Zigmond, Wallace, and Marston (2011) noted in their book, Models for Implementing Response to Instruction: Tools, Outcomes, and Implications that while there is a wealth of information on the RTI process, the same cannot be said for real-life implementation of RTI. A gap in the current research on RTI would be examples of actual program outcomes (Fuchs & Deshler, 2007; Glover, DiPerna, and Vaughn, 2007; Kerins, et al. 2010; Nelson-Walker, et al., 2013; O’Conner, et al, 2005; Vaugh & Fuchs, 2003). The sheer volume of
information available on RTI does not provide the focus practitioners need to validate the
protocols, procedures, and resources for the implementation of RTI through best practices
The individual components and the conceptual foundation of RTI have a solid research base, but
as a unit there are many gaps remaining. While a growing number of experimental and quasi-
experimental studies have evaluated outcomes across multi-tier reading intervention models,
Denton (2012) identified the research needs related to response to intervention (RTI) in reading.
The identified needs included (a) intervention effectiveness in regard to programming, the
interventionists, group size, and other organizational and technical aspects, (b) impact on limited
English proficiency, (c) next step measures for students who continue to struggle in Tier 3, (d)
criteria for determining responsiveness to the intervention, and (e) the impact and effects of a
fully implemented RTI framework.

Educators and researchers alike know that one size does not fit all when teaching students
to read. This is one reason Chhabra and McCardle (1998) stated that reading research must
follow large numbers of real children, in real classrooms for long periods of time to evaluate the
effectiveness of instructional programs. When more of the studies with a cross-section of
children, in a wide range of settings, study a phenomenon and yield similar results, the
confidence level in the findings increases. Converging evidence is critical as a basis for policy
and operational definitions. RTI is grounded in strong theory, but the mass of empirical data at
this time does not yield a strong operational definition for practitioners. This research attempts to
help add to the body of knowledge on the impact and effects of a fully implemented RTI
framework by providing a longitudinal study of several cohorts of students.
CHAPTER 3
RESEARCH METHODS

This chapter describes the mixed-methods design employed to examine the outcome of the Response to Intervention (RTI) process on student achievement. The chapter begins with the rationale for the selection of a mixed-methods approach followed by details of the quantitative and qualitative phases of the study. The chapter is organized into the following sections: (1) purpose of the study, (2) research questions, (3) research design, (4) population and sample, (5) instrumentation, (7) data collection procedures, (8) data analysis, and (9) limitations of the study.

Purpose of the Study

The purpose of this study was to examine the outcomes of the RTI process, such as student achievement in reading and/or placement in special education. In addition, the data provided evidence to the district to guide program decisions. The study investigated whether students identified as lower achieving readers in kindergarten or first grade, who received deliberate intervention in reading within the RTI program, had higher levels of reading achievement by the end of third grade than students who did not receive the RTI reading intervention. The independent variables in this study were inclusion in the intervention or the non-intervention group, gender, socio-economic status, implementation year, cohort, school, and entry grade. Academic achievement in reading was the dependent variable and was measured by reading assessment results and placement in special education.
Research Questions

Given this is a mixed-methods study, this research encompassed a primary or central question with sub-questions to guide the exploration of the data collection (Creswell, 2003). This overarching question guided this study:

1. What are the long-range (3 or 4 years) differences on reading achievement for students who participated in RTI Tier 2 reading intervention as compared to students who scored slightly above the intervention group and did not receive treatment?

The following sub-questions guided the research and data analysis for this study:

2. What differences occur among the RTI Tier 2 groups as compared to those students just outside the intervention group in regards to special education identification by the end of third grade?

3. What differences are seen in students' reading achievement among and within the various nested groups (e.g. school, implementation year, achievement measures)?

4. From the perspective of the focus group, what understandings result from the quantitative data analysis of the academic outcomes between and among the nested groups?

Research Design

Yin (2009) stated, that the research question commands “a logical plan for getting from here to there” (p. 20) to meet the purpose of a study. As evidenced by the number of studies in the literature, researchers have found mixed-methods research to be useful when studying educational issues (Truscott, Swars, Smith, Thornton-Reid, Zhao, & Dooley, 2010). The logical path for these research questions was a mixed-methods approach that included a longitudinal, quasi-experimental design followed by an explanatory study.
Researchers consider a mixed-methods design an effective way to deepen the understanding of a complex phenomenon by analyzing one type of data with another type (Creswell, 2002; Plano-Clark et al., 2014; Richards & Morse, 2013; Tashakkori & Teddlie, 2003, 2010). The statistical nature of quantitative research provides a general explanation (Creswell & Plano, 2010); while a qualitative study provides a more detailed understanding of the phenomena (Crewell & Plano Clark, 2011). The mixed-methods approach is guided by the principle of pragmatism (Alise & Teddlie, 2010). The rationale behind pragmatism is that the general and the specific together can tell researchers more than either one alone (Alise & Teddlie, 2010). The strengths of the designs are merged, while the weaknesses are mitigated (Caruth, 2013; Gay, Mills, Airasian, 2006; Green, Caracelli, & Graham, 1989; Tashakkori & Teddlie, 1998). Additionally, questions can be uncovered for further study (Caruth, 2013; Lopez-Fernandez & Molina-Azorin, 2011). In this study, the general nature of the statistical analysis was enhanced by a sample of stakeholders' collective perspectives.

By not limiting the data collected, a mixed-methods approach seeks convergence through triangulation of the data. Triangulation reduces bias by substantiating results and as a result increasing reliability (Greene, Caracelli, & Graham, 1989). Triangulation seeks to obtain clarity and interpretation through mutual viewpoints about the phenomenon (i.e. complementarily), to enhance the findings by revealing more details (i.e. expansion), and to extend the findings with an aim for one method to inform the other (i.e. development) (Caruth, 2013; Creswell, 2003; Greene et al., 1989). This research used multiple data sources (i.e. a range of student achievement data, identification for an SLD in reading, and focus group data) to deepen the understanding of the outcome of RTI on student achievement. As a result of triangulating the data, the study's purpose was realized by providing expansion and development of the statistical
RTI: DIFFERENCES IN READING ACHIEVEMENT

analysis through complementary viewpoints.

Following the lead of the research questions, the key principles of a mixed-methods study that need to be determined include: (a) analytic weighting of the components, (b) completeness of the core component, (c) theoretically driven by the core component, and (d) appropriate sequencing (Morse, Niehaus, Wolfe, & Wilkins, 2006). I was cognizant of the value of the statistical analysis of the student achievement data; therefore, the theoretical drive was dictated by the quantitative component that held the primary weight of the project and could have stood alone. Moreover, the quantitative study was longitudinal including three research sites and six existing cohorts of students that provided nested groups for analysis. The nested group analysis provided completeness to the core study.

The use of longitudinal research on the outcome of an intervention has a solid base in recent literature. In general, longitudinal studies provide the opportunity to examine patterns over time, provide evidence of long range impact, reduce the likelihood of cultural differences (i.e. participant background knowledge, teacher changes) impacting results, and address gaps in the literature (Gay, Mills, and Airasian, 2006; Plano-Clark et. al, 2014). All would be applicable in this research as the study followed cohorts of students through repeated data collections. In addition, tracking several cohorts permitted the researcher to look at changes over time in the program effectiveness.

Finally, the timing (e.g. concurrent or sequential) of a mixed study refers to the order and nature of the components (Creswell & Plano, 2011; Teddlie & Tashakkori, 2009). This study followed a fixed-emergent, sequential design. The consecutive nature required data collection and analysis in two phases (Creswell, Plano Clark, Gutmann, & Hanson, 2003; Caruth, 2013; Onwuegbuzie & Teddlie, 2003; Tashakkori & Teddlie, 1998). According to Creswell and Plano
(2011), the fixed-emergent design allows details of the qualitative phase to emerge based on the interpretation of the results from the initial quantitative phase. In this research, the role of the qualitative data was used to provide insight on the statistical data analysis from the perspective of stakeholders. The two phase design provided opportunity to extend the statistical findings with anecdotal data. Beyond enhancing the results of the first phase of the study, the sequential approach was used to offset the sample size limitations. Based on the weight difference and sequential arrangement of the data collection, this research may be characterized as QUAN --> qual, more specifically an explanatory sequential mixed-methods design (Creswell & Plano, 2010; Gay, Mills, & Airasian, 2006; Tashakkori & Teddlie, 1998).

Explanatory sequential design is well characterized in the literature (Creswell, et al. 2003; Tashakkori and Teddlie, 1998). The classic version of an explanatory sequential design includes quantitative data collection and analysis followed by qualitative data collection and analysis. This pairing results in a richer interpretation of the results. For this research, the design supports the purpose to provide data on the implementation of the RTI, as well as the continuous improvement of the district’s program.

Details of each phase of the study will be described in the next sections. Figure 6 presents a visual depiction of the longitudinal nature and nested cohort groups.
Figure 6. Visual Model for Explanatory Sequential Mixed-Methods. Adaptation based on Plano-Clark et al. (2014).

Conceptualizing Longitudinal Mixed-Methods Designs.
Quantitative Research Design

A quasi-experimental design (QED), specifically non-equivalent group design (NEGD), was selected to address the research questions. NEGD is similar to pre-test/post-test control group experimental design, but uses intact groups and consequently lacks random assignment (Gay, Mills, Airasian, 2006). The rationale behind NEGD is to provide intervention to those in most need (Trochim, 2006), which parallels the underlying philosophy of the RTI framework. Even though random assignment increases the internal validity, many situations do not permit the option of random assignment (Gay, Mills, & Airasian, 2006). For ethical reasons, this research would not have permitted random assignment of students to the Tier 2 intervention group; therefore, a form of QED was the most appropriate.

NEGD provides improved control over other designs by partially controlling history, maturation, testing, instrumentation, selection, and mortality (Gay et al., 2006; Trochim, 2006). NEGD assigns participants to the treatment group based on cut scores and compares them to the control group comprised of participants just outside the cut score range. This research assigned students to cohort groups following the pre-test results. Kindergarten and first grade students who scored the lowest on the assessments were placed in Tier 2 and the next lowest were placed in the non-intervention group. This allowed for a partial match of relevant characteristics (e.g. age, instructional experience) between the two groups. The testing variable difference may be controlled by statistical methods to further equate the groups (Johnson & Christensen, 2010). NEGD allowed the use of existing RTI Tier 2 groups and the district assessments (e.g. DIBELS and Scholastic Reading Inventory) consequently enhancing this research design from an administrative perspective.
This longitudinal NEGD study included six cohort sets and three waves. The cohort data will be captured at the point of entry (e.g. kindergarten or first grade) into RTI Tier 2 and then compared to the results in a snapshot fashion three or four years later (e.g. third grade). It is a blend of a cohort study (i.e. same participant characteristics over time) and a trend study (i.e. an examination of changes in the population over time) (Creswell, 2012). While participants within the study remained the same, a panel study (Creswell, 2012) was not conducted since only group data were analyzed. Figure 7 represents the multiple cohorts (e.g. N_1 and N'_1), as well as the multiple pre/post-test (e.g. O) data. Further details of this chart are explained in the sampling and instrumentation sections.

<table>
<thead>
<tr>
<th>N_{1,2,3,4,5,6}</th>
<th>O_{1,2,4}</th>
<th>X</th>
<th>O_{1,2,3,5,6}</th>
</tr>
</thead>
<tbody>
<tr>
<td>N'_{1,2,3,4,5,6}</td>
<td>O_{1,2,4}</td>
<td></td>
<td>O_{1,2,3,5,6}</td>
</tr>
</tbody>
</table>

*Figure 7. Design Notation*

N represents the Tier 2 intervention groups while N' represents the non-intervention group. Each N represents the sum of the three school (School A, School B, and School C) groups. O represents pre-test and post-test results from the various measures. O1 represents pre- and post-test DIBELS results. O2 represents pre- and post-test mClass TRC results. O3 represents pre- and post-test SRI results. O4 represents TERA 3 summative results at the onset of kindergarten (pre-test) O5 represents IREAD-3 summative results during 3rd grade. (no pre-test) O6 represents ISTEP+ summative results during 3rd grade. (no pre-test)

X represents the treatment.

**Qualitative Research Design**

In this sequential mixed-methods research, the second phase was an explanatory study. An explanatory study is typically concerned with non-statistical methods of inquiry and draws findings from themes and categories through the data analysis process (Grethel, 1998; Helenka, 2010; Richards & Morse, 2006). The purpose was to collect perspective data to deepen the understanding of the phenomenon (Creswell & Clark, 2010; Creswell & Plano, 2010). In this research, the explanatory phase consisted of four focus groups made up of a purposeful sample.
of stakeholders from each school. A focus group is a dynamic group conversation used to gather information. It has been shown to be an effective means to gather qualitative data about educational or instructional effectiveness (Ledeman, 1990). The central assumption of a focus group is that candor is acceptable because participants understand the common issue, feel comfortable with each other, and draw strength from each other. This belief is grounded in group therapy methods. In addition, the focus group helps unfold the “why” behind a behavior or outcome (Ledeman, 1990). According to the research, there are five assumptions atop which focus group structure rests: (1) people are a valuable source of data, (2) group input is sounder than individual, (3) people can articulate their beliefs and actions into words, (4) people need guidance of the facilitator to extract information, and (5) the natural interactions and exchange of the group can be used to bring information to light about the phenomenon. The focus groups discussed the first phase of the data results of the study in relationship to their experiences and personal knowledge of the students and implementation process. Focus groups were used to (1) gain perspective on the quantitative results, (2) provide a holistic view of the outcome of RTI implementation, and (3) address research sub-questions.

Trustworthiness. Trustworthiness of qualitative work equates to validity and reliability in a quantitative study (Bowen, 2005). Aspects to be considered include credibility, transferability, dependability, and confirmability. Credibility, one of the most important aspects when affirming trustworthiness, may be established by creating an environment of trust, using conventional research techniques, and triangulation of the data (Shenton, 2004). In this study, cultural familiarity of the researcher with the focus group participants was previously established. In addition the focus group interviews were analyzed using conventional techniques of initial coding and member checks, which will be described in more detail in the "Qualitative
Data Collection" section. Transferability and dependability are established through the researcher's detailed explanation of the study (Shenton, 2004), as are provided. Additionally, the multiple focus groups increased the dependability. Finally, confirmability is to qualitative research as objectivity is to quantitative research. The key to confirmability is to ensure the findings are the result of the experiences of the participants rather than the researcher. Within this study, the triangulation of the data supports the confirmability aspect of trustworthiness.

**Population and Sample**

The research utilized six years of archival data from six cohorts of students from three demographically similar elementary schools (School A, School B, and School C) within a public school district. The selected district is located in a moderate-size, semi-rural Midwest community consisting of three small towns adjacent to a metropolitan city and a university town. The community had a varied economic base and a large commuter base (Indiana Department of Work Force Development, 2014; STATS Indiana, 2014). The school district had one high school (grades 9-12), one middle school (grades 7-8), and three elementary schools (grades K-6). Each elementary school had 90 to 160 students per grade level. For the purpose of the research study, the school selection was a convenience sample. Convenience sampling, or opportunity sampling, met the research requirement of providing evidence to the district leadership as to the district's RTI program implementation efforts.

A review of the individual school data reflects the homogeneous nature of the district (Appendix A). In general, the student population of each school was between 90-95% Caucasian with an approximate socio-economic status (SES) of 30% free/reduced lunch. The populations were similar academically as noted by kindergarten TERA 3 Reading Quotient data (i.e. 45-54% of the students fell below average based on national norm data) and similar district
state testing results (Appendix B).

Given the specific inclusion criteria of the RTI framework, random sampling was not possible, so a purposeful sample was taken. According to Tashakkori and Teddlie (2010) and Patton (1990), purposeful sampling addresses the specific purpose in regards to the research question and each member of the sample would be selected because of the rich information they would provide. To answer the research questions in this study, participants’ archival data were tracked from the time of intervention through third grade. Kindergarten is the first opportunity for educators to provide instruction and evaluate students’ responses and, as such, provides the first opportunity for early intervention (Al Otaiba et al., 2011). To establish the Tier 2 intervention and non-intervention comparison groups, all students in the district were assessed using standardized benchmark assessments. Following the assessment, each RTI team rank ordered students within the grade level according to district protocols (i.e. rank order based on benchmark assessment results along with teacher ranking). Table 4 provides a detailed visual representation of the cohort groups. It is noteworthy to mention that each cohort group represents the sum of the three schools’ sub-cohort groups. In other words, \( n_{1} \) includes School A \( (n_{A1}) \), School B \( (n_{B1}) \), and School C \( (n_{C1}) \), and so forth throughout the cohorts. In this NEGD study, a school Tier 2 group \( (n \approx 20) \) consisted of students with the lowest reading achievement. A total of approximately 360 students met those criteria. All criteria listed hold true for the comparison group with the exception of the first criteria. The comparison group included the next lowest 15-20% of the grade-level population.
Table 4

Cohort Groups

<table>
<thead>
<tr>
<th>School Year</th>
<th>Cohort Name (x)</th>
<th>( n_x )</th>
<th>( n_{xA} ), ( n_{xB} ), or ( n_{xC} )</th>
<th>( n'_x )</th>
<th>( n'<em>{xA} ), ( n'</em>{xB} ), or ( n'_{xC} )</th>
<th>Data Collection Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kindergarten cohorts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008-2009</td>
<td>1</td>
<td>60</td>
<td>20</td>
<td>60</td>
<td>20</td>
<td>2008-2009 to 2011-2012</td>
</tr>
<tr>
<td>2009-2010</td>
<td>2</td>
<td>60</td>
<td>20</td>
<td>60</td>
<td>20</td>
<td>2009-2010 to 2012-2013</td>
</tr>
<tr>
<td>2010-2011</td>
<td>3</td>
<td>60</td>
<td>20</td>
<td>60</td>
<td>20</td>
<td>2010-2011 to 2013-2014</td>
</tr>
<tr>
<td>1st grade cohorts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009-2010</td>
<td>4</td>
<td>60</td>
<td>20</td>
<td>60</td>
<td>20</td>
<td>2009-2010 to 2011-2012</td>
</tr>
<tr>
<td>2010-2011</td>
<td>5</td>
<td>60</td>
<td>20</td>
<td>60</td>
<td>20</td>
<td>2010-2011 to 2012-2013</td>
</tr>
<tr>
<td>2011-2012</td>
<td>6</td>
<td>60</td>
<td>20</td>
<td>60</td>
<td>20</td>
<td>2011-2012 to 2013-2014</td>
</tr>
</tbody>
</table>

Note: \( n_x \) represents nonequivalent groups from each school year and \( n \) represents the Tier 2 intervention groups while \( n' \) represents the non-intervention group. Each \( n \) represents the sum of the three school (School A, School B, and School C) groups.

Within the district of study, the percentage of students in Tier 2 was not fixed, but rather determined by the RTI team. The RTI team consisted of all grade-level teachers, the building administrators, the grade-level special education teacher, the building reading coach, and the school psychologist. The district and school administrators believed the constraints of up to 20% inclusion in Tier 2 provided the best learning opportunities for those students in need. This value followed national trends for RTI Tier 2 percentages and parallels state testing proficiency results for the district (i.e. ISTEP+ ELA pass rate between 78.5% and 87.4%).

One additional caveat exists within the 20% mark; students who were receiving special education services for reading were not provided services within the Tier 2 framework of RTI. The services provided to these students were provided by the special education staff. As such, these students were excluded, or skipped, in the selection process for Tier 2 and, therefore, are not represented in either group.

While similar groups provide more sound results (Gay et al., 2006), the treatment and control groups are not identical in NEGD research. However, the longitudinal nature of this study created cohort combinations that were repeated over three years, within and among the schools. This provided similarly-demographic nested groups to study (e.g. comparison across
schools, comparison across years within the implementation development). A sample of the nested groups is represented in Figure 8 below.

![Diagram](image_url)

*Figure 8. Sample Nested Groups. Adapted from Upton & Cook, 2014*

**Environment**

In addition to similar student demographics, each school's RTI program was similar. Each began in 2008 and followed the district protocols, which were an adaptation of national and Indiana RTI frameworks (Table 5). In addition, teaching and instructional support staff allocations were similar (Appendix C). Noteworthy differences to the district's RTI program include:

- Tier 2 groups \((n = 3 - 7)\) were conducted outside of the classroom by interventionists, instead of the traditional classroom approach where the teacher fits the intervention into the classroom structures. This provided consistent implementation of the intervention, standardized intervention programming, and consistent training for the interventionist. However, the interventionists were not all certified educators.

- The RTI team conducted 4-5 meetings per year to discuss the student data and selected the treatment group based on whole grade-level data instead of the traditional classroom data. Once again, by providing a grade-level analysis, the
students with the most needs across the grade-level were provided the treatment.

Overall the environment was homogeneous and provided six sets to study. In essence, parallel ‘experiments’ were underway within the district providing opportunity for analysis of non-equivalent nested groups with pre/post-test longitudinal data.

Table 5

| Distinctions in RTI, both Indiana and the study district, and Traditional Approaches |
|----------------------------------|---------------------------------|---------------------------------|
| **Traditional Approach**            | **Indiana Response to Instruction** | **Study District RTI** |
| Students who are unsuccessful with the core curriculum are referred to and often placed in special programs that include instruction in pull-out classrooms. | When students are unsuccessful in the core curriculum or have demonstrated proficiency, their teachers review the student data and adjust instructional practices including intensity and duration to meet the students’ goals. | T1: All students - differentiated instruction, scaffolding to support learning, flexible grouping  
T2: Small group (n = 3 to 6), beyond core instruction, on weakest skill area(s), 30 minutes daily, 6-9 weeks  
T3: Additional time (30 minutes daily) beyond T2, small group (n = 1 to 3), on weakest skill area(s), 6-9 weeks  
Minimum of two T2 cycles and one T3 cycle prior to referral for special education |
| Teachers primarily use the same instructional methods for all students. | Teachers modify, support, and extend instructional practices based on individual student goals/needs. | Scaffold instruction to support release of instructional responsibility (I do, We do, You do model)  
Differentiation  
Small group work within the classroom  
Additional student support provided to teachers by instructional aides |
| Student growth and needs are determined sporadically or over extended periods (e.g., end-of-semester exams; ISTEP+). | Student growth and goals are measured weekly or biweekly through progress monitoring. | Benchmark assessments given 3 - 5 times/year  
Progress monitoring of students in T2 and T3  
Formative and summative assessments within instruction |
| Data review and instructional decisions are made in isolation by individual teachers. | Data analysis and instructional decisions are made through discussions in school teams. | Analysis by  
• Individual teachers (ongoing)  
• RTI data meetings by peer team (4-5/year at the end of each cycle of RTI)  
• School Improvement Team (periodic)  
• Professional learning communities  
• AdvancEd Accreditation Team (periodic) |
| Special education teachers hold the main responsibility for students who are not succeeding in the core curriculum. | All teachers are responsible for meeting the goals of all students with program specific teachers (e.g., special education, Title I, ELL) being integrated with classroom teachers. | All teachers are responsible for meeting the goals of all students. Inclusionary model of service for most students  
• Co-teaching or collaboration of grade-level teacher with special education teachers or IA  
• Extra support through intervention time |

*Note*: The RTI framework is a major shift from the traditional education system, which typically employed a pull-out model for special education students. The table summarizes the Indiana RTI framework and provides details on the research district’s implementation of the RTI framework.
Treatment/Intervention Protocols

The intervention students were placed into homogeneous groups based on each participant's reading weakness using the benchmark data and teacher input. The groups of 4-7 participants were instructed daily for 30 minutes beyond the 90-minute core instruction. Participants' growth was progress monitored every two weeks. The intervention cycle ran between six and nine weeks. Following the intervention cycle, each student’s growth was scrutinized based on the progress monitoring data and program adjustments were made as deemed necessary. Pending achievement outcomes, a student could: be moved back a tier, remain in the tier with modifications to the intervention plan, advance to Tier 3, or receive a referral for special education. In contrast, the control group attended instruction in other academic areas during this time and did not receive additional reading instruction.

Intervention curriculum used an array of scientifically-based programs. Selection of the programs for each group was based on the area of weakness in reading, the student's age, and the level of need of the student. The specific program was not associated with the data in this study, but may be a point of further investigation for the research district.

The interventions were conducted by interventionists who were either certified teachers or instructional aides under the guidance of certified teachers. The interventionists were provided training and ongoing support in accordance with each program's prescribed requirements. Instructor logs were maintained to monitor the interventions and observations were conducted periodically by the reading coach to maintain program fidelity. In research, the fidelity to the implementation is integral as it helps substantiate the relationship to between student outcomes and the program effectiveness (O' Donnell, 2008).
Focus Group Participants

The selection of the representative focus group was driven by two key principles of qualitative research: (1) the sample is purposively selected and (2) the selection of participants is directed by the emerging analysis of the data (Richards & Morse, 2006). The criteria for selection included (1) knowledge of the district's RTI protocols; (2) exposure to the full implementation; (3) involvement with the primary grade students throughout the study timeline; and (4) willingness to participate in the study. To meet the previously mentioned criteria, all participants were employed from 2008 through 2014 and worked with primary grades. By working with primary grade students, the participant was familiar with the district's protocols, the curriculum, and the student population. This familiarity provided a level of expertise necessary to enhance the research project quantitative data analysis.

The focus groups included a teacher focus group per school and an administrator focus group. All focus group members volunteered to participate and signed the consent form (Appendix D). Table 6 presents the layout for each of the focus groups. To sustain conversation, research indicates a minimum of 6-12 participants would be preferred (Johnson & Christenson, 2010; Morgan, 1997). To maintain appropriate group size, if more than one staff member met the above criteria, seniority was used for the selection. The justification for seniority was that the more experienced members would provide a better historical perspective. The administrators were separated from the teacher to allow participants to anonymously reflect on the data without the oversight of their job performance evaluator. All focus group participants will remain confidential as per the agreement within the consent form. In this regard, minimal demographic statistics will be presented on the focus groups.
Table 6
Focus Group Participants

<table>
<thead>
<tr>
<th>Focus Group #</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1: Administrators</td>
<td>3 building administrators, 1 per school</td>
</tr>
<tr>
<td></td>
<td>School psychologist, assigned to the district</td>
</tr>
<tr>
<td></td>
<td>Special Education Director, assigned to the district</td>
</tr>
<tr>
<td>Group 2-4: Teachers</td>
<td>4 Grade-level teachers (K-3), preferably 1/grade (one per school)</td>
</tr>
<tr>
<td></td>
<td>1 Special Education teacher</td>
</tr>
<tr>
<td></td>
<td>1 Reading Coach</td>
</tr>
</tbody>
</table>

Note: Neither Reading Coach was employed in this position throughout the study. They had previously been a teacher and were familiar with the process. They were included due to their intimate relationship with the process.

Instrumentation

All kindergarten through second grade students within the district were benchmarked three times per year utilizing mClass Reading 3D™ (Table 7). In addition, Scholastic Reading Inventory™ (SRI), Test of Early Reading Ability™ (TERA-3), and Indiana state assessments were delivered to grade-appropriate levels. Each assessment permitted a standard score analysis. A standard score is a measure of relative position that expresses how far a raw score is from the reference point (Gay, Mills, Airasian, 2006). One method of expression would be a score’s standard deviation from the mean. Standard scores allow comparison of a student's performance to that of other students or another test score.

Table 7
District Comprehensive Assessment Plan

<table>
<thead>
<tr>
<th>Assessment Category</th>
<th>Assessment Schedule</th>
<th>Students Assessed</th>
<th>Assessment Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark (or Universal</td>
<td>August</td>
<td>All K students</td>
<td>TERA-3</td>
</tr>
<tr>
<td>Screener)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benchmark</td>
<td>August</td>
<td>All K-2 students</td>
<td>mClass Reading 3D,</td>
</tr>
<tr>
<td></td>
<td>January</td>
<td></td>
<td>which includes DIBELS</td>
</tr>
<tr>
<td></td>
<td>May</td>
<td></td>
<td>and TRC</td>
</tr>
<tr>
<td>Benchmark</td>
<td>August</td>
<td>All Grade 2-6 students</td>
<td>Scholastic Reading Inventory (SRI)</td>
</tr>
<tr>
<td></td>
<td>January</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>April/May</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Progress Monitoring (PM)</td>
<td>Determined by level</td>
<td>All RTI T2 and T3 students</td>
<td>DIBELS (K-6)</td>
</tr>
<tr>
<td></td>
<td>of intervention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnostic</td>
<td>As needed</td>
<td>Students moving from T2 to T3</td>
<td>Various norm-referenced assessments</td>
</tr>
<tr>
<td>Summative</td>
<td>Spring</td>
<td>All 3rd grade students</td>
<td>ISTEP+ ELA</td>
</tr>
<tr>
<td>Summative</td>
<td>Spring</td>
<td>All 3rd grade students</td>
<td>IREAD-3</td>
</tr>
</tbody>
</table>

Note. The assessment calendar was similar each year of the study.
TERA-3

The third edition of the Test of Early Reading Ability (TERA-3) is a measure of reading ability of children ages 3-6 through 8-6 with strong validity and reliability (Reid, Hresko, & Hammill, 2001). It assesses mastery of early developing reading skills. The individually administered, norm-referenced assessment provides standard scores for three subtests (e.g. alphabet knowledge, conventions of print, and meaning) ($M = 10$, $SD = 3$), a Reading Quotient, percentiles, age equivalent score, and a grade equivalent score based on the normative sample. The normative information is stratified by age (i.e. 3 to 6 month intervals) and showed an absence of gender, racial, disability, and ethnic bias. According to Rathvon (2004), for standardization purposes the stratification data (e.g. 3 month intervals) for ages 5 and 6 fell below acceptable levels. Based on 1999-2000 normative sample ($n = 875$) of the U.S. population, reliability is consistently high (e.g. all but 2 of the 32 coefficients reported were near or above .90) (Rathvon, 2004; Reid, Hresko, & Hammill, 2001). The Reading Quotient is an indicator of the potential to reach an independent reading level and a child with a low score may have difficulties achieving this level (Reid, Hresko, & Hammill, 2001). TERA-3 is not recommended as a stand-alone assessment tool as it is not comprehensive enough to determine the basis of the poor performance and therefore recommendations are to coordinate the results with other assessments (i.e. DIBELS) to provide a more in-depth analysis (Reid, Hresko, & Hammill, 2001).

mClass Reading 3D

A commercially-available screening measure, mClass Reading 3D provides a balanced approach for assessing the foundational skills of reading. The assessment is comprised of two components: (1) Text Reading and Comprehension (TRC) and (2) Dynamic Indicators of Basic
Early Literacy Skills (DIBELS)\(^6\). Achievement standards for DIBELS and TRC were defined by the University of Oregon and Amplify\(^7\), respectively. When paired together, DIBELS and TRC have demonstrated an internal and external predictive validity, as demonstrated by the correlation found between ORF, TRC, and Indiana's ISTEP+ summative test \((r = 0.66 \text{ to } 0.67)\) and between ORF, TRC, and North Carolina's EOG summative test \((r = 0.69 \text{ to } 0.72)\).

Furthermore DIBELS and TRC have correctly classified 84.22\% and 84.81\% of students' pass rate on the end of third grade ISTEP+ and EOG respectively (Amplify, 2014).

The assessment is administered individually by a certified educator with specific training in the delivery of the assessment and with the aid of a digital device. To increase inter-rater reliability, all test administrators were trained and refreshed periodically on the assessment procedures. Further, the digital device aids in inter-rater reliability by automating instructions and timing of each assessment.

**DIBELS.** DIBELS is designed to assess growth and development of the essential early literacy skills established by the National Reading Panel and the National Research Council for students from kindergarten through sixth grade (University of Oregon, 2011). As a series of short, one-minute fluency assessments, DIBELS is based on measurement procedures of curriculum-based measures, commonly referred to as CBM (Deno & Mirkin, 1977; Deno & Fuchs, 1987). DIBELS provides criterion-referenced target scores (i.e. individual test scores and

\(^6\) During the study window, the DIBELS assessments altered slightly. Initially, DIBELS 6th Edition was used. In the fall of 2010, DIBELS Next began. Both consist of short, standardized repeated measures for assessing the acquisition of the foundational early literacy skills. While direct comparisons of some aspects are not reliable, both editions can serve as screening or progress monitoring tools for the same foundational skills. In addition, the composite score cuts was adjusted to a new standard for successful reading achievement.

\(^7\) Formerly Wireless Generation.
a composite score) that change across time, which permit the comparison of a student's performance to other students based on national norms for adequate reading progress. Several scores are produced including a raw score for each sub-test, a composite score, and a developmental benchmark score.

DIBELS has been compared to various standardized assessments\(^8\) and has been found to be reliable and valid. DIBELS measures showed moderate reliability individually (e.g. single-form \( r = .72 \) to \(.89\)); repeated measure aggregate results between \(.91\) to \(.96\); a median concurrent validity of DIBELS measures with Woodcock-Johnson Broad Reading Cluster\(^9\) between \(.36\) and \(.75\); and predictive validity (i.e. at least \(12\) months) for oral reading fluency between \( r = .38\) to \(.69\) (Good et al., 2004; Gushta & Wang, 2013).

Finally using a latent variable approach, DIBELS composite scores showed a composite reliability in the range of \(.74\) to \(.91\) (e.g. acceptable reliability \(≥ 0.7\)) (Amplify, 2014). DIBELS has been used as a screening and progress monitoring tool.

However, two common criticisms of DIBELS are its inadequacy to measure comprehension and its use of a stepping stone model that suggests mastery of certain skills are sequential in nature (Riedel & Samuels, 2007). The remedy in this study for the first concern would be mClass TRC.

---

\(^8\) Growth data from DIBELS 6th ed. and DIBELS Next have been compared to GRADE (Group Reading Assessment and Diagnostic Evaluation), 4th Grade Reading Passage NAEP (National Assessment of Education Progress), and other standardized assessments (Amplify, 2014).

\(^9\) The Woodcock-Johnson Broad Reading Cluster is a sub-test of the Woodcock-Johnson Test of Achievement (WJ III). WJ III is a normative assessment that provides scores based on age norms and grade norms.
TRC. Text Reading and Comprehension (TRC) is based on an assessment technique called running records\textsuperscript{10} and uses leveled readers to determine the student's instructional reading level (i.e. a feasible, but challenging reading level). After reading the book, the student answers follow-up questions including oral comprehension, written comprehension, recall, and retelling. Using a combination of the student's accuracy and comprehension, an instructional reading level is established based on Fountas and Pinnell\textsuperscript{11} levels.

TRC has been found to be precise showing a marginal reliability median of 0.86 across various K-3 student populations (Amplify, 2014). Inter-rater reliability was found to have a high degree of consistency with a high average Spearman rank-order correlation ($\rho = 0.73$) (Amplify, 2014). In addition, TRC has been shown to be valid with strong predictive power. A two-year study showed a predictive accuracy between 67.8% and 86.5% for students (K-3; $n = 24,509$) with the low percentage underestimating student achievement in 3rd grade (Amplify, 2014).

**Scholastic Reading Inventory**

Scholastic Reading Inventory (SRI) is designed to measure how well readers comprehend text. It focuses on core reading comprehension skills (i.e. referring to details in the text, drawing conclusions, making comparisons, and generalization) and assigns a Lexile\textsuperscript{12} measure (BR to

---

\textsuperscript{10} Running Records was first developed by Marie Clay in 1993. The observational survey provides data on a student's reading ability. From Amplify, 2014.

\textsuperscript{11} Fountas and Pinnell (F&P) Text-Level Gradient™ was first established in 1996 by Irene Fountas and Gay Su Pinnell. F&P are often referred to as guided reading levels. The research behind F&P is rooted in Marie Clay's work on text complexity. F&P evaluates text characteristics (e.g. text structure, vocabulary, literacy features) to determine a text complexity level. From Fountas & Pinnell, 2011.

\textsuperscript{12} A Lexile measure provides a reading level for an individual or the difficulty of a text. The intent of a Lexile is to provide guidance to match text to the reader. A Lexile is not a grade equivalent. Instead of measuring a person's
RTI: DIFFERENCES IN READING ACHIEVEMENT

1700L) to the student. In addition to the Lexile, SRI generates criterion- and norm-referenced results, including percentile rank, stanine, normal curve equivalent (NCE), grade-level equivalent, and a performance standard. The computer-based assessment may be administered to students in 1st through 12th grade. In this research, SRI was administered by the teacher to students three times per year.

SRI has been found to be both reliable and valid. In a field test of students in grade 3, 4, 5, and 7 (N = 879), SRI was found to measure a construct similar to that of other measures of reading comprehension (e.g. correlation value of 0.73 and 0.67 with the 3rd and 4th grade North Carolina EOG respectively) (Scholastic, 2007). Several studies (Scholastic, 2007) have demonstrated criterion-related validity due to the positive relationship between SRI scores and enrollment in a reading intervention program as well as solid construct validity (r = .789 to .824 in relation to the Stanford Achievement Test, 9th or 10th edition). Finally, the predictive accuracy of the spring SRI scores for students grade 2-10 (N = 3465) in relation to achievement on the California CST English-Language Arts assessment was found to be strong (Knutson, 2008).

**Indiana State Assessments: ISTEP+ and IREAD-3**

Indiana Statewide Testing for Educational Progress Plus (ISTEP+) is designed to measure student achievement according to the Indiana Academic Standards that were adopted by the Indiana State Board of Education. The results of ISTEP+ provide information on the extent an individual student has mastered the Indiana Academic Standards, the extent a group of students are attaining mastery on the Academic Standards, and finally where an individual reading level against a norming sample, it represents a persons' ability level on a developmental scale of reading ability. However, typical Lexile ranges have been determined. From MetaMetrics (2014).
student stands relative to the Indiana State Board of Education’s three performance levels (i.e. Did not pass, Pass, and Pass+). The Indiana Reading Evaluation and Determination (IREAD-3) assessment measures the reading skills for a student based on the foundational reading standards through grade 3 relative to two performance levels (i.e. Did not Pass and Pass). A student who passes demonstrates proficient understanding when reading and responding to grade-level literacy and informational text. As with the previously mentioned assessments, the ISTEP+ and IREAD-3 tests were administered in a standard fashion as per the Indiana Department of Education (IDOE) protocols (see IDOE Assessment website for details, www.doe.in.gov/assessment). Both assessments are criterion-referenced tests and do not provide norm-reference information based on a comparison to students from across the nation. The criterion-referenced scores only indicate where the student stands relative to the cut score based on Indiana Academic Standards.

The criterion-referenced score for each is developed based on Item Response Theory (IRT). IRT is a family of statistical models. To characterize the items and generate the student scale scores, one model was used for the multiple-choice items and another model was used for the open-ended items. Both models determined the difficulty of each item and how well each item distinguished students who do and do not have the skills being tested by the item. Scale scores used pattern scoring which considers question difficulty and therefore does not correlate directly to a number of questions correct, or a raw score. According to the Guide to Test Interpretation: Grade 3-8, raw scores should not be used to make comparisons between performances on different tests. The scale score represents an ability score. These three-digit, equal-interval scores are expressed on a vertical scale by content area.
In regards to item-level reliability, intra-class correlations for all items on the English/Language Arts (ELA) assessment fell between 0.74 to 0.97 with a mean of 0.89 and the kappa statistics fell between 0.72 to 0.96 with a mean of 0.78 (IDOE, 2009). Both fall within acceptable limits. The differential item functioning (DIF), which provided a view of the systemic errors by subgroup performance compared to the whole group performance, indicated about 7% of the operational items exhibited gender or ethnic DIF at the moderate and large levels (IDOE, 2009). ISTEP+ test-level reliability was evaluated by Cronbach's $\alpha$ to be 8.1 (i.e. reliability coefficients at or above 8.0 are considered of sound reliability) and the proficiency-level reliability was found to be most accurate close to the cut score (IDOE, 2009). According to the IDOE report (2009), the validity of ISTEP+ was found to be strong (e.g. KMO for the total group ranged from 0.96 to 0.98 and for subgroups from 0.90 to 0.97 with 0.90 or greater is considered strong; MA 74-96% for the total group and subgroups indicating that each component has one major factor present; and the RMSEA values were all below 0.04 for each subgroup thus considered a good fit). Finally, as for IREAD-3 no specific statistical data was available at the time of this study; however, we can assume since it is a state-developed assessment it is reliable and valid.

**Qualitative Instrument: Guiding Interview Questions**

The researcher interviewed each focus group using a semi-structured format (Ledeman, 1990). The semi-structured questions were based on the research questions and the analysis of the quantitative data. Semi-structured interviews further probed the phenomenon and provided insight not available from the statistical data. The focus group questions were designed to guide the discussion without hindering the educators' input and reflection on the data or reflections on the outcome of the RTI framework on student achievement and growth (Appendix E).
Data Collection

Following the sequential nature of the study, the data collection occurred in two phases. The first phase included the collection of the archival student assessment data along with the demographic student data. Subsequent to the analysis of this data, the qualitative phase commenced. In the qualitative phase focus group data was collected to further analyze the numeric outcomes of the quantitative phase.

Quantitative Data Collection

As mentioned previously, Figure 7 and 8 represent the nested groups within this longitudinal study. The intervention groups and non-intervention groups are represented by $n$ and $n'$, respectively, with subscript notations as to the study cohort. Four years of archival data were gathered for each cohort. For example, data were collected for cohort $n_1$ from 2008-2009 to 2011-2012. Assessments were conducted during both pre- and post-treatment cycles. For the sake of convention, the benchmark assessments (i.e. TERA-3 and SRI) will be discussed throughout the study as beginning-of-year (BOY), middle-of-year (MOY), and end-of-year (EOY). In addition, the third grade IREAD-3 reading assessment and ISTEP+ assessment were utilized as culminating achievement data points, or for the sake of consistency, as an EOY.

The district studied in this research used a mixture of previous (where applicable) and current year's data to rank order each grade level within a school. This process aligns with the current thinking of the field (Fuchs & Fuchs, 2006). According to Fuchs and Fuchs (2006), benchmark assessments may be norm-referenced measures or criterion-referenced measures. Criterion, such as scores below the 25th percentile for norm-referenced measure or below a cut score for criterion-referenced measure, is set to ascertain risk and make a placement. One difference between this study group and Fuchs' description would be the use of percentile for
placement or cut score within a tier. Due to staffing constraints and utilizing national RTI data trends for typical populations within Tier 2, approximately 20% of the students within a grade-level received Tier 2 services rather than a hard fast percentile or cut score.

Data collection for student benchmarks and progress monitoring was conducted by a trained, certified educator as a component of the RTI framework. The researchers compiled the archival data from Pivot®, a data warehouse.

**Qualitative Data Collection**

With the permission of the focus group participants, each focus group discussion was recorded for analysis and coding. Beyond the written acknowledgement from the consent form, each focus group was reminded that the discussion would be recorded for later coding. The recordings were stored on a secure hard drive.

At the onset of the focus group discussion, the data and the statistical analysis from the first phase of the quantitative study were shared with the group prior to the discussion. Explanations of the meaning of the various statistics were provided along with the data. Interpretation of the results was avoided. The participants were asked if further explanation of the study data or statistics was necessary.

Next, the focus group was asked open-ended questions to guide the discussion. Each question focused on an aspect of the research questions. When necessary, clarifying, probing, or inquiring questions were asked to better understand, focus thinking, or broaden thinking based on the participant's comment. As the focus group facilitator, I used coaching strategies (e.g. acknowledging, pausing, paraphrasing, and questioning) based on Costa's Cognitive Coaching model. When reflecting on a complex intellectual activity such as teaching, the use of these coaching skills has been shown to build trust and rapport, capitalize upon teachers' cognitive
processes, and enhance pattern finding and analysis (Costa & Garmston, 2005). The intent of these strategies was to deepen the participants' reflection on the implementation of RTI.

For this research, Initial Coding was utilized due to the appropriateness for almost any study and the beginning qualitative researcher's skill level. Initial coding requires time to digest, provides analytical leads, is cyclic in nature, and may alert the researcher to more data needs to support results (Saldana, 2009). The steps employed for the initial coding were

1. Pre-coding responses that rises to the top upon the initial review
   a. During focus group conversation initial notes were taken based on keywords that stood out based on the research questions.
   b. All focus group recordings were reviewed to gain a sense of the whole.

2. Decoding and encoding for the real meaning
   a. Each focus group's audio was reviewed and labeled with content descriptors, data examples, and the role of the speaker (Saldana, 2009).
   b. A code ledger was created and similar codes were clustered. The ledger columns included topics, unique topics, and leftovers (Roberts, 2010).

3. Member checks
   a. Following the coding by the researcher, member checks (Shenton, 2004) were conducted by providing the notes to a focus group member to verify the content matched the discussion. Completing a member check is the single most important provision to bolstering credibility of a focus group's results (Shenton, 2004).

4. Analysis of the codes for meaning
   a. Patterns and unique insights among respondents
   b. Patterns with the quantitative results
(c) Frequency of an expression

(d) Expression of implied causation or outcome that may support research findings or provide potential research questions for further study.

An important aspect to recall with coding is that it is done through the eyes of the researcher with the goal of organizing evidence into methodical categories and linking ideas together, specifically for this research to the quantitative results. As coding is a craft, this researcher completed basic codes to coordinate with the inferential statistics with an attempt to enhance the depth of understanding of the data. My dissertation chair also reviewed my coding.

**Data Analysis**

I am hypothesizing that if a group of students with higher benchmark scores (i.e. control group) were to participate in high-quality instruction, the group's scores would remain above or possibly increase the gap between themselves and the lower group (i.e. treatment group). This being said, the hypothesis would be H1: the treatment group will grow at an equal or slightly lesser rate than the control group. This outcome would be based on achievement expectancy results that represent a typical pattern of growth. Trajectory analysis (Bakerson & Cothberg, 2006) typically predicts that lower achieving students will fall further behind academically. Based on these facts, a positive outcome for the RTI framework would be data showing the treatment group participants at least maintained the same amount of growth as their peers who began at higher levels.

**Quantitative Data Analysis**

The measure of the long range outcome on reading achievement for students who participated in RTI Tier 2 intervention as compared to students who scored slightly above the intervention group was measured. Three or four years of archival student achievement data were
analyzed and data concerning special education identification were monitored. The pre-test and the effect of the intervention were the independent variables, while the post-test results were the dependent variable. Additional variables that permitted the study of the nested groups included the school, timing of the intervention, entry into school relative to the year of the study, and reading area of most need. Within- and between-group analyses were conducted using descriptive and inferential statistical tests. In addition to reviewing performance of the cohort groups, attention was given to the classifications of gender, socio-economic status, and student age relative to grade level peers. All statistical tests were conducted using IBM's SPSS predictive analytics software.

Specific measures used to analyze the reading assessment data included traditional descriptive statistics (i.e. mean, standard deviation, frequency) to present the data in a manageable form thus providing a concise picture of the data (Trochim, 2006). Means and standard deviations were computed for each cohort group: intervention versus non-intervention, School A, B, and C, and initial year of intervention (i.e. kindergarten 2008-2009 or first grade 2010-2011). In addition to descriptive data, I reported frequencies for gender (i.e. male vs. female), socio-economic status (i.e. free/reduced lunch vs. paid lunch), and school entry age data (i.e. birth month).

Beyond descriptive statistics, inferential statistics (i.e. ANOVA, t-test and chi square) were used. The data were analyzed using t-tests to determine if a significant difference existed between the means of any of factors such as gender or socio-economic level. To ascertain if an event (e.g. identification for an SLD) occurred more frequently in one group than another, chi square t-tests were conducted to compare group frequencies. An analysis of variance (ANOVA) was run to conclude if a significant difference exists between the intervention and non-
intervention groups based on variables such as the school, year of implementation, and the interaction among factors. Prior to running test, the equality of variance between groups was examined using Levene's Test to confirm homogeneity of variances.

**Qualitative Data Analysis**

The focus group interviews were recorded and later analyzed using initial coding of the open-ended question responses. Initial coding uses descriptive coding to break down the qualitative data in order to summarize and find patterns in the participants' feedback. According to Saldana (2009), "a code in qualitative inquiry is most often a word or short phrase that symbolically assigns a summative, salient, essence-capturing, and/or evocative attribute for a portion of language-based or visual data" (p. 3). Basically, a code acts like a book title; it captures the heart of the thought.

The qualitative data were analyzed using initial coding as detailed in the data collection section along with conducting member checks for trustworthiness. This method began with a sense of the whole and followed with detailed analysis of each recording using the codes. Through multiple reviews of the recordings by this researcher, the interviews were analyzed to establish commonalities and differences in the participant feedback. This form of coding allowed trends in the reflective thought of the participants to emerge and a depth to the quantitative results. Segments of the recording were transcribed to share the voices of the focus group participants in this research.

Interpretation allowed for the development of the underlying story within the data. Through the initial cataloguing, the codes (32) were collapsed into nine (9) broad categories of information (Creswell & Plano Clark, 2004). This analysis was used to augment the quantitative results with phenomena perspective of a sample of stakeholders. The codes, along with anecdotal evidence, will be partnered with the quantitative results.
Researcher's Perspective

To instill confidence in the research, the researcher must provide full disclosure to all methods used to encourage transparency and replication (O’Leary, 2004). To provide transparency, I am a district administrator within the study school district and have provided organizational and leadership support of the RTI protocols throughout its implementation. In addition, as the researcher, I conducted and analyzed the focus group interviews. Since some of the data gathered was qualitative in nature, I had to be sensitive to any personal bias to avoid influencing the outcome of the study (Merriam, 1998). When collecting, analyzing, and reporting the qualitative data, every effort was made to be unbiased and remain neutral. No previous opinion on the specific outcomes of the RTI framework was formed. It would be my belief that the RTI process has helped students through its proactive intervention approach; however, neither the district personnel nor I have analyzed the program's impact. The transparency within the study allows other researchers to replicate part or all of the study aspects.

Creswell (1998) said an important step in the research process was to “establish rapport so that participants will provide good data” (p. 110). One advantage that I brought to the research is that I had already established rapport with the staff. The staff felt comfortable speaking with me and they understood the underlying purpose of the study was improvement.

Pseudonyms were given to participants (e.g. students and focus group participants, as well as the schools) to ensure their confidentiality. Consent forms were secured following all necessary ethical standards adhered to in the study. In addition, the data analysis was accessible to school administrators.
Limitations of the Study

As in all research this study included limitations. This study was limited by the selection of the instructional tools (e.g. benchmark assessments, progress monitoring tools, and intervention programs) that were selected by the school district; sample size variants across the study groups; and the interventionists' appropriate facilitation of the intervention. The use of multiple cohorts across three schools over multiple years decreases the limitations presented. In addition, any irregularities within the data were investigated.

The generalizations of findings from this study may be affected by the fact that this study was completed within a school district and with a relatively limited sample. One of the limitations is the generalizability due to the study of a single district that is relatively homogeneous.

Summary

Chapter three explained the research methodd including details regarding the RTI framework, benchmark assessments, intervention tools, and the participants. Since the research was conducted in a complex social environment and not in a sterile laboratory setting, a combination of quantitative and qualitative research was implemented. The mixed-methods approach has been shown to be beneficial when investigating human behavior (Patton, 1990.) Green et al., (1989) highlighted the purposes including triangulation, complementary, and expansion. Mixed-methods builds on the strength of each method and minimizes the weaknesses to increase validity and reliability. This study used an explanatory sequential mixed-methods design, specifically a pre/post-test longitudinal, non-equivalent group design (NEGD). The quasi-experimental design was structured with the follow-up by focus group discussions. The research utilized both descriptive and inferential statistics to measure the effectiveness of a
variety of intervention strategies. The mixed-methods approach allowed the statistical analysis to be enhanced through qualitative data collection. The qualitative data from the focus group was triangulated with the statistical data and data on student placement in special education for a Specific Learning Disability (SLD).

The outcomes were measured by student academic growth through third grade. Using various assessments, kindergarten and first-grade students demonstrating lower reading skills received Tier 2 intervention. The intervention group's growth was compared to the growth of students who fell just outside the treatment criteria and did not receive Tier 2 intervention. The qualitative phase was conducted as a follow-up to the quantitative results to deepen the understanding of said results. Focus groups from each school were interviewed using a semi-structured format to explore the outcomes of the RTI program on the primary grade students. Chapter 4 will summarize the results of this study.
CHAPTER 4
RESULTS

This chapter outlines the analysis of the research questions regarding the long-range outcome on reading achievement and special education placement following the implementation of the RTI process in one district. The mixed-methods research was utilized for the purpose of triangulating the data through quantitative and qualitative data. A comparison of the assessment results from the students who received intervention and the students slightly above the intervention groups who did not receive the treatment will be presented.

First, this chapter will present the demographic data for six cohorts. The cohorts include students entering kindergarten in fall 2008, 2009, and 2010 with low reading scores relative to peers during either kindergarten or first grade. Each cohort is comprised of intervention and non-intervention sub-groups that include students from three elementary schools. The use of multiple cohorts bolsters the results by increasing the sample size. The nested group analysis decreases the study limitations and provides completeness. Next, this chapter will present the results for each research question in terms of descriptive and inferential statistics. The final aspect of each research question will be the qualitative data analysis of the focus group interviews.

Purpose of the Study

The purpose of this study was to examine the outcomes of the RTI process, such as student achievement in reading and/or placement in special education. In addition, the data provided evidence to the district to guide program decisions. The study investigated whether students identified as lower achieving readers in kindergarten or first grade, who received
deliberate intervention in reading within the RTI program, had higher levels of reading achievement by the end of third grade than students who did not receive the RTI reading intervention. The independent variables in this study were inclusion in the intervention or the non-intervention group, gender, socio-economic status, implementation year, cohort, school, and entry grade. Academic achievement in reading was the dependent variable and was measured by reading assessment results and placement in special education.

**Research Questions**

Given this is a mixed-methods study, this research encompassed a primary or central question with sub-questions to guide the exploration of the data collection (Creswell, 2003). This overarching question guided this study:

1. What are the long-range (3 or 4 years) differences on reading achievement for students who participated in RTI Tier 2 reading intervention as compared to students who scored slightly above the intervention group and did not receive treatment?

The following sub-questions guided the research and data analysis for this study:

2. What differences occur among the RTI Tier 2 groups as compared to those students just outside the intervention group in regards to special education identification by the end of third grade?

3. What differences are seen in students' reading achievement among and within the various nested groups (e.g. school, implementation year, achievement measures)?

4. From the perspective of the focus group, what understandings result from the quantitative data analysis of the academic outcomes between and among the nested groups?
**Participant Demographics**

The participant demographics are shared in the next section. To provide a general overview of the student population in the study, the student cohort demographic data are introduced. In this same section, the student cohort data are disaggregated by the various nested groups to present a basis for the third research question. Subsequently, the focus group demographics will be provided along with the focus group research question.

**Student Cohort Demographics**

The study sample, combining the six cohort groups based on fall benchmark data (i.e. TERA-3, teacher input, and mClass instructional level) from kindergarten or first grade into intervention and non-intervention categories, is presented in Table 8. The initial combined cohort group included 1109 students. Omitted from the study were 155 students (14%) due to a lack of pre- or post-test data. It may be assumed that these students were not enrolled in the district from the beginning of kindergarten to the end of third grade. According to the National Center for Education Evaluation and Regional Assistance (2009), typically between 10-20% of a sample has missing data. The method selected to handle the missing data was the listwise deletion (i.e. removing the cases with missing assessment data). Problematic issues with this method included reduction in the sample size causing a reduction in statistical power and potential for bias associated to one or more factors in the study (Peugh & Enders, 2004). In Rubin's theoretical framework for missing data, one of the mechanisms for why data are missing is missing completely at random (MCAR) mechanism. Peugh & Enders (2004) provide several examples of MCAR in longitudinal achievement studies including absence from school for the assessment and permanently leaving the student due to relocation. The question then arises are these reasons unrelated to other factors such as socio-economic status. In this study, the
demographics with and without the omitted cases prove to be very similar (see Table 9 note). Therefore according to Rubin's framework, the remaining participants represent a random sample of the complete data set (Peugh & Enders, 2004).

Beyond the demographic aspects that are typically associated with at-risk students, a newer terminology for the more fragile learners in the education system would be school dependent student (Nichols, n.d., Urban School Services, n.d.). School dependent students are those who succeed or fail depending on the school they attend. Social dependency status (SDS) includes some of the basic needs historically met at home, such as physical needs (e.g. nutritional meals and health care), emotional needs (e.g. security, self-worth, and structure), intellectual needs (e.g. cognitive skills development and homework help), social needs, cultural needs, and language needs (Urban School Services, n.d.). SES and SDS are not interchangeable, but may be analogous to each other. Shifting from at-risk to school dependent students changes the focus to variables that educators can control (Nichols, n.d.). Variables such as quality instruction, interventions, and opportunities to become part of the school community are items that schools can support and focus on improving. For this study, SDS was not determined for the students omitted from the study due to incomplete data. Only a consideration of the SES was reviewed, which does not align completely with SDS. I would note that during a student's time at the district, the three variables listed above are thoughtfully implemented and that progress monitoring results could be analyzed in a further study to explore school dependency status. For this study, the use of listwise method to remove the students without complete assessment data does not seem to pose an issue; however, does impose limitations.
Since RTI intervention is an annual determination, the data were next filtered for duplicate students. An example of a duplicate student would be a student who was provided intervention during kindergarten in 2008 (cohort 1) and then again during first grade in 2009 (cohort 4). To yield a single entry for each student, the duplicate for a specific student was omitted based on the following protocols. First, students were removed from a non-intervention cohort if they were identified in an intervention cohort. Second, if a student was selected for two intervention or non-intervention cohorts, a single entry was entered in the earliest chronological cohort. This removed 315 (28%) student entries. After filtering the data for duplicate students and exclusion of students with incomplete data, the final study sample included 639 unique students (58%) in the initial study sample. While removal of duplication is different than omitting of students due to the lack of complete data, the frequency distribution supports the demographic similarities of the data set prior and post duplicate removal. An alternative approach considered for the duplicates had been analysis by individual cohort only. When considering the sample population, this did not seem to be the best option. The limitations of these decisions will be discussed in chapter 5.

All students within the final study sample \((n = 639)\) experienced similar core reading instruction based on the district's adopted curriculum between kindergarten and third grade. Out of the study sample, 327 (51%) students were included in the intervention group and 312 (49%) were not in intervention during kindergarten or first grade. Holding the core curriculum consistent allows for exploration of the differences between the intervention and non-intervention groups.
### Table 8: Cohort Frequency

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Intervention n</th>
<th>%</th>
<th>Non-Intervention n</th>
<th>%</th>
<th>Total n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008-2009 Kindergarten</td>
<td>69 (44)</td>
<td>11</td>
<td>57</td>
<td>9</td>
<td>129</td>
<td>20</td>
</tr>
<tr>
<td>2009-2010 Kindergarten</td>
<td>70 (35)</td>
<td>11</td>
<td>73</td>
<td>11</td>
<td>143</td>
<td>22</td>
</tr>
<tr>
<td>2010-2011 Kindergarten</td>
<td>70 (38)</td>
<td>11</td>
<td>64</td>
<td>10</td>
<td>134</td>
<td>21</td>
</tr>
<tr>
<td>2009-2010 1st Grade</td>
<td>53 (6)</td>
<td>8</td>
<td>46</td>
<td>7</td>
<td>99</td>
<td>15</td>
</tr>
<tr>
<td>2010-2011 1st Grade</td>
<td>30 (6)</td>
<td>5</td>
<td>39</td>
<td>6</td>
<td>69</td>
<td>11</td>
</tr>
<tr>
<td>2011-2012 1st Grade</td>
<td>35 (9)</td>
<td>5</td>
<td>33</td>
<td>5</td>
<td>68</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>327 (127)</td>
<td>51</td>
<td>312</td>
<td>49</td>
<td>639</td>
<td>100</td>
</tr>
</tbody>
</table>

Note: The counts listed in parenthesis are students who experienced two years of intervention. Within the subgroup of students who are indicated as having experienced two years of intervention, a student in a 1st grade cohort labeled as part of this subset would have been a student who was retained in 1st grade. A kindergarten student labeled as part of this subset would have either been in intervention during both kindergarten and 1st grade or had been retained in kindergarten and received intervention both years of kindergarten. (Side note: The district reports a typical kindergarten retention rate of approximately 8% and noted that approximately 50% of these retentions are due to maturity concerns by the teacher and parent.)

As displayed in Table 8, the kindergarten intervention and non-intervention groups ranged from 16-21% per cohort of the larger grade-level population (i.e. estimated grade-level population is 350 students) and the first-grade groups ranged from 9-15% per cohort of the grade-level population. Through the filtering process detailed above, cohorts 4, 5, and 6 were reduced. Initially, each component of a cohort would have represented fewer than 20% of grade-level population. This is a critical piece of evidence that supports high-quality Tier 1 instruction based on current standards, 80-85% of students should have their needs met by Tier 1 (Fuchs & Deshler, 2012; Fuchs & Fuchs, 2006; NASDSE, 2006; Shapiro, 2011). Viability of the core program is an important aspect when articulating the results of RTI due to the fact that a poor core program would yield non-responders that would be easier to remediate in Tier 2 (Bean, 2008; Gersten & Dimino, 2006; Hill, et al., 2012; Newmann et al., 2011).

A final note regarding the general demographics of the student sample, following the filtering process, a new subgroup emerged from the data as noted in parenthesis. This subgroup would include a student who had been selected for intervention twice, such as during both kindergarten and first grade. This outcome will be discussed in Chapter 5.
Demographic data for students included in the study are presented in Table 9.

Table 9

<table>
<thead>
<tr>
<th>Demographic Details</th>
<th>Intervention</th>
<th>Non-intervention</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>187</td>
<td>29</td>
<td>161</td>
</tr>
<tr>
<td>Female</td>
<td>140</td>
<td>22</td>
<td>151</td>
</tr>
<tr>
<td>Total</td>
<td>327</td>
<td>51</td>
<td>312</td>
</tr>
<tr>
<td>Socio-economic Status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Free Lunch</td>
<td>96</td>
<td>15</td>
<td>47</td>
</tr>
<tr>
<td>Reduced Lunch</td>
<td>33</td>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td>Paid Lunch</td>
<td>198</td>
<td>31</td>
<td>230</td>
</tr>
<tr>
<td>Total</td>
<td>327</td>
<td>51</td>
<td>312</td>
</tr>
<tr>
<td>Ethnicity *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>8</td>
<td>1.3</td>
<td>4</td>
</tr>
<tr>
<td>Asian</td>
<td>5</td>
<td>0.8</td>
<td>3</td>
</tr>
<tr>
<td>Hispanic</td>
<td>6</td>
<td>0.9</td>
<td>2</td>
</tr>
<tr>
<td>White</td>
<td>298</td>
<td>46.6</td>
<td>293</td>
</tr>
<tr>
<td>Multiracial</td>
<td>10</td>
<td>1.6</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>327</td>
<td>51.2</td>
<td>312</td>
</tr>
<tr>
<td>Special Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No identification</td>
<td>241</td>
<td>38</td>
<td>281</td>
</tr>
<tr>
<td>Identified</td>
<td>86</td>
<td>14</td>
<td>31</td>
</tr>
<tr>
<td>Total</td>
<td>327</td>
<td>51</td>
<td>312</td>
</tr>
<tr>
<td>Limited English Proficient</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>322</td>
<td>50</td>
<td>309</td>
</tr>
<tr>
<td>No</td>
<td>327</td>
<td>51</td>
<td>312</td>
</tr>
</tbody>
</table>

Note: * No students were identified as Indian American or Hawaiian. ** The demographics of the initial sample, the data set prior to removal of the student who moved in after the fall of kindergarten testing window or moved out prior to the spring testing window in third grade and the elimination of the duplicate students, appear to be the same as the demographics of the filtered sample. Males made up 54.8% of the initial sample, as compared to 54.5% after the filtered sample. White students were 92.8% of the initial sample, as compared to 92.5% after the filtered sample. As for SES, the free, reduced, and paid lunch percentages were 23.4%, 10.8%, and 65.8% prior and 22.4%, 10.6%, and 67.0% after, respectively. Finally, the original sample had 18.5% of the students who had an IEP in comparison to the filtered sample at 18.3% with an IEP.

Consistent with the district’s population (see Appendix A), the study sample represents approximately 591 (92.5%) White and 8 (1%) are Limited English Proficient (LEP) students.

From this point forward, ethnicity and LEP will not be included in the analysis because of the small sample size in those categories (e.g. between 8-10 students indicated under each minority) and the potential lack of anonymity. Throughout this dissertation, lower poverty level is considered to include both students on free lunch and reduced lunch. It is noteworthy that the
percentage of males in the sample (55%) is elevated as compared to the district population (51%). Furthermore, the sample exhibited a higher percentage of students in the lower-poverty-level (33%) than the district (30%). This shift parallels research results for struggling readers, as more males demonstrate challenges in learning to read than females and lower income populations tend to show a higher percentage of reading struggles (NCTE, 2006; NCES, 2009; NCES, 2011). Finally, as one might anticipate considering an intended outcome of RTI process is the identification for special education services, students are identified more often for services in the intervention group compared to the non-intervention group, 26% and 10% respectively. Inferential statistics will conclude if these differences are statistically significant related to the research questions.

**Gender by nested group.** Table 10 presents the cross tabulation by gender based on placement within or not within the intervention group by cohort as well as the overall. Along with Table 10, Figure 9 visually represents the intervention group gender distribution. Note that while 55% of the overall study population is male, the intervention group included more males (57%) than the non-intervention group (52%). Once again the male population appears to represent a higher percentage within the intervention group of the school population. One anomaly that stands out in Figure 9 is within cohorts 5 and 6 in which the intervention subgroups show a balance between males and females. Further investigation would be merited to determine if this anomaly was due to the removal of the duplicate students upon filtering the study group. In addition, note that I initially balanced the cohort population size.
Table 10

Cross tab: Receive intervention by cohort with gender

<table>
<thead>
<tr>
<th>Demographic</th>
<th>Cohort 1</th>
<th>Cohort 2</th>
<th>Cohort 3</th>
<th>Cohort 4</th>
<th>Cohort 5</th>
<th>Cohort 6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>41</td>
<td>59</td>
<td>45</td>
<td>64</td>
<td>37</td>
<td>53</td>
<td>31</td>
</tr>
<tr>
<td>Female</td>
<td>28</td>
<td>41</td>
<td>25</td>
<td>36</td>
<td>33</td>
<td>47</td>
<td>22</td>
</tr>
<tr>
<td>Total</td>
<td>69</td>
<td>70</td>
<td>70</td>
<td>53</td>
<td>30</td>
<td>35</td>
<td>16</td>
</tr>
<tr>
<td>Non-intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>32</td>
<td>56</td>
<td>34</td>
<td>47</td>
<td>32</td>
<td>50</td>
<td>21</td>
</tr>
<tr>
<td>Female</td>
<td>25</td>
<td>44</td>
<td>39</td>
<td>53</td>
<td>32</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>57</td>
<td>73</td>
<td>64</td>
<td>46</td>
<td>39</td>
<td>33</td>
<td>31</td>
</tr>
</tbody>
</table>

Note: * Cohort 1, 2, and 3 are the kindergarten cohorts and Cohort 4, 5, and 6 represent the 1st grade cohorts. Numbers are lower in the 1st grade cohorts due to the previously discussed filtering of students to be included in only one cohort.

Figure 9. Gender Distribution by Cycle within the Intervention Subgroup

When reviewing gender based on school in Figure 10, it is noteworthy that School C has a higher percent of male students in intervention (63%) as compared to School A (53%) and School B (54%). Inferential statistics will establish if these differences in the descriptive statistics are statistically significant related to the research questions.
Socio-economic by nested group. Examining the data on socio-economic status, the intervention subgroup shows a higher percentage of students with a lower poverty classification (39%) than both the non-intervention subgroup (26%) and the combined district (29%). Note the shift in cohort 6 may be due to the removal of duplicate students from the study data.

Table 11
Cross tab: Receive intervention by cohort with SES

<table>
<thead>
<tr>
<th>Demographic</th>
<th>Cohort 1</th>
<th>Cohort 2</th>
<th>Cohort 3</th>
<th>Cohort 4</th>
<th>Cohort 5</th>
<th>Cohort 6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Free/Reduced</td>
<td>26</td>
<td>38</td>
<td>32</td>
<td>46</td>
<td>31</td>
<td>44</td>
<td>20</td>
</tr>
<tr>
<td>Paid</td>
<td>43</td>
<td>62</td>
<td>38</td>
<td>54</td>
<td>39</td>
<td>56</td>
<td>33</td>
</tr>
<tr>
<td>Non-intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Free/Reduced</td>
<td>13</td>
<td>23</td>
<td>12</td>
<td>16</td>
<td>21</td>
<td>33</td>
<td>16</td>
</tr>
<tr>
<td>Paid</td>
<td>44</td>
<td>77</td>
<td>61</td>
<td>84</td>
<td>43</td>
<td>67</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 12 shares the cross-tabulation based on socio-economic status for students receiving intervention by school. When investigating the average school SES for the three enrollment years (i.e. 2008, 2009, and 2010 = kindergarten enrollment), one notices that the percent of lower poverty students in the intervention group (35%, 43%, and 40%) is higher than the percent within school population (28%, 36%, and 29%) for each school.
Table 1
Cross tab: Receive intervention by school with socio-economic status

<table>
<thead>
<tr>
<th>Frequency</th>
<th>School A</th>
<th></th>
<th>School B</th>
<th></th>
<th>School C</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Free/Reduced</td>
<td>37</td>
<td>35</td>
<td>44</td>
<td>43</td>
<td>48</td>
<td>40</td>
</tr>
<tr>
<td>Paid</td>
<td>68</td>
<td>65</td>
<td>58</td>
<td>57</td>
<td>72</td>
<td>60</td>
</tr>
<tr>
<td>Non-Intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Free/Reduced</td>
<td>18</td>
<td>21</td>
<td>27</td>
<td>27</td>
<td>37</td>
<td>29</td>
</tr>
<tr>
<td>Paid</td>
<td>69</td>
<td>79</td>
<td>72</td>
<td>73</td>
<td>89</td>
<td>71</td>
</tr>
</tbody>
</table>

Special education by nested group. Table 13 reflects the demographic data by cohort for inclusion in special education. Cohort 2 and 4 (2009-2010; year two of RTI implementation, kindergarten and first grade) and Cohort 6 (2011-2012; year four, first grade) show fewer students identified as needing an IEP at 13% and 11% respectively; as compared to, Cohort 1, 3, and 5 resulting in 28%, 21%, and 20% respectively. Once again, Cohort 6 shows unusual results in the fact that the percentage of students resulting in needing an IEP in the intervention group (11%) is lower than the non-intervention group (18%). A column proportionality test (see note in Table 13), confirmed that the subsets of the cohorts 1-6 do not differ significantly from each other ($p = .05$), thus rendering them comparable.

Table 13
Cross tab: Receive intervention by cohort with including in Special Education

<table>
<thead>
<tr>
<th>Demographic</th>
<th>Cohort 1</th>
<th>Cohort 2</th>
<th>Cohort 3</th>
<th>Cohort 4</th>
<th>Cohort 5</th>
<th>Cohort 6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No IEP</td>
<td>44_a</td>
<td>64</td>
<td>55_a</td>
<td>79</td>
<td>49_a</td>
<td>70</td>
<td>43_a</td>
</tr>
<tr>
<td>Has an IEP</td>
<td>25_a</td>
<td>36</td>
<td>15_a</td>
<td>21</td>
<td>21_a</td>
<td>30</td>
<td>10_a</td>
</tr>
<tr>
<td>Non-Intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No IEP</td>
<td>47_a</td>
<td>82</td>
<td>69_a</td>
<td>94</td>
<td>57_a</td>
<td>89</td>
<td>45_a</td>
</tr>
<tr>
<td>Has an IEP</td>
<td>10_a</td>
<td>18</td>
<td>4_a</td>
<td>6</td>
<td>7_a</td>
<td>11</td>
<td>1_a</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No IEP</td>
<td>91</td>
<td>72</td>
<td>124</td>
<td>89</td>
<td>106</td>
<td>79</td>
<td>88</td>
</tr>
<tr>
<td>Has an IEP</td>
<td>35</td>
<td>28</td>
<td>19</td>
<td>13</td>
<td>28</td>
<td>21</td>
<td>11</td>
</tr>
</tbody>
</table>

Note: The subscript a denotes a subset of Cohort categories whose column proportions do not differ significantly from each other at the significance level of 5% based on the $z$-test adjusted p-values by Bonferroni method.
Population data for inclusion in special education by school, Table 14, indicates a higher percentage of students who received intervention receiving an IEP. Since one outcome of the RTI process is identification for special education, this would be predicted. However, it is noteworthy to mention that the percentage of students with an IEP in the intervention group and the non-intervention group are much closer for School B than for School A or C. Another method of identification for special education is parent referral. Parent referrals may impact the percentage of students identified through the RTI process.

Table 14
Cross-tab: Receive intervention by school with including in Special Education

<table>
<thead>
<tr>
<th>Frequency</th>
<th>School A</th>
<th></th>
<th>School B</th>
<th></th>
<th>School C</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No IEP</td>
<td>70</td>
<td>67</td>
<td>82</td>
<td>80</td>
<td>89</td>
<td>74</td>
</tr>
<tr>
<td>Has an IEP</td>
<td>35</td>
<td>33</td>
<td>20</td>
<td>20</td>
<td>31</td>
<td>26</td>
</tr>
<tr>
<td>Non-Intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No IEP</td>
<td>79</td>
<td>91</td>
<td>85</td>
<td>86</td>
<td>117</td>
<td>93</td>
</tr>
<tr>
<td>Has an IEP</td>
<td>8</td>
<td>9</td>
<td>14</td>
<td>14</td>
<td>9</td>
<td>7</td>
</tr>
</tbody>
</table>

Lastly, an analysis of identification for special education by year of implementation specifically for the study population would yield 27.8%, 12.4%, 20.6%, and 14.7% for the years 2008, 2009, 2010, and 2011 respectively by the end of third grade. Note that the percentages are higher than one may anticipate because this is a subset of the school population and therefore the results are skewed in comparison. The school special education population is expressed in Appendix A for each of the implementation years.

The student cohort sample was divided into two groups, intervention and non-intervention. The entry-year of the cohort spanned a four year period and included two grade levels, kindergarten and first grade. Overall the demographics of the sample aligned closely with the district demographics with small elevations in male, SES, and special education. The intervention group had slightly higher percentage of males, lower-poverty, and special education
when compared to the non-intervention group and the district. There were a few inconsistencies between the schools and within the cohorts. The implications of these differences will be analyzed in the results section following the completion of the demographics for this study to understand if any significance exists.

**Research Questions Results**

This section is organized based on the research questions. Both quantitative and qualitative results will be explained. The quantitative results were based on student data to investigate the long-range outcomes following intervention. The qualitative results were based on open-ended questions related to the research questions that were posed to four focus groups. A sample of the detailed coding can be found in the results. Coding can be found in Appendix F. The results of the coding will be used to support the quantitative results by research question.

First will be the analysis regarding the long-range differences on reading achievement. Second will be the results regards to special education identification by the end of third grade. The third section will present the students' reading achievement outcomes among and within the various nested groups and the final section will present the data from the focus groups.

**Research Question 1: What are the long-range (3 or 4 years) differences on reading achievement for students who participated in RTI Tier 2 reading intervention as compared to students who scored slightly above the intervention group and did not receive treatment?**

Using the pre-test (TERA-3) and the post-test (IREAD-3, ISTEP+, and SRI) to formulate the analysis for the overarching research question, inferential analysis, such as independent t-tests, and descriptive statistics, such as odds ratios, were conducted to understand if there were significant differences among the means. Descriptive statistics, including frequency
distributions, were considered to understand generalities around the data set. A range of tests were conducted due to the type of data available for each assessment. These analyses will be elaborated on throughout chapter 4.

First, an independent t test for between-group comparisons was conducted to establish if significant differences existed between the intervention and the non-intervention subgroup reading achievement from entry year to third grade. Determination of achievement was based on the NCE gains for the post-test, Scholastic Reading Inventory (SRI), as compared to the pre-test NCE scores on TERA-3. The norm curve-equivalent (NCE) scale scores from TERA-3 and SRI allowed me to make meaningful comparisons among different achievement tests and among different nested groups.

The intervention group exhibited more growth on the NCE scale score than the non-intervention group. This growth indicates that the intervention group is catching up, while still slightly behind. The outcome of the t-test demonstrated that the intervention group had statistically significant gains over the non-intervention group, t(614) = 7.17, p < .001. The intervention group had a mean gain of 31.03 with a standard deviation of 22.02; while the non-intervention group had a mean gain of only 18.83 with a standard deviation of 20.09. The multivariate result was significant, Pillai’s Trace effect size indicated medium practical significance for the difference in the gains for students between intervention and non-intervention (V = .583, F = 858.85, df = p = .01). A visual representation of these results is shown by the graphical representation of the general linear model in Figure 11. In the figure, the rate of change for the intervention group is dramatically sharper than the rate of change for the non-intervention groups. More specifically based on the means, the intervention groups gained twelve (12) points of growth over the non-intervention group growth rate. The intervention
group is closing the achievement gap, but still behind the non-intervention group.

![Graphical Representation of the General Linear Model Results of the Difference in Gains between the Intervention and Non-intervention Subgroups](image)

*Figure 11. Graphical Representation of the General Linear Model Results of the Difference in Gains between the Intervention and Non-intervention Subgroups*

Next frequency data for the pass/fail rate for each assessment is presented based on inclusion in intervention or non-intervention groups (Figure 12). TERA-3 pre-test scores are compared to IREAD-3, ISTEP+ Grade 3, and SRI post-test scores. When considering the results of the IREAD-3 and ISTEP+, I wanted to note that both are criterion-referenced assessments that provide a minimum competency measure based on mastery of the Indiana Academic Standards. The primary outcome for each is a pass-fail rating. According to the *Guide to Test Interpretation: Grade 3-8 Indiana*, ISTEP+ results provide information on where a student stands relative to the three performance levels established by the Indiana State Board of Education and IREAD-3 provided two performance levels. ISTEP+ and IREAD-3 do not provide norm-referenced information, but instead categorical data.
To parallel the categorical data from ISTEP+ and IREAD-3, the TERA-3 national norm descriptive ratings were combined into two categories: (1) Very Poor to Below Average, and (2) Average and Very Superior. The Lexile score from Scholastic Reading Inventory (SRI) were grouped into below-grade-level and on-grade-level based on SRI Proficiency Placement Scale (Scholastic, 2009). The two categories created were (1) At Risk to Basic 2, and (2) Low Proficient to Advanced.

The TERA-3 pre-test results signify that 80% of the intervention group and a 49% the non-intervention group demonstrated achievement levels at Very Poor to Below Average. A chi-square ($\chi^2$) analysis was performed to learn whether or not there was a significant difference between expected observations of TERA-3 results by group participation and observed rates across the groups. The results concluded indeed there was very strong evidence of a relationship between subgroup inclusion (intervention or non-intervention) and failing TERA-3 ($\chi^2 = 65.01$, df = 1, $p < .001$). The probability that the variables, subgroup and pass/fail, are dependent upon each other. The next step is to substantiate the degree of strength of the relationship. According to Wiley (2006), the phi coefficient ($\phi$) is utilized in educational testing instead of Pearson r or Cramer V when a continuous variable, such as a test score, has a binary scale imposed. Since this analysis is based on a 2 x 2 contingency table with a pass/fail category obtained in relationship to a threshold, this standard holds true. Based on the results of this analysis, failing TERA-3 has a moderate association with inclusion in intervention ($\phi = -.322$, $p < .001$). Next the same measures were run for each post-test.

The post-test results for IREAD-3 included a 16% did-not-pass rate for the intervention group as compared to a 5% did not pass rate for the non-intervention group. There is strong evidence of a relationship between inclusion in intervention and not passing IREAD-3 in third
grade ($\chi^2 = 18.15$, df = 1, $p < .001$). However, based on the phi coefficient, the association between did-not-pass IREAD-3 and intervention is minor ($\phi = -.170$, $p < .001$).

Next, the ISTEP+ post-test results show a 22% did-not-pass rate for the intervention group as compared to a 7% did-not-pass rate for the non-intervention group. There is very strong evidence of a relationship between inclusion in intervention and not passing ISTEP+ in third grade ($\chi^2 = 28.35$, df = 1, $p < .001$). From the results of the phi test, there is a minor association between did-not-pass ISTEP+ and inclusion in intervention over non-intervention students ($\phi = -.213$, $p < .001$).

Finally, Lexile scores from SRI post-test show a 27% below grade-level rate for the intervention group and a 20% below grade-level for the non-intervention group. A chi-square test of independence provided evidence of a relationship between below grade level Lexile score and inclusion in the intervention group ($\chi^2 = 4.37$, df = 1, $p = .036$). Nevertheless, the association is not statistically significant between these factors according to the phi coefficient test ($\phi = -.083$, $p = .036$).

To summarize, the association between failing the assessment and inclusion in the intervention group was moderate for the pre-test, TERA-3. This association would be expected based on the grade-level rank order process for determination of students in inclusion. However, a less expected result would be that the post-tests (IREAD-3 and ISTEP+) only exhibited a minor association. This shift in association will be discussed in Chapter 5.
<table>
<thead>
<tr>
<th>Student Assessment Results</th>
<th>Legend: Pass: Green  Fail: Blue</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intervention</strong></td>
<td><strong>Non-intervention</strong></td>
</tr>
<tr>
<td><strong>TERA-3 (Pre-test)</strong></td>
<td>Pass = Average to Superior ranking on the NCE  Fail = Very poor to Below average on the NCE</td>
</tr>
<tr>
<td><img src="image1" alt="Pie Chart" /></td>
<td><img src="image2" alt="Pie Chart" /></td>
</tr>
<tr>
<td>80% DNP</td>
<td>49% DNP</td>
</tr>
<tr>
<td><strong>IREAD-3 (Post-test)</strong></td>
<td></td>
</tr>
<tr>
<td><img src="image3" alt="Pie Chart" /></td>
<td><img src="image4" alt="Pie Chart" /></td>
</tr>
<tr>
<td>16% DNP</td>
<td>5% DNP</td>
</tr>
<tr>
<td><strong>ISTEP+ (Post-test)</strong></td>
<td>Pass = Pass and Pass+</td>
</tr>
<tr>
<td><img src="image5" alt="Pie Chart" /></td>
<td><img src="image6" alt="Pie Chart" /></td>
</tr>
<tr>
<td>22% DNP</td>
<td>7% DNP</td>
</tr>
<tr>
<td><strong>SRI (Post-test)</strong></td>
<td>Pass = On or above grade level based on Lexile  Fail = Below grade level based on Lexile</td>
</tr>
<tr>
<td><img src="image7" alt="Pie Chart" /></td>
<td><img src="image8" alt="Pie Chart" /></td>
</tr>
<tr>
<td>27% DNP</td>
<td>20% DNP</td>
</tr>
</tbody>
</table>

*Figure 12. Pre/Post-Test Descriptive Levels by Subgroup of Intervention and Non-intervention*
Inspection of the visual representation, Figure 12, shows an association between membership in the intervention group and not passing the assessment. This association appears to be strongest for TERA-3 and weaker for ISTEP+, IREAD-3, and SRI Lexile. This visual inspection aligns with the phi coefficient values. While the rate of failure is still higher for the intervention group, the shift implies greater growth for intervention students relative to non-intervention students. The appearance is that the two groups are closer aligned following the intervention. This parallels the results of the t-test for mean gain on the NCETERA-3 to SRI.

Having analyzed at the comparison within an assessment, the next logical step would be to examine across assessments. Nevertheless due to the type of assessments and the available results data, the means for the assessment results cannot be directly compared across assessments through inferential statistics. However for each assessment, the means of the intervention compared to the non-intervention groups show results with the non-intervention group slightly higher than the intervention group for each assessment. The test means along with other characteristics of the data set are listed in Table 15. In all cases, the distribution is relatively normal since the values are near zero. On each post-test, the tail is slightly longer on the left side, which may be due to the group selection of struggling readers in comparison to a typical population. The data sets have low kurtosis values so has a light tail. The ISTEP+ post-test for the non-intervention group does show an elevation in the number of outliers as compared to the other assessments ($\alpha_4 = 1.17$). In general, the distribution of each test is relatively normal; however, no further testing can be completed on the means to further compare them. This is due to the type of scores available for the state assessments. From this point forward, comparisons among assessments will be based on pass/fail scale.
Table 15

Mean Based on the Assessment Score for Each Pre- and Post-test by Subgroup

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Min.</th>
<th>Max.</th>
<th>M</th>
<th>SD</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-Test (TERA-3 Reading Quotient)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>325</td>
<td>43</td>
<td>119</td>
<td>80</td>
<td>13.2</td>
<td>.126</td>
<td>-.220</td>
</tr>
<tr>
<td>Non-intervention</td>
<td>303</td>
<td>55</td>
<td>139</td>
<td>91</td>
<td>12.4</td>
<td>.048</td>
<td>.279</td>
</tr>
<tr>
<td><strong>Post-Test (IREAD-3)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>320</td>
<td>304</td>
<td>650</td>
<td>493</td>
<td>53</td>
<td>-.143</td>
<td>.902</td>
</tr>
<tr>
<td>Non-intervention</td>
<td>308</td>
<td>335</td>
<td>650</td>
<td>520</td>
<td>56</td>
<td>.506</td>
<td>.774</td>
</tr>
<tr>
<td><strong>Post-Test (ISTEP+ Gr 3)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>318</td>
<td>285</td>
<td>581</td>
<td>448</td>
<td>48</td>
<td>-.503</td>
<td>.831</td>
</tr>
<tr>
<td>Non-intervention</td>
<td>307</td>
<td>298</td>
<td>598</td>
<td>473</td>
<td>42</td>
<td>-.098</td>
<td>1.17</td>
</tr>
<tr>
<td><strong>Post-Test (SRI Lexile)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>320</td>
<td>0</td>
<td>1007</td>
<td>621</td>
<td>118</td>
<td>-.459</td>
<td>-.032</td>
</tr>
<tr>
<td>Non-intervention</td>
<td>308</td>
<td>0</td>
<td>1002</td>
<td>646</td>
<td>173</td>
<td>-.430</td>
<td>.118</td>
</tr>
</tbody>
</table>

Note: TERA-3 Reading Quotient range from 35 to 165 (M=100, SD=15) (Reid, Hresko, & Hammill, 2001). The 2012, 2013, and 2014 ISTEP+ scale score range from 120 to 780 and the 2015 ISTEP+ scale score range from 200 to 800. For the 2012 to 2015 ISTEP+ assessments, the overall average for all students was 462.60 (i.e. Low SES 445.76 and Special Education 426.64) (IDOE and CTB/McGraw-Hill, 2012, 2013, 2014, 2015). IREAD-3 scores range from 200 to 650 with no Indiana mean data available. A value of zero (0) for SRI Lexile denotes a BR or beginning reader level. A Lexile between 500 and 800 is considered on grade level and a mean in the 600s would be considered proficient ranging between the 53-67% based on work by Metametrics, Inc incorporating nationally norm-referenced scores on standardized tests linked to the Lexile Framework® for Reading and medium-sized U.S. State’s End of Grade Test to the Lexile Framework (Scholastic, 2009).

Putting all assessments on an equal rating, Table 16 shows the pre- and post-test means by intervention and non-intervention. A rating of zero (0) was given to a failing score and a rating of one (1) was given to a passing score, as previously defined. In this table and the subsequent figure, the growth in the mean for the intervention subgroup appears to be greater than the non-intervention subgroup from the pre-test to the post-test. Even though the visual interpretation of the numbers does not provide inferential support for the intervention process, it does support the previous evidence of growth.
Table 16
Mean Based on Failure [Pass (1) and Fail (0)] for Each Pre- and Post-test by Subgroup

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Min.</th>
<th>Max.</th>
<th>M</th>
<th>SD</th>
<th>SEM</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-Test (TERA-3 NCE)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>325</td>
<td>0</td>
<td>1</td>
<td>.20</td>
<td>.403</td>
<td>.022</td>
<td>79.6</td>
</tr>
<tr>
<td>Non-intervention</td>
<td>303</td>
<td>0</td>
<td>1</td>
<td>.51</td>
<td>.501</td>
<td>.029</td>
<td>48.8</td>
</tr>
<tr>
<td><strong>Post-Test (IREAD-3)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>320</td>
<td>0</td>
<td>1</td>
<td>.84</td>
<td>.364</td>
<td>.020</td>
<td>15.6</td>
</tr>
<tr>
<td>Non-intervention</td>
<td>308</td>
<td>0</td>
<td>1</td>
<td>.95</td>
<td>.222</td>
<td>.013</td>
<td>5.2</td>
</tr>
<tr>
<td><strong>Post-Test (ISTEP+ Gr 3)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>318</td>
<td>0</td>
<td>1</td>
<td>.78</td>
<td>.417</td>
<td>.023</td>
<td>22.3</td>
</tr>
<tr>
<td>Non-intervention</td>
<td>307</td>
<td>0</td>
<td>1</td>
<td>.93</td>
<td>.258</td>
<td>.015</td>
<td>7.2</td>
</tr>
<tr>
<td><strong>Post-Test (SRI Lexile)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>320</td>
<td>0</td>
<td>1</td>
<td>.73</td>
<td>.444</td>
<td>.025</td>
<td>26.9</td>
</tr>
<tr>
<td>Non-intervention</td>
<td>308</td>
<td>0</td>
<td>1</td>
<td>.80</td>
<td>.399</td>
<td>.023</td>
<td>19.6</td>
</tr>
</tbody>
</table>

Note: For this table, failure equaled zero (0) and pass equaled one (1). TERA-3 NCE failure rate was based on the descriptive categories of Very Poor, Poor, and Below Average. SRI Lexile failure rate was based on Metametrics, Inc. proficiency ranges. Failure was defined as below 500L or At-Risk, Basic 1, and Basic 2 (Scholastic, 2009).

Figure 13. Mean for the Rate of Passing by Inclusion in Intervention

In analyzing the Odds Ratio (OR) for each assessment, the OR of not meeting the academic standard set as the grade-level proficiency level for students in intervention is greater than that of students in the non-intervention group; however, a comparison of the odds ratio from pre- to post-test shows a reduction in the odds for the intervention group in comparison to the non-intervention group. For the TERA-3 pre-test the odds of not meeting the standard for grade-
level proficiency are 4.1 times greater for the intervention group than the non-intervention group within a 95% confidentiality interval. In comparison, the results show a 3.4 times greater odds for IREAD-3, a 3.7 times greater odds for ISTEP+, and a 1.5 times greater odds for SRI.

<table>
<thead>
<tr>
<th></th>
<th>OR</th>
<th>(Intervention/ Non-intervention)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TERA-3</td>
<td>4.10</td>
<td></td>
</tr>
<tr>
<td>Post-test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IREAD-3</td>
<td>3.78</td>
<td></td>
</tr>
<tr>
<td>ISTEP+ Gr. 3</td>
<td>3.72</td>
<td></td>
</tr>
<tr>
<td>SRI</td>
<td>1.49</td>
<td></td>
</tr>
</tbody>
</table>

Note: OR was on intervention/non-intervention. The reciprocal for non-intervention/intervention would be OR(TERA-3) = 1/4.10 = 0.244; OR(IREAD-3) = 1/3.38 = 0.296; OR(ISTEP+) = 1/3.72 = 0.269; and OR(SRI) = 1/1.49 = 0.672.

In summary, the intervention group was closing the gap on the non-intervention group. However, the intervention group was still slightly behind. Closing the gap is indicated by a mean increase of 12 points by the intervention group over the non-intervention group in regards to the NCE gains from TERA-3 to SRI (m = 31.03 and m = 18.83 respectively). The gain was statistically significant and exhibited a medium effect size. While still slightly behind, one can see by referencing Figure 11 and 12, the intervention group is closing the gap on the non-intervention group. This is also portrayed by the improvement in the odds ratios for not meeting the academic standard set for grade-level proficiency on the pre- and post-tests.

Research Question 2: What differences occur among the RTI Tier 2 groups as compared to those students just outside the intervention group in regards to special education identification by the end of third grade?

When comparing the intervention and non-intervention group special education percentages, it should be noted that one purpose of the RTI process is the identification of special education needs. Within the cohorts, there were 117 (18.3%) students within the study sample
that were identified as needing an Individualized Education Program (IEP) for Specific Learning Disability (SLD) in reading. Out of this, 86 (13.5%) of the students had been in intervention during kindergarten and/or first grade and 31 (4.9%) had not. In other words, of the students identified with a SLD, 73.5% had been in intervention. A chi-square test of independence was performed to examine the relation between inclusion in intervention and special education identification. The relation between these variables was statistically significant, ($X^2 = 28.58$, df = 1, $p < .001$). A phi coefficient was computed to assess the strength of the association. There was little association found between the two variables, ($\phi = .211$, $p < .001$). Figure 18 shows the results by school, implementation year, and entry year. A statistically significant difference was found between identification for special education and each of the factors. For each of the statistically significant findings a small association was found. Exceptions to this were found for School B and implementation year four. Identification for SLD for School B was split 59% from the intervention groups and 41% from the non-intervention, as compared to 81% and 78% for School A and C, respectively. The lack of statistical significance for identification of special education in implementation year is demonstrated by 60% of the students who were identified for special education having not been in intervention, while only 40% had been in intervention. As a comparison, the years of implementation ranged from 16.7% to 28.6% of the students identified as needing an IEP for SLD not participating in intervention. As a reminder, the 2011 implementation year included only first grade and duplicated had been included in the kindergarten cohort (Cohort 3). Cohort 6 included 68 students.
Table 18
*Cross-tab: Chi-square and Phi Coefficient for the cross-tab of receiving intervention with inclusion in Special education by school, implementation year, or intervention grade*

<table>
<thead>
<tr>
<th>Chi-square and Phi test</th>
<th>X²</th>
<th>df</th>
<th>p</th>
<th>ϕ</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>School</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>15.95</td>
<td>1</td>
<td>.000</td>
<td>.288</td>
<td>.000</td>
</tr>
<tr>
<td>B</td>
<td>1.07</td>
<td>1</td>
<td>.301</td>
<td>.073</td>
<td>.301</td>
</tr>
<tr>
<td>C</td>
<td>15.77</td>
<td>1</td>
<td>.000</td>
<td>.253</td>
<td>.000</td>
</tr>
<tr>
<td><strong>Implementation year</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>5.43</td>
<td>1</td>
<td>.020</td>
<td>.208</td>
<td>.020</td>
</tr>
<tr>
<td>2009</td>
<td>14.67</td>
<td>1</td>
<td>.000</td>
<td>.247</td>
<td>.000</td>
</tr>
<tr>
<td>2010</td>
<td>15.06</td>
<td>1</td>
<td>.000</td>
<td>.272</td>
<td>.000</td>
</tr>
<tr>
<td>2011</td>
<td>0.618</td>
<td>1</td>
<td>.432</td>
<td>-.095</td>
<td>.432</td>
</tr>
<tr>
<td><strong>Entry year</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kindergarten</td>
<td>20.93</td>
<td>1</td>
<td>.000</td>
<td>.228</td>
<td>.000</td>
</tr>
<tr>
<td>First grade</td>
<td>7.55</td>
<td>1</td>
<td>.006</td>
<td>.179</td>
<td>.006</td>
</tr>
</tbody>
</table>

Note: 2008 includes only a kindergarten cohort and 2011 includes only a 1st grade cohort.

Considering the selection process for intervention, the minimal or lack of association will be discussed in Chapter 5.

Next a 2 x 2 between-subject ANOVA was conducted that examined the effect of inclusion in intervention or non-intervention (subgroup) and identification for a SLD by the end of third grade on NCE gains from TERA-3 to SRI. A two-way ANOVA analyzed the independent and joint effects of the variables in a single study. Both the subgroup and identification included two levels. There was no statistically significant interaction between the effects of group and identification for a SLD by the end of third grade on NCE gains from TERA-3 to SRI, \( F(1, 612) = 1.233, \ p = .267 \). Since the interaction is not significant, no main effects were reported. However, the graph of the estimated marginal means of the gain was included. This graph visually displays the growth of the groups, as well as, showing the similar slope of the two lines. Parallel lines indicate no interaction effect between inclusion in special education and intervention for academic gains.
Figure 14. Graph of Estimated Marginal Means of Gain for the Two-way ANOVA of the Effect of Group Inclusion and Identification for Special Education

Subsequently, I delved deeper into any relationships between identification for special education and other variables. Being that the cohorts did not show a significant difference in frequency among them, I turned to other nested groups for analysis. Two-way ANOVAs were run for gender/identification in special education, entry grade/identification, and SES/identification. No significant interaction was found for any of these results.

Table 19

<table>
<thead>
<tr>
<th>Variable</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender•Spec Ed</td>
<td>.751</td>
<td>.387</td>
</tr>
<tr>
<td>SES•Spec Ed</td>
<td>1.615</td>
<td>.200</td>
</tr>
<tr>
<td>Entry Grade•Spec Ed</td>
<td>.200</td>
<td>.655</td>
</tr>
</tbody>
</table>

Note: The degrees of freedom for each source was 1 and the error for each variable was 612.

Compared to the historical perspective from the research district, special education placement has shown a decrease over the course of the implementation of the RTI. Prior to the implementation of RTI in 2008, the district 3-year average for 2005, 2006, and 2007 was 19.0%
and specifically for 3rd grade the 3-year average was 23.8%. For the first three years of implementation of RTI, the district showed a decrease of special education identification to 17.6% and third grade decreased to a 20.9% average. From 2011 to 2013, the years in which the study cohort students were in 3rd grade, the average percentage of students in the district and third grade in special education both dropped again, 16.9% and 18.9% respectively. While these numbers include all classifications of special education, traditionally identification for other areas does not change over time.

![District Special Education Percent](image)

*Figure 15. Graph of the District Special Education Percent*

A review of the third grade special education percentages by school indicated that School A and School C saw a decrease of 8.8% and 4.8% respectively, while School B actually saw an increase of 1.0%.
This data provides a basis for discussion in Chapter 5 on differences in special education rates relative to the RTI implementation.

**Research Question 3: What differences are seen in students' reading achievement among and within the various nested groups (e.g. school, implementation year, achievement measures)?**

The first nested grouping would be intervention and non-intervention. The IREAD-3 and ISTEP+ results both indicated that the intervention group scores lower than the non-intervention groups. Both tests show a significant difference in the scale score means between the groups. The IREAD-3 scale score mean was 493 for the intervention group whereas the non-intervention group mean was 520. The means for ISTEP+ were 448 for the intervention and 473 non-intervention groups. An independent t-test was run for each assessment. Levene’s test was satisfied in all cases. The t-test results for the each post-assessment indicate that the intervention group is still behind the non-intervention group (IREAD-3: \( t(626) = 6.270, p < .001 \); ISTEP+: \( t(623) = 6.790, p < .001 \)). However, when reviewing the SRI Lexile t-test results, there is no
statistical significance between the means of the groups ($t(626) = 1.70, p = .090$). The difference in the assessment outcomes may be due to the test type and will be discussed in Chapter 5.

Referring back to Table 9, the population of the intervention subgroup is 39% in the lower poverty level and 57% male, as compared to the non-intervention subgroup at 26% in the lower poverty level and 52% male. Statistical analysis (two-sided Fisher’s exact test) confirmed that inclusion in the lower poverty level group and the intervention subgroup co-occurred more frequently than expected by chance ($p < .001$, Fisher’s exact test). The co-occurrence of lower poverty level and inclusion in the intervention group does pose a limitation for this study and needs to be considered when stating comparisons. However, based on the literature, one would anticipate this co-occurrence since poverty is the number one predictor of achievement (Marzano, 2003). Nevertheless, by delving one layer deeper and disaggregating the data by school, the Fisher’s exact test results indicated that while School A and School B show the same co-occurrence as the whole population, School C did not show a statistically significant co-occurrence ($p_{(School\ A)} = .037; p_{(School\ B)} = .027; p_{(School\ C)} = .083$, Fisher’s exact test). In other words, at School C inclusion in intervention broke the stereotypical poverty relationship with academic struggles. The difference between the schools as well as the outcome of the post-test will be discussed in the findings.

Next, the results of the two-sided Fisher’s exact test for gender and intervention, do not demonstrate statistical difference beyond that expected by chance ($p > .05$, Fisher’s exact test) for the whole group, by school, or by cohort with the exception of Cohort 2 ($p = .04$, Fisher’s exact test). The results indicate that gender distribution was balanced and did not play a factor in participation in intervention. This is unforeseen considering the literature on the relationship between males and reading achievement.
Examining the pre- and post-test results for pass-fail by school based on inclusion in intervention, Table 20 expresses the descriptive data for each assessment. For TERA-3, School A results comprise slightly fewer students in the Very Poor to Below Average category, defined previously as the failure group, than School B and C. For ISTEP+ and IREAD-3, all schools show similar pass rates with one exception of the non-intervention group at School C. As a point of reference, School A had the lowest SES and School C housed special populations of students (i.e. students on the autism spectrum or with emotional disorders) who may have been higher functioning and in the general population but do have unique needs.

Table 20
_Cross tab: Pre-test and Post-test by Intervention by School_

<table>
<thead>
<tr>
<th>Frequency</th>
<th>School A</th>
<th></th>
<th>School B</th>
<th></th>
<th>School C</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td><strong>TERA-3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very Poor to Below Average</td>
<td>76</td>
<td>72</td>
<td>84</td>
<td>82</td>
<td>98</td>
<td>84</td>
</tr>
<tr>
<td>Average and Above</td>
<td>29</td>
<td>28</td>
<td>18</td>
<td>18</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td>Non-Intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very Poor to Below Average</td>
<td>29</td>
<td>34</td>
<td>49</td>
<td>51</td>
<td>70</td>
<td>58</td>
</tr>
<tr>
<td>Average and Above</td>
<td>57</td>
<td>66</td>
<td>47</td>
<td>49</td>
<td>51</td>
<td>42</td>
</tr>
<tr>
<td><strong>IREAD-3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did Not Pass</td>
<td>19</td>
<td>18</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>Pass</td>
<td>86</td>
<td>82</td>
<td>84</td>
<td>85</td>
<td>100</td>
<td>86</td>
</tr>
<tr>
<td>Non-Intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did Not Pass</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Pass</td>
<td>86</td>
<td>99</td>
<td>94</td>
<td>97</td>
<td>112</td>
<td>90</td>
</tr>
<tr>
<td><strong>ISTEP+</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did Not Pass</td>
<td>24</td>
<td>23</td>
<td>24</td>
<td>24</td>
<td>23</td>
<td>20</td>
</tr>
<tr>
<td>Pass</td>
<td>79</td>
<td>75</td>
<td>69</td>
<td>70</td>
<td>83</td>
<td>73</td>
</tr>
<tr>
<td>Non-Intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did Not Pass</td>
<td>3</td>
<td>3</td>
<td>7</td>
<td>3</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Pass</td>
<td>75</td>
<td>86</td>
<td>81</td>
<td>84</td>
<td>96</td>
<td>77</td>
</tr>
<tr>
<td>Pass+</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>8</td>
<td>16</td>
<td>13</td>
</tr>
</tbody>
</table>

Note: Twelve (12) participants did not have TERA-3 data. Eleven (11) participants did not have IREAD-3 data. Fourteen (14) participants did not have ISTEP+ data.
Using the same gain analysis data (i.e. NCE SRI:TERA-3) as with the whole group, a one-way analysis of variance (ANOVA) was calculated on achievement gains for a variety of nested groups to further explore student achievement growth based on membership. In all cases homogeneity of variance was met according to the Levene’s test unless otherwise noted.

In comparing outcomes by the year of implementation through a one-way ANOVA, the analysis shows a statistically significant difference between years, [F(3, 612) = 4.176, p = .006]. The means for the years of implementation were 29, 27, 23, and 18 respectively. The mean differences between the first/fourth and second/fourth proved to be significant (Table 21) based on the post hoc comparisons using the Bonferroni test. However, no significant difference was ascertained between the other implementation years. Taken together, these results suggest differences in implementation over the time frame. This data will be explored further in the focus group summary. It is noteworthy at this time that central office direct involvement in protocols was greater in the first two years of implementation. In addition, the fourth year of implementation included only students who were identified in first grade (Cohort 6) and this is a reduced population due to the removal of duplicate students who participated in kindergarten (Cohort 3).

<table>
<thead>
<tr>
<th>Bonferroni</th>
<th>MD</th>
<th>SE</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008 (1st) compared to</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009 (2nd)</td>
<td>2.06</td>
<td>2.42</td>
<td>1.000</td>
</tr>
<tr>
<td>2010 (3rd)</td>
<td>5.40</td>
<td>2.48</td>
<td>.181</td>
</tr>
<tr>
<td>2011 (4th)</td>
<td>10.54</td>
<td>3.32</td>
<td>.010</td>
</tr>
<tr>
<td>2009 (2nd) compared to</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010 (3rd)</td>
<td>3.34</td>
<td>2.12</td>
<td>.692</td>
</tr>
<tr>
<td>2011 (4th)</td>
<td>8.48</td>
<td>3.06</td>
<td>.035</td>
</tr>
<tr>
<td>2010 (3rd) compared to</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011 (4th)</td>
<td>5.14</td>
<td>3.12</td>
<td>.596</td>
</tr>
</tbody>
</table>
Next, a comparison by school was run with no statistically significant difference found, \( F(2, 613) = .565, p = .569 \). This can be seen in the close proximity of the means by school, \( M_{\text{School A}} = 25; M_{\text{School B}} = 26; \) and \( M_{\text{School C}} = 24 \). The lack of statistical difference will be explored in the focus groups; however, it is noteworthy to recall that the RTI protocols, staff training, and student selection process for each school were the same.

In comparing gains by cohort, the Levene’s test rejected homogeneity of variance. Since the assumption of homogeneity of variance was not met, the Welch adjusted F ratios test (4.82) was calculated, which was significant \( (p < .001) \). My conclusion was that at least two of the six cohorts differed significantly on their mean gains between TERA-3 and SRI. To learn where the mean differences reside Games-Howell post hoc procedure was used. Mean differences were found between cohort 1 and cohort 3 \( (\text{MD} = 7.99, \text{SE} = 2.71, p = .050) \); cohort 1 and cohort 6 \( (\text{MD} = 10.54, \text{SE} = 3.31, p = .023) \); cohort 2 and cohort 3 \( (\text{MD} = 8.13, \text{SE} = 2.63, p = .032) \); and cohort 2 and cohort 6 \( (\text{MD} = 10.68, \text{SE} = 3.24, p = .016) \). Referring back to previous data, the mean gain in the NCE results from TERA-3 to SRI for these four cohorts were 29, 29, 21, and 18 respectively. As a point of reference, district administration was heavily involved during cohort 1 and 2 with ongoing support the additional years. In addition, cohorts 1, 2, and 3 were kindergarten groups and cohort 6 was a first grade group. Further exploration of these points of reference will be discussed in Chapter 5.

The comparison by entry grade was found to have no significance between entry in kindergarten \( (M = 26, SD = 21.3) \) and entry in first grade \( (M = 23, SD 23.1) \). The lack of significance will need to be reviewed in Chapter 5 in comparison to the other nested groups that were found to be significant. One specific area of interest would be the difference in cohort 6, which is only first grade, with cohort 1 and cohort 2.
Finally, an analysis based on the TERA-3 pass/fail descriptor was conducted. This analysis found a very significant difference in the means ($M_{\text{fail}} = 34.4$ and $M_{\text{pass}} = 8.2$), ($t(614) = 17.3, p < .001$). Taking this analysis one layer deeper, t-tests were run on the split on intervention and non-intervention, as noted in Table 19. These results revealed a significant difference in the failure rates between the subgroups, as well as in the pass rates.

<table>
<thead>
<tr>
<th>Table 22</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean Gain from TERA-3 NCE to SRI NCE</strong></td>
</tr>
<tr>
<td><strong>Means</strong></td>
</tr>
<tr>
<td><strong>Non-intervention</strong></td>
</tr>
<tr>
<td>Fail</td>
</tr>
<tr>
<td>Pass</td>
</tr>
<tr>
<td><strong>Intervention</strong></td>
</tr>
<tr>
<td>Fail</td>
</tr>
<tr>
<td>Pass</td>
</tr>
</tbody>
</table>

Note: Non-intervention subgroup, $t(297) = 11.4, p < .001$ and Intervention subgroup, $t(397) = 10.8, p < .001$)

Overall, the analysis of the various nested groups indicated a difference between the initial years of implementation (i.e. year 1 and year 2) and year 4 (i.e. cohort 6). However, the results show no significant difference among schools or between entry year (i.e. kindergarten or first grade) into intervention. Chapter 5 will explore the similarities and difference between the nested groups.

Once again exploring odds ratios (OR), the analysis by implementation year is listed in Table 23. It is interesting that the 2008 OR for TERA-3 is larger than the other years. This year includes only kindergarten students, as whereas 2009 and 2010 include both kindergarten and first grade students and 2011 includes first grade students. Another point worth review is the general incline in the odds ratio for the potential of failure in the non-intervention group over the intervention group for each post-test.
Table 23

*Odds Ratio for the Failure Rate by Year of Implementation*

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR (Non-intervention/Intervention)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2008</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-test: TERA-3</td>
<td>.800</td>
<td>.295, 2.172</td>
</tr>
<tr>
<td>Post-test: IREAD-3</td>
<td>.157</td>
<td>.034, .727</td>
</tr>
<tr>
<td>Post-test: ISTEP+</td>
<td>.202</td>
<td>.064, .637</td>
</tr>
<tr>
<td><strong>2009</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-test: TERA-3</td>
<td>.121</td>
<td>.066, .221</td>
</tr>
<tr>
<td>Post-test: IREAD-3</td>
<td>.216</td>
<td>.060, .781</td>
</tr>
<tr>
<td>Post-test: ISTEP+</td>
<td>.198</td>
<td>.072, .544</td>
</tr>
<tr>
<td><strong>2010</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-test: TERA-3</td>
<td>.291</td>
<td>.159, .533</td>
</tr>
<tr>
<td>Post-test: IREAD-3</td>
<td>.291</td>
<td>.123, .690</td>
</tr>
<tr>
<td>Post-test: ISTEP+</td>
<td>.243</td>
<td>.103, .569</td>
</tr>
<tr>
<td><strong>2011</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-test: TERA-3</td>
<td>.260</td>
<td>.093, .728</td>
</tr>
<tr>
<td>Post-test: IREAD-3</td>
<td>1.650</td>
<td>.258, 10.560</td>
</tr>
<tr>
<td>Post-test: ISTEP+</td>
<td>1.036</td>
<td>.270, 3.971</td>
</tr>
</tbody>
</table>

In exploring a several of nested groups within this study, significant differences were seen between the following:

- intervention and non-intervention group results on the scale scores for IREAD-3 and ISTEP; yet no significant difference on SRI Lexile,
- the percentage of students who were identified a lower poverty in the intervention group as compared to the non-intervention group for School A, School B, and the entire study group, and
- cohort groups in regard to the mean gain for NCE from TERA-3 to SRI, specifically 1-3, 2-3, 1-6, and 2-6.

However, no significant difference was seen for the following nested groups:

- gender between the intervention and non-intervention groups
- gender by school, or by cohort, with the exception of Cohort 2,
- implementation year, with the exception of the first/fourth and the second/fourth years,
• school to school, or
• kindergarten to first grade.

This summary provides a foundation for discussion in Chapter 5 on reading achievement among and between nested groups.

**Research Question 4: From the perspective of the focus group, what understandings result from the quantitative data analysis of the academic outcomes between and among the nested groups?**

As a foundation for the focus group analytics, let us begin with the focus group demographic information. Four focus groups were conducted. Participants comprised the special education director, school psychologist, elementary principals, reading coaches, special education teachers, and classroom teachers. All participants have been with the district the duration of the study, have worked with primary students, and volunteered to participate. As mentioned in the methodology, due to the small sample size of the focus groups, in-depth demographic data were not collected.

For the administrators’ focus group five (5) administrators met the criteria. All five participants were included in the group. The experience level of the administrator cohort ranged from 9 to 38 years, yielding an average of 20 years.

For the teacher focus groups, the criteria previously established in the methodology were met by 16 educators at School A, 13 at School B, and 12 at School C. The selection process was decided by willingness to participate. The seniority ranking was not required based on participation interest/availability. The teacher focus groups included 7 teachers from each school and included grade-level teachers, special education teachers, and reading coaches. The teaching experience within the three focus groups ranged from 8 to 29 years with an average of
15 years at School A, 14 years at School B, and 19 years at School C.

A total of 26 educators participated in four focus group interviews. In general, the focus groups included 4 males, 100% Caucasian, and ranged in teaching experience from 8 to 38 years. The exclusion of further detailed demographic data maintained anonymity in participants' responses. This lack of detailed focus group demographic data did not degrade the research study data presented in this section.

Overall results from the focus group indicated that the educators positively viewed the RTI process currently implemented in the district. Several teachers did express concern regarding the number of IEPs acquired, the length of time some students were involved in the RTI process, and whether each student who had been in intervention was translating the growth into class work. The majority of the focus group participants associated the following positive outcomes with the implementation of RTI: gains in student assessment scores, high IREAD-3 assessment results, using the data to inform instructional planning, using the data to monitor student growth, improved identification process for the student who needs an IEP, and generally catching more students before they fell through the cracks. The overall theme was that the quality of the early intervention can influence student achievement.

**Perspective of practitioners on long-range differences on reading achievement.** In reflecting on the data analysis, participants discussed the moderate significance of the 12 point mean gain in the NCE score from TERA-3 to SRI by the intervention group over the non-intervention group. The impact of this result provided confidence in the focus group participants that the RTI program was benefiting students and improving reading skills. Several commented on the amazing gains from TERA-3 to SRI NCE results and the reduction in the odds of failure for the intervention group in relation to the non-intervention group. One person commented that
"the intervention group showed rapid growth over a relatively short period of time." Several participants commented that it was nice to see that the hard work of the staff, as well as the students, was working. It was noted that in the day-to-day experience it is sometimes difficult to see the final outcome. Almost unanimously it was agreed upon that the correct students were being identified and that the high-quality, early intervention was helping those students. In one focus group, the comment was made that "The huge gap is closing." Another participant stated, "We are catching more students before they fall between the cracks and I am not referring to qualifying a student for an IEP."

Two key pieces of data were mentioned multiple times: IREAD-3 results and the slope of the line for the intervention groups from the TERA-3 to SRI gain. The discussion revolving around the IREAD-3 results focused on the success rate of taking students who were substantially behind in kindergarten or first grade and seeing the high pass rate on the core reading assessment. It was noted that even though this is a very high stakes test and some students find high stake test settings overly challenging, our students were demonstrating their reading ability. They felt that for many students the inclusion in the intervention process provided not only the skills, but the human support necessary to motivate the student and give him or her the efficacy to meet the goal on the IREAD-3 test. As for the graphic of the mean gains, one simple comment captured the group’s sentiment, "The lines coming together is what we want to see." In summary, while the intervention students were still behind on the SRI assessment in third grade, the means for the two groups were relatively close. Both groups had shown gains, but the intervention group's gain had outshined the non-intervention's gain.
**Perspective of practitioners on special education outcomes.** In the focus group discussion, it was noted that while the data seems to be trending down, the percentage of special education students within this population seemed very high. However, the conversation continued as I reminded the participants that this was a subset of the overall grade-level population and that this subset was intentionally selected for reading challenges. Therefore the student population was more susceptible to needing special education services. In looking at it this way, the sentiment was that the identification process along with the ongoing progress monitoring of the students was providing a valuable service to the student. An observation was that the district seems to be matching the needs of the student fairly well since the majority of IEP students have been supported through the RTI process for at least a year, if not several.

In the administrator focus groups, the student support through the RTI process was noted, as well as the fact that it did not occur in any formal or consistent manner prior to the implementation of the RTI process. In the previous structure, the district had waited for the student to fail for long periods of time before designing a custom program to support them. This was expanded upon by a comment that "previously most data was anecdotal records and it was difficult to make determinations based on that data and one series of assessments."

During School C's focus group, a participant remarked on the higher than normal percentages of IEPs at School A. School A enjoyed the lowest poverty level, did not house any special programs, and was receiving non-resident students (who were accepted based on screening data). The conversation invoked questions regarding the impact of parent dynamics, such as parent requests and parent follow-through on the number of IEPs. It was noted in the focus group that this a reflective thought and not intended to blame parents. It was noted that some parents, more emphasis was mentioned at School B and School C focus groups, have
difficulty following through with the paperwork for testing for unknown reasons. This does impact the referral rate completion, but the teachers noted that it does not halt student support through the RTI program. In this conversation, the potential outcomes of parent requests were discussed. School B had a higher level of non-intervention students with an IEP and possibly this was the result of more parent requests. One participant expressed that while School A had a lower poverty rate, the rates were relatively close and may not be a differentiating factor. Overall this discussion posed interesting questions regarding the differences between schools and implementation years.

One final discussion from one of the teacher focus groups revolved around the impact of social media and finances on student identification for special education. The participant stated that there have been times when the advocates were heavily active on social media and parent requests increased. This question was then posed to the administrators' focus group by the researcher. The administrators discussed the process of explaining the RTI process to parents in an attempt to allow them to give permission for intervention and to allow the program to work before moving directly to an IEP. Different administrators noted their personal success with this parent intervention. As for the financial influence, as the economy has changed so have the families attending the school. A participant commented that she had been told by a parent that a few parents seek out the IEP for the potential financial gains such as student financial aid or supplemental security income. In general, the parent’s engagement with their child’s learning can be a positive partnership with the school to provide a supportive team for their child.

**Perspective of practitioners on differences in nested groups.** The focus groups did not delve into the nested groups as much as the overarching intervention and non-intervention aspect. However, they did provide some interesting insights on gender, ethnicity, and SES.
Each focus group commented on their surprise on the topic of the distribution based on gender. The percent of males in intervention was not as high as they had anticipated. This perspective was based on perception rather than facts. Multiple aspects were discussed in regards to the classroom teachers' involvement in the tiers. Remember, the district protocol had intervention groups facilitated by either a classroom teacher or an instructional aide. Two different focus group members noted that if a teacher took an intervention group, the group would typically be the most challenging and in most cases larger. This may be the reason for the skewed perspective. On a positive note, they felt that the male population appeared to be improving slightly as the process evolved.

One participant stated that while the Black population was very small, it seemed to be high in the intervention group when compared to the district population. I restated at this time that ethnicity was not studied due to the small student population. In addition, considering the small minority population at the district within a particular grade level in a particular year, the minority percentage within that grade level may be higher than the general district population. An example of this occurred last year at School B. The number of English Language Learners within the kindergarten class was substantially higher than the school or the district.

The factors of between school, implementation year, and entry year were rarely mentioned. Only a cursory comment or two were made in each focus group. No clear themes arose from this discussion.

The nested group for socio-economic status was mentioned in each focus group multiple times. Several participants from each group commented on the connection between lower economic status and inclusion in intervention. The analysis of the SES results did show a relationship between inclusions in intervention and lower SES. This was not a surprising
revelation to the group. It was noted that the number of free lunch in the intervention group was larger than not only the paid lunch but the reduced lunch as well. A participant mentioned that low SES is the number one indicator of academic challenges. One participant stated, "Our placement results demonstrate that students are not a single test score and that the whole child influences individual's achievement." Subsequent to this comment, it was stated that the intervention process appeared to be showing an impact for these children and how do district educators inform policy makers about the critical nature of early childhood education. I will be discussing this more in Chapter 5 because of the negative unintended consequences that can result from such expectations.

In the administrators' focus group, it was mentioned that while the poverty rates at the schools are similar the type of student was different. A participant shared that School A had a different demographic of low-income students than School B or C. School A tends to include more heritage families who have had changes in their socio-economic status, were small farmers, or were grandparents raising children. On the other hand, School B and C have more transient and new families to the community. The difference in poverty would be more situational poverty versus generational poverty. In part, these differences are caused by proximity to interstate highways, land development, and farming. A final statement was how the district, and on a larger scale the United States, might intervene with students in lower poverty sooner.

**Perspective of practitioners on logistics and future needs.** In each cohort, the discussion turned to the logistics of the RTI process within the district that supported student achievement, along with future needs. Logistical items they felt impacted the growth included (1) the quality of Tier 1 instruction, (2) the size of the intervention group, (3) the organization of the program to meet individual student's needs with an appropriate program, and (4) proper
staffing arrangements. Future needs focused on maintaining a program that appears to be working, or program stability.

**Quality of Tier 1 instruction.** The first factor involved the quality of the work prior to that a child participated in intervention, or Tier 1 instruction. The administrator focus group witnessed that both student and teacher indicators demonstrated that the district’s teaching staff were highly-qualified and subsequently provided a solid instructional delivery. The focus group measured student indicators by benchmark assessment growth and summative assessment data for the district. Both measures showed consistent growth and by the end of the research study, the state summative assessments were approaching the district’s 90% pass rate goal. Teacher indicators mentioned included: (1) staff evaluation (i.e. district historically releases teachers who are deemed to be weak, district has strong hiring processes that includes ongoing professional development and mentors, and staff evaluation process is rigorous and intentional) and (2) academic degrees/licensure (i.e. 100% of teachers hold appropriate teaching license, 100% of teachers are highly-qualified based on Indiana’s Title II requirements, and 100% of teachers hold a minimum of a bachelorette degree with continued studies). More than a few focus group participants commented on the critical nature of assuring quality core instruction.

**Size of the intervention group.** The second item designated as a critical logistical component was delivery of intervention in a small group. Specifically, the focus group members believed that a group size between 4 and 8 was appropriate as evidenced by the student growth outcomes mentioned previously. It was noted by several members that most groups were comprised of 4 to 6 students, and for the larger group, 7 or 8, to work well the program structure required online components, such as Read Naturally, to permit individualization by the group facilitator. One teacher shared that the small group permitted the personalized attention
necessary for strong individualized growth.

**Organization of the program to meet individual student's needs.** The next component of logistics discussed by the focus group was the organization of the Tier 2 intervention groups, including the placement of students, rank order determination, the program alignment, and the support of all stakeholders. Participants talked about the care taken by all stakeholders to place students, not only in the appropriate content, but with the appropriate staff member and fellow students. In fact, reading coaches were thanked by the groups for their attention to the whole child. The rank order process was discussed and the consensus was that the teacher ranking component was a critical aspect of the process. It provided a picture of the whole child and not one snapshot from an assessment.

In addition, the use of the hybrid model for intervention was seen by a number of participants as beneficial. One administrator comments positively on the support provided to instructional aides to deliver quality instruction through the standard protocol. Yet another administrator mentioned the fidelity component that would be more difficult if structured programs were not utilized. Others appreciated the flexibility provided by a periodic review through a team planning session. It allows each student's needs to be addressed instead of pigeonholing a student in a pre-determined program. Several small aspects of the organizational structure were discussed by individual members of the focus groups. The big picture was the structures supported staff and provided a level of confidence in regards to quality for students.

**Staffing aspects.** The final critical component articulated by the focus group included a range of staffing aspects. Three aspects of staffing were mentioned multiple times were matching the staff member to the Tier 2 program, ongoing professional development for Tier 2 instructors, and leadership.
One participant commented on the essential need to match the group facilitator to the student and yet another commented on the longevity of classified staff supporting student growth. To explain the meaning of matching students, I asked for elaboration. The participant talked about personality strengths of the instructional aide, such as a firm approach or a quiet demeanor and the benefit of each to different students. In regards to longevity, instructional aides that return each fall provide additional consistency for students as well as a stronger knowledge base on the intervention process and programs. According to the focus groups, pairing the IA’s skills, knowledge, and personality with the student group helped provide the structural support the students needed. This staffing aspect was enhanced by ongoing support of the aides.

Another key staffing factor mentioned by each teacher focus group was the ongoing support of the instructional aides. This support came from continued professional development by the reading coach, more awareness of student’s needs, and strategies for managing behavior concerns. Participants concurred that since these support tools have been implemented the instructional aides have viewed RTI as a positive learning environment. It was noted that many IAs feel valued and have taken ownership of the process.

The final staffing component was leadership. The teacher focus groups talked about several layers of leadership, including the reading coach, building administrators, district administrators, and the final leadership aspect of teacher leaders. One example of the reading coach leadership was mentioned previously in this section. As for the district and building administrators role, several comments were made regarding the involvement of administrators in the process and the support provided by administrators. It was commented that the district administrator involvement provided the organizational, staffing, and financial support necessary
to allow the program to be implemented with fidelity and quality. In general, the focus groups believed the organizational structures, quality of the resources, and staffing played a critical role in students’ growth.

*Future needs for stability.* Every aspect of the focus group conversations was not positive. A few areas of concern were mentioned. The concerns dealt with program stability in respects to staffing changes, including administration; shifts in the focus of administration; student intervention groups remaining small as finances change; and an increase in the transient student population. One comment regarding the results was that the district administration had a stronger presence in the beginning and the group wondered how this impacted the results. Overall the feeling was that the district was targeting the correct students and accruing positive outcomes.

In addition to the response to the focus group questions, the focus group participants posed several reflective questions that provided additional insight into the study outcomes. These questions may prove valuable to future district decision-making. A few questions that seemed meaningful to the participants included:

1. As we look back at the intervention process and the current implementation, do we feel we are utilizing teachers in the intervention process appropriately to acquire the greatest student achievement?

2. The numbers of students in RTI seems to remain the same. Are we replacing students and should the number of students in RTI be reducing? (i.e. On a side note, a participant did provide a rebuttal that the bell curve will always exist.)

3. What percentage of the students cycle in and out of RTI tier 2? If a student continues to cycle through RTI, what else might be done? Should they receive an IEP?
4. How have the non-resident students and mobility impacted the RTI process?

5. For students in intervention who did not receive an IEP, what is their outcome beyond third grade?

Finally, the focus group participants asked that the results of this research study be shared with all teachers, instructional aides, and administrators as evidence of the value of our current implementation and as a conversation starter regarding future decisions.

Summary

The results reported in this chapter provide data to answer the four research questions. The purpose of this study was to examine the outcome of the RTI process on student achievement and the consequences of the implementation of RTI on the identification of special education. To recap the analysis, a medium practical significance was seen when evaluating the gain based on TERA-3 to SRI results in favor of the intervention group. This was visually represented in two ways: (1) Figure 11 illustrates the achievement gap closing between the intervention and non-intervention groups and (2) Figure 12 presents a closer alignment between pass and fail for the post-test results. In addition, the subtle differences between and among the various nested groups were expressed. The data reported in this chapter will be discussed in detail in Chapter 5. Moreover, the data already has provided evidence to the district to guide program decisions.
CHAPTER 5

CONCLUSIONS

This research was conducted in an effort to examine the outcomes of the RTI process on student achievement and special education identification across three schools within a school district in Indiana. It endeavored to provide information that would support the district leadership in decision-making as it relates to the implementation of RTI. Finally, the study sought to add to the knowledge base on RTI implementation to support practitioners. Both quantitative and qualitative research methods were employed in this longitudinal study.

The chapter begins with a brief summary of the study, which specifically addresses the research questions. Then the findings are compared to the research literature on cultural factors influencing student achievement and the intended outcomes of the RTI process. Finally, the chapter will conclude with implications for practice and suggestions for further research.

Overview of the Problem

Today's society is dependent on a literate population to drive economic and social progress (EnGage, 2003; Carnevale, Smith, & Strohl, 2013). Reading is a foundational skill to continued learning. Nevertheless, educational literature does not show a positive trend towards an increase in the literacy skills of society (NCTE, 2006; NCES, 2011). Higher-risk populations (e.g. students from lower income households) are not achieving at the same rate as their peers. Students are over-identified for special education, which in some cases provides an excuse for a lack of student achievement, rather than focusing upon continual learning improvement. Previously, the discrepancy model had been used to support struggling readers. The discrepancy
model has been said to be too little and too late, which has resulted in the unfortunate outcomes mentioned above (Fuch, 2006; Hintze, 2008). To ensure a literate population, schools must develop quality reading instruction with the aim of supporting each child. One component of supporting students would be to intervene as soon as reading problems are identified.

To increase literacy skills, reduce the achievement gap, and diminish over-identification of students for special education, a strong intervention strategy would place emphasis on prevention and intervention when a student began to struggle academically (Bursuck & Blanks, 2010; Coyne, Zipoli, & Ruby, 2006). Research indicates that intervention should begin no later than first grade. Response to Intervention (RTI) is one method that provides early intervention through research-based instructional strategies. The literature provides robust evidence on the merits of early intervention to prevent reading problems and aid in the identification of a disability (Coyne, Zipoli, & Ruby, 2006: Showers, Joyce, Scanion, & Schnaubelt, 1998; Slavin, 1994). However, minimal studies have been conducted on the effectiveness, efficiencies, and best practices for practitioners to follow. A longitudinal study of an implementation of RTI by a school district has the potential to provide additional evidence to support practitioners; therefore, enhancing its impact on student achievement and the identification process.

**Purpose of the Study**

The purpose of this study was to examine the outcomes of the RTI process, such as student achievement in reading and/or placement in special education. In addition, the data provided evidence to the district to guide program decisions. The study investigated whether students identified as lower achieving readers in kindergarten or first grade, who received deliberate intervention in reading within the RTI program, had higher levels of reading achievement by the end of third grade than students who did not receive the RTI reading intervention. The independent variables in this study were inclusion in the intervention or the
non-intervention group, gender, socio-economic status, implementation year, cohort, school, and entry grade. Academic achievement in reading was the dependent variable and was measured by reading assessment results and placement in special education.

**Research Questions**

Given this is a mixed-methods study, this research encompassed a primary or central question with sub-questions to guide the exploration of the data collection (Creswell, 2003). This overarching question guided this study:

1. What are the long-range (3 or 4 years) differences on reading achievement for students who participated in RTI Tier 2 reading intervention as compared to students who scored slightly above the intervention group and did not receive treatment?

The following sub-questions guided the research and data analysis for this study:

2. What differences occur among the RTI Tier 2 groups as compared to those students just outside the intervention group in regards to special education identification by the end of third grade?

3. What differences are seen in students' reading achievement among and within the various nested groups (e.g. school, implementation year, achievement measures)?

4. From the perspective of the focus group, what understandings result from the quantitative data analysis of the academic outcomes between and among the nested groups?

**Review of the Research Methods**

This longitudinal study was a mix-method research design to provide a richer interpretation of the complex phenomenon related to the implementation of RTI. Therefore, the study was conducted in two phases. First, the quantitative phase employed a non-equivalent group design (NEGD) that included six cohorts of students from three elementary schools within
a school district. Each cohort was comprised of kindergarten or first-grade students and included two sub-sets. The first sub-set included students who had the lowest benchmark scores within the grade-level at the school and therefore received intervention. The students in the second sub-set had received the next lowest scores and did not receive intervention. This provided a larger sample, as well as, providing parallel implementations to compare. To support ethical practice for the implementation of RTI of selecting the students in the most need, the sample was purposeful rather than random. Random sampling would have increased internal validity; however, the use of non-equivalent intact groups supports the intent of the RTI process (Gay, Mills, Airasian, 2006; Trochim, 2006). Second, the qualitative phase included four (4) focus groups that reflected on the quantitative results of the study and the focus group participants' professional insight on the implementation of the RTI process. An intentional choice was made to position the quantitative data as the primary data source and to use the qualitative data in a subsequent role with the intention of deepening the explanation of the results.

In the quantitative phase, the intervention student subgroups participated in small-group focused reading instruction in addition to the core reading instruction; whereas, the non-intervention group only received the core reading instruction. The sample population (n = 639) consisted of kindergarten and first-grade students from three demographically similar elementary schools within a school district (92.5% White, 33% low SES). Archival data was captured at the entry into kindergarten (i.e. TERA-3) and at the end of third grade (i.e. ISTEP+, IREAD-3, SRI, and identification for specific learning disability or SLD). Descriptive and inferential statistics were run for each cohort based on inclusion in the primary sub-group, socio-economic status (SES), gender, school, year of implementation, and entry year. For this study, ethnicity and ELL were not analyzed based on the district demographics. Chi-square and phi coefficient were used
to establish any associations between variables. T-tests and analysis of variance (ANOVA) tests were used to compare the means of the dependent variables to measure if student achievement gains, student pass rates, and identification for special education achieved statistical significance among the various nested groups. When appropriate, post hoc tests were used to ascertain where the differences exist and odds ratios were used to compare relative odds of the occurrence of the outcome given exposure to the intervention.

In the qualitative phase, participants in the focus groups (n = 26) included educators who had professional in-depth experiences with the district's RTI process and the student population. The semi-structured format of the focus groups added a perspective of the quantitative results and provided a holistic overview of the RTI implementation. Initial coding developed themes by deconstructing the data from the open-ended question responses (Saldana, 2009). To enhance the credibility of the coding, member checks were conducted and my dissertation chair reviewed the coding.

**Limitations of the study.** As in all research, this study had several limitations. One goal of this study was to add to the research base on the current implementation of RTI in schools. However, this study did have limits geographically and demographically, which limited the ability for generalization. The data came from one district with a relatively homogeneous population. Geographically, the study is from one mid-west state, and specifically one rural school district. Demographically, the student population at the district was predominately White with a moderate level of poverty. Conversely, the use of multiple cohorts, over several years and from multiple schools, aided in the generalization of the results.
The availability of the same pre- and post-test was also a limitation of this study. Since the same pre- and post-tests were not available, I had to compare assessments with similar scaled scores, such as NCE, and pass/fail rates on other assessments.

The study had additional limitations with implementation. Inherently, even with training, teachers introduce bias when administering the standardized assessments. Every teacher was trained in the delivery of the assessments; nevertheless, every teacher brings his/her own professional background leading to inherent bias. The data did not provide information on who conducted the assessments, so no inter-rater reliability analysis could be conducted among the teachers. Secondly, the quality or fidelity of the implementation of each intervention group may be different based on the training of the facilitator, the location of quiet space to meet, the consistency in the schedule, and periodic monitoring. All of these items were dependent on humans and imposed some limitation. Third, the student sample had internal bias because students were not randomly chosen from the population. Due to the unethical and practical nature of excluding students in need of support, random selection would not have been permitted in the district. The study used archival data from enrollment rather than random selection. Student participation was based on enrollment at the district between kindergarten and third grade during the years of the study. Some loss of data did occur due to either a student leaving prior to the end of third grade or enrolling after the pre-test window in kindergarten. It is noteworthy to mention that the district's mobility (5.4%) is less than the average state mobility (11.5%). Regardless, mobility posed limitations as students were omitted due to a lack of pre- or post-test data. Fourth, not having paired tests is a limitation to my study; realizing this fact, the normal curve equivalent tests did permit comparison. When working with an actual implementation and not an experimental study, a researcher must be flexible and work within the
boundaries of the environment to design the best possible study. However, I also needed to understand the limitations the environment imposed.

To verify consistency among the study sample, schools, and district, demographic comparisons were evaluated and each group was found to be similar to the others. These comparisons provided evidence that questions pertaining to the potential limits of the study due to the omission of these students (e.g. Are the results impacted by the most at-risk students leaving the study?) can be minimized.

The quality of the general classroom instruction, Tier 1, can be seen as a limitation to any study concerning the outcome of Tier 2. Information on the general curriculum, quality of the teaching staff, and detailed information validating Tier would increase the reliability of the results. I mentioned earlier that the knowledge base on the RTI framework would indicate at least 80% of the students should have their needs met in Tier 1. I presented data previously sharing that over 80% of the students in this district were meeting the state assessment measures. This provides some evidence to support quality Tier 1 instruction.

Finally, the size and construct of the focus groups presented study limitations. The total study group participation was 26 within the four focus groups. This sample size is adequate and the repetition increased the dependability of the results. The construct of the focus groups was to use volunteers. Volunteer participants may impact the results based on research that human volunteers may not represent the population in general (Rosenthal & Rosnow, 1975). Other limitations of volunteer sampling include potential bias towards a certain type of person who has a personal interest in the research topic and incomplete conclusions based on this imbalance.
Major Findings

The following is a brief description of the findings supported in this research.

- Students who received intervention grew significantly more than their peers who were just above the cut score and did not receive intervention.
- Students who received intervention substantially closed the gap between themselves and their peer group who were just above the cut score.
- TERA-3 NCE scores seemed to provide adequate data to discriminate students’ reading needs.
- Program fidelity impacts program outcomes. In this study, the role of leadership was highlighted as an area that impacts the outcomes.
- Following intervention, disproportionalities in gender and SES were non-significant.
- Identification of specific learning disabilities was greater for the students in the intervention group than their peers who score slightly above the intervention group. However, there are students in both groups who were identified for special education.
- The hybrid model, problem-based structure with standard treatment protocols, appears to have worked well to increase student achievement.

Findings Related to the Literature

In this section I will expound on the findings. A brief summation of the major results will be compared to the literature base. Some of the literature substantiates my findings and provides evidence that social influences can be overcome. My findings will be organized by research question through the lens of the quantitative study along with the perspective of the
qualitative focus group feedback. The qualitative data aided in triangulating the findings in the quantitative data. By incorporating both aspects into the study, the conclusions drawn were richer. When a finding impacts more than one research question, the finding will be discussed in the first research questions and be referenced thereafter.

Research Question 1: What are the long-range (3 or 4 years) differences on reading achievement for students who participated in RTI Tier 2 reading intervention as compared to students who scored slightly above the intervention group and did not receive treatment?

The first research question sought to understand the long-range achievement differences for students who received intervention as compared to the non-intervention group who scored slightly above them on the RTI grade-level ranking and received no intervention. Long-range achievement differences between the groups were understood by analyzing the differences on pre- and post-test measures and testing for significance. The results suggest that students in the intervention group made significantly greater gains compared to the non-intervention group. This finding is in agreement with the existing literature on the benefits of early intervention and the RTI framework. Previous research had found that early intervention increases the achievement rate and provides a strong likelihood of academic success (Denton, 2012; Flippo, 2001; Lopez, 2009; Vaughn et al., 2007; Marston, 2005). Wanzek and Vaughn (2007) found a larger effect size for interventions provided to kindergarten and first grade students than to upper elementary students. In a subsequent systematic literature review, Denton (2012) revealed a larger effect size for interventions provided in the early stages of reading over those provided after second grade.
Initially, my intent was to run inferential statistics, such as ANOVA, comparing the pre-test with each of the post-test result. The pre-test, TERA-3 provided a range of scores that could be analyzed through inferential statistics. However, a challenge arose with the ISTEP+ and IREAD3 post-tests. The state tests are criterion-referenced assessments that have already been converted to a scale score based item response theory (IRT). This scale score cannot be compared with other scale scores because the test content was “scaled” differently. Therefore, many inferential statistics were not able to be utilized. The next appropriate action would be to compare a student’s score to the average performance of a peer group on the test or convert the scale score to a derived score such as a normal curve equivalent (NCE). Both of these options were not feasible due to the lack of additional information in the *ISTEP+ Technical Manual (2012, 2013, and 2014)* from the Indiana Department of Education. According to the manual, the main outcome provided for ISTEP+ or IREAD-3 was categorical data in the form of a performance level. Often, categorical data from a criterion-referenced test does not sufficiently discriminate to assess incremental growth. This type of test is not constructed to maximize score variability (Hambleton, 2009). The results from ISTEP+ and IREAD-3 will be discussed later in this section. Fortunately, Scholastic Reading Inventory (SRI) provided a reading quotient, percentile and a norm curve equivalent (NCE) score to compare to the pre-test results.

T-tests were used to analyze the relationship between the gains in NCE from TERA-3 to SRI for the student sample. The results indicated stronger growth for students in the RTI intervention (M =31.03, SD = 22.02) than those not in intervention (M = 18.83, SD = 20.09), \( t(614) = 7.17, p < .001 \). This statistically significant gain for the intervention group was further analyzed using the Pillai’s Trace test to determine effect sizes. Comparing the gains of the intervention group to the non-intervention groups, Pillai’s Trace revealed a medium effect size.
based on Cohen’s description ($V=0.583$, $F=858.85$, $df=1$, $p<.001$). Visually, Cohen illustrated a medium effect size as one that could be seen with the naked eye, such as the height difference between a 10 year old and 18 year old girl (Coe, 2002).

To further evaluate these results, I converted the effect size to a percentile and a rank contingent to Coe’s report. Using Coe’s interpretation of effect size, an effect size of 0.58 for the gains in NCE suggests that 73% of non-intervention students would exhibit less gain between the post and pre-test than the average intervention student. Next, I ascertained the rank of a student in the non-intervention group who would have grown the same as the average student in the intervention group. Based on the gains in NCE for a group of 25 students, a medium effect size is the same as saying that the average person (i.e. 13th) in the intervention group would exhibit comparable growth to the student who ranked 7th in the non-intervention group. This outcome parallels the analysis of the research by Marston (2005) that 50 to 70% of students receiving Tier 2 intervention responded to the instruction. When recalling that the intervention group was slightly below the non-intervention group in this non-equivalent group design (NEGD), I deduced that the RTI framework in this district had proven effective for students’ academic growth.

Interestingly, a report by Coe (2002) brought to light Cohen’s acknowledgement of the potential danger of using terms in and of themselves to relate effect size. This report asserted that rather than a generic effect size label, an intervention should be construed relative to other interventions that sought to generate the same effect. Noting that the practical effect of an intervention is relative to the benefits and cost of the intervention, Coe stated that an effect size of 0.1 with minimal cost or an easy implementation change could be considered a significant improvement for academic growth.
Considering the concept of practical effect, the interpretation of effect sizes could be seen as a relationship between the current study and the literature. In other words, as Coe discussed, a small effect size may produce a relative influence on academic growth that is important to practitioners. In the prevailing research, the average effect size for student’s performance on a reading test is 0.30 and one study exhibited a 0.63 effect size based on the achievement following a targeted intervention for at-risk students (Coe, 2002). I, as did Coe, wish to emphasize that this illustration was not intended to be a definitive judgment on effectiveness, but instead as a reference point. With respect to this research question, a finding of a medium effect size (0.58) for the gains in NCE from the pre- to the post-test indicate a long-range achievement difference between the groups that is positive for the intervention group. This result is larger than the average student performance on reading test effect size and consistent with the intervention study mentioned previously. Consequently, this finding has important implications for the RTI framework.

To be more specific, the intervention group grew twelve (12) points more than the non-intervention group on the SRI NCE compared to the TERA-3 NCE. According to a CTB/McGraw-Hill document, Interpreting Norm-Referenced Score, NCE was developed specifically for Title I programming. In schools, RTI procedures often follow the regulations for Title I, since Title I frequently funds RTI. In this document, CTB stated that a difference in points from a pre- post-test of 4 NCE points at the district level, 5 NCE for a school, 7 NCE for a class, or 11 NCE for an individual student is an educationally important difference (McGraw-Hill, n.a.). The student growth in this study is educationally important and more specifically, the gains by the intervention group are educationally important. As previously discussed in Chapter 4 Figure 11, a graphical representation of the general linear model comparing the gains for the
two groups, shows a clear trend of closing the gap between the intervention and non-intervention groups.

Based on the method by which the raw scores were converted to scale scores, the state assessment data provided only categorical data to analyze (see definitions for further details). Instead of scaled scores or NCE for ISTEP+ and IREAD-3, pass/fail data needed to be used.

To analyze the data on pass/fail for the pre-test (TERA-3) and post-test (ISTEP+, IREAD-3, and SRI Lexile level), a chi-square test of independence test was the suitable analysis to appraise the data. The chi-square test was performed to examine the relationship between the specific test outcome and subgroup inclusion. For each test, the results indicated that the relationship between these variables was significant. Intervention students were more likely to not pass TERA-3, IREAD-3, ISTEP+, and SRI than the non-intervention students. This would be anticipated based on the RTI premise, which is to help students with academic needs (Batsche et al., 2005; NASDSE, 2006). The positive finding was that the association between belonging to the intervention group and not passing the test was weaker for each of the post-tests than for the pre-test. Put differently, the ability to deduce if a student was in the intervention group as compared to the non-intervention group would have been easier on the pre-test than on post-test. A relationship still existed, but it was smaller.

Not only was the association reduced, but the likelihood of passing on each of the post-test was greater for both groups than on the pre-tests. Again, the gap is narrowing between the intervention students and their peers. While the risk was lowered for the intervention group more than the non-intervention group, the pass/fail results do need to be interpreted with caution, because pass/fail data does not discriminate well. With that, these results match the earlier results in this study and seem to be consistent with the preliminary research on RTI.
implementation (Fuchs & Fuchs, 2001; Fuchs, 2003; Fuchs & Deshler, 2007; Flippo, 2001; Kovaleski, Prasse, 2004; Lopez, 2009; Pikulski, 1994; Vaugh et al., 2007; Ziolkowska, 2007).

The final analysis for the long-range achievement differences between the groups was a comparison of Odds Ratio (OR) for each assessment. For each assessment, the odds of not meeting the academic standards are greater for the intervention group; however, this difference between the two groups is reduced on each of the post-tests. It is also noteworthy to mention, that the rate of not passing reduced for each post-test in comparison to the pre-test. For example, an OR of 4.10 on the TERA-3 pre-test would not seem drastically different from an OR of 3.78 on the IREAD-3 post-test, if it were not mentioned that the pass rate for the IREAD-3 was approximately 90% as compared to the TERA-3 at 35%. Although these findings do not show that the intervention group meet or surpassed the non-intervention group, the difference between the groups was reduced. This reduction would seem to align with the results of the gain in NCE discussed previously and may help to develop a broader picture of the long-range achievement differences between the intervention and non-intervention groups for this research study. Considering all of this evidence, it seems that the conceptual framework of the RTI framework has a relationship to the long-range academic growth of the intervention students in my study.

Conceptual Framework. Instruct, assess, and analyze are steps of the cyclic decision-making approach for which the conceptual framework of Response to Intervention is based (Fuchs & Fuchs, 2006; Heller, Holtzman, & Messick, 1982; Linan-Thompson, Vaughn, Prater, & Cirino, 2006; Reschly, 2002). The fluid nature of this inclusive model would have seemed to have provided the much needed time and instruction that many students with academic reading struggles required. One example of meeting individual student’s needs would be the subgroup of students who were exposed to two years of intervention. Through the cyclic decision-making
approach, students who were in need of additional support were provided additional intervention based on the annual universal screener. By beginning with universal screeners, the educational team was provided rich data whereby the team could make sound decisions. These decisions were supported throughout the process with ongoing progress monitoring to continually reflect on the students’ needs. Through ongoing reflection, a course realignment could occur to attempt to maximize the benefit to each student. The student results in my study would seem to support the benefits of the cyclic decision-making approach.

In addition, the focus groups’ overall perceptions of the RTI process were largely optimistic. One focus group member said, “We appear to be targeting the right kids.” The general view of the focus group participants was supported by the student achievement data for intervention students compared to non-intervention students in this study. Not all comments were positive. A few statements of concern were noted for the lack of growth by individual students. The goal of most educators is to meet the needs of each student and frustration can arise when challenges continue. Considering these outcomes, the cyclic decision-making approach of RTI would seem to have supported students (Fuchs & Fuchs, 2006; Shapiro, 2011) and provided the necessary well-founded confidence required by staff when implementing a program (Fashola & Slavin, 1996).

**Section summary.** Taken together, these results suggest that the intervention group substantially closed the gap between themselves and their peer group who scored just above them; however, the intervention group did not surpass these non-intervention peers. Returning to the research question posed, it is now possible to state that the evidence presented thus far supports the concept that RTI provides academic benefits to students. This research study supports the literature on early intervention through the RTI framework aiding academic
Research Question 2: What differences occur among the RTI Tier 2 groups as compared to those students just outside the intervention group in regards to special education identification by the end of third grade?

While the results from my first research question disclosed substantive gains in the intervention group compared to the non-intervention group, the focus of research question two did not prove as fruitful. Unfortunately, this research question could not be answered as well as initially planned for various reasons. First, data collected for this study did not include the process by which a student was initially referred for a disability. In addition to the evidence-based method required of schools, IDEA (2004) permits parents to request a referral. In this study, further analysis to discriminate between parent and RTI referrals would need to be conducted to deduce any change in identification based solely on RTI. Unfortunately for this research, gathering this data on the historical cohorts of students was beyond the scope of the study. Second, the historical special education enrollment data was not specific to SLD. However, this data did provide an overall picture of change. As mentioned previously, most areas of special education typically do not fluctuate drastically over time with the noted exception of autism and SLD (NCES, 2009; Reschly, 2002). SLD and autism have both been on the rise. The implication of historical numbers will be presented later in this section. Notwithstanding these issues, the scope of the RTI program in respects to changes in identification of special education process was scrutinized through multiple data sources.

By the end of third grade, 18.3% of the students in this study were identified as needing an IEP for SLD. Although this percentage may seem high, it is important to understand that this percentage represents only students in this study and not the entire grade level. As a point of
reference, the district statistics indicate 16.4% of students have been identified for a special education, which aligns with state (14.6%) and national (13.0%) percentages.

Continuing with this finding, statistically significant differences were found between identification for special education and group inclusion ($\chi^2 = 28.58, df = 1, p<.001$). Specifically, 13.5% of the study sample had been in kindergarten and/or first grade intervention. In other words, three-fourths of the students who were identified for special education had been in intervention. This was not surprising considering the structure and intent of the RTI framework. Indiana’s premise for RTI is that each student should be provided a continuum of services to ensure meeting or exceeding the proficiency standards (IDOE, 2010). In fact, the surprise was that the association was not higher given the student selection process for intervention. As a final point, the use of intervention, specifically RTI, relates to the findings in the literature on evidence-based methods for determination of special education needs (Shapiro et al., 2011; Vaughn, Wanzek, Woodruff, & Linan-Thompson, 2007).

For further analysis, I used Chi-square to understand whether school or year of implementation were related to identification for special education. All combinations, with the exception of School B and implementation year 2011, were found to be statistically significant with a small association between participation in intervention and identification for an IEP for SLD. Again, under the RTI framework, this would be expected. With regards to the lack of significance for the 2011 implementation year, this year included only first grade students and the sample size was reduced by the removal of students who were also in Cohort 3. Of the individual data sets, this set of data would include the smallest sample size and be more prone to irregularities. Based on the focus group input for School B, a possible explanation for the lack of significance would be parent referrals. The administrator focus group reported attempts to
persuade parents to permit the RTI process to drive the SLD referral; nevertheless, not all parents agreed. It was the perception of the teacher groups that the principal’s leadership style impacted the number of parent referrals that were redirected to the RTI or continued as a parent referral.

These statements are not meant to malign the parent referral option or the principal for his/her leadership style; it is simply stated to provide an alternative explanation based on the qualitative data. In fact, a positive aspect of the parent referral relates to a drawback of the RTI framework. When a student is performing at grade level, but has an area of weakness, this student is often overlooked due to the ranking process. According to a member of the administrator focus group, this could be the case in this district.

Finding a relationship between various factors and inclusion in special education, next I examined the effect of various factors on academic gains of students with and without a special education identification. Two-way ANOVA tests were run. A variable in each test was identification for special education. The counter variables explored were gender, SES, or entry grade. No relationship was found for any of the counter variables with identification for special education in regards to gains on NCE. A cautionary conclusion would be that the district saw no disproportionality in the identification for a SLD based on gender or SES. In other words, the disproportionate number of males seen under the discrepancy model (Bocian, Beebe, MacMillan, & Gresham, 1999; Coffey & Obringer, 2000; Cortiella & Horowitz, 2014; Reschly, 2002) does not seem to be present under the RTI framework in this study. The same positive outcome would be true for lower poverty students. These findings aligned with the research on the RTI framework (Hintze, 2008; Shapiro, et al., 2011).
Finally, comparing the percentage of students in special education prior to (2005-2007) and post-RTI implementation (2011-2013, years cohorts were in 3rd grade), it would appear that the overall special education population for the district, the district’s third grade, School A, and School C all decreased. The district reduced 2.1% and district third-grade special education percentage decreased 4.9%. However, School B increased slightly (1%). As mentioned previously, these percentages included all categories of special education; however, with the exception of autism, which has been on the rise, special education categories typically do not fluctuate over time (NCES, 2013; Reschly, 2002). The special education administrators in the focus group agreed that the changes within the district appear to be a reduction in SLD and an increase in autism identification. While this supports the assumptions of RTI, it does not provide irrefutable evidence for this conclusion. All that could be said would be that the downward trend in special education percentages for the district parallels the implementation of RTI in the school district. Further data would be needed to substantiate this statement.

**Conceptual Framework.** Going back to the basis of the RTI framework, the cyclic decision-making approach has been recognized as having the potential to intervene and therefore reduce the number of students identified as having an SLD. This potential was recognized at the national level between 2004 and 2011 with a reduction of 0.9%, or 400,000 students (Cortiella & Horowitz, 2014; U.S. Department of Education, 2013). This analysis implies that the cyclic decision-making approach has impacted the percentage of SLD in the study district as well. The structures of this inclusive model would seem to have provided the support many students required.
Section summary. In considering the result, the parent referral method did seem to obstruct the results for this research question. With this being said, the dual referral method does not negate the benefits of the RTI framework. Even with a parent referral, decision-making can be based on student data, intervention occur earlier, and the percentage of special education students in the district was reduced. A member of the administrator focus group commented that when a parent refers their child for identification, we have a wealth of quality data to review compared to the prior methodology. Mesner and Griffith (2005) found that assessment data in the RTI framework shaped instruction to better meet student needs. The literature provides strong evidence that the use of data can help to prevent reading difficulties (Denton, 2012; Fuchs & Fuchs, 2006; Fuchs & Vaughn, 2005; Gersten et al., 2009; Hale et al., 2013; National Reading Panel, 2000). In addition, the use of data overcomes one of the challenges of the discrepancy model to help eliminate the widespread arbitrariness of SLD identification (Kavael, 2002; Fuchs & Fuchs, 2006; Vaughn & Fuchs, 2003). Considering both identification methods used the formative assessment data, the components of the RIT framework provided valuable information to support the identification of students with a learning disability.

The focus group member also noted that in many cases a parent referral would prompt earlier support for their child than under the old method. Research has shown that explicit, intense, and monitored, early intervention increases the achievement rate and provides a stronger likelihood of academic success (Flippo, 2001; Fuchs & Fuchs, 2006; Gersten et al., 2009; Lopez, 2009; Vaughn et al., 2006; Vaughn et al., 2007). Early intervention supports another flaw in the discrepancy, or “wait-to-fail”, model (Berkley, et al., 2009; Fletcher, et al., 2004; Meyer, 2000). Each of these beliefs would be supported by the literature on the benefits of early intervention and the RTI framework (Brown-Chidsey & Steege, 2005; Danielson & Doolittle, 2007; Flippo,
RTI: DIFFERENCES IN READING ACHIEVEMENT

2001; Fuchs & Fuchs, 2006; Hintze, 2008; Lopez, 2009; Mask & McGill, 2010; Shapiro, Zigmond, Wallace, & Marston, 2011; VanDerHeyden, Witt, & Gilberson, 2007; Vaughn et al., 2011). In general, the majority of students in this study who were identified with an SLD participated in intervention and those who did not benefited from the established procedures.

Despite the fact that the evidence is not extensive, indications would point to the RTI framework being a contributor to the reduction in special education. This question looked at the differences from RTI based on one subcategory, special education. Other subgroups have been influenced by the discrepancy model along with special education and will be discussion in research question 3.

**Research Question 3: What differences are seen in students' reading achievement among and within the various nested groups (e.g. school, implementation year, achievement measures)?**

Expanding the analysis beyond identification for SLD, the underlying question was whether membership in a subgroup indicated a greater likelihood of weaker academic achievement relative to the group’s counterpart. I analyzed the data based on inclusion in subgroups and academic gains for gender, SES, school, implementation year, and entry year. As a reminder, ethnicity was not investigated due to the low incident of minorities at the school district. In interpreting my findings, I realize that no statistical difference between groups would mean that the subgroups were similar or exhibited analogous outcomes. Conversely, a statistical difference may indicate bias within the nested group. I will first discuss the statistical findings by placement within the nested group followed by academic achievement by nested group.

Gender results in this study indicated that number of males and females were generally balanced in regards to placement in the intervention and non-intervention groups. The Fisher’s
exact test indicated no statistically significance difference beyond that expected by chance between gender and placement in intervention or non-intervention groups ($p = .177$). In addition, the difference between the genders by school or by cohort was found to be not statistically significant ($p > .05$). The one exception was in Cohort 2 ($p = .04$), where 64% of the intervention sample consisted of males compared to 47% of the non-intervention sample being males. In general, no gender bias for presence in intervention group appeared to transpire. This positive outcome for a lack of a gender bias would be counter to much of the literature that reports more males have academic difficulty in reading than females (Connell & Gunzelmann, 2004; Gurian & Stevens, 2004). This discrepancy prominent in the literature base would be substantiated by the achievement data required from the federal No Child Left Behind (NCLB). In fact, the focus group commented that they were surprised by the balance between boys and girls. It was their feeling that more males were involved in intervention. Thoughts on this perception will be discussed in the research question 4. Overall, the number of males in the intervention group was statistically equivalent to the number of females. With equivalent enrollment in intervention, I now shift to the academic outcome based on gender.

An analysis of NCE gains from TERA3 to SRI resulted in no statically significant difference between genders ($F(1, 614) = .125, p = .724$) with a mean gain for males of 25.4 and females of 24.8. From this result, I examined the academic gains with gender and group inclusion to find no evidence of interaction effect between the variables. In other words, any difference in academic gains between groups was not dependent upon which gender the subject was and that any differences between females and males were not dependent upon which group the student participated in. These results are somewhat unexpected as the literature base would indicate that girls outscore boys in reading. In discussing this result, a member of the focus
group commented that the activity-based learning in intervention groups may have helped the boys’ engagement and with increased engagement came increased understanding of the content. Guigon (2010) and Hattie (2011) both found that teaching in multiple ways had been shown to improve learning for both boys and girls. Another focus group participant mentioned that the district had been intentional about adding to its non-fiction book selection and the current reading series provided a wealth of non-fiction stories, which would typically engage boys more than fiction. While I have no definitive evidence, one proposition for the balance in academic gains would be the classroom structures and the nature of the intervention programs used within the district. A few components of the district’s instruction that would support diverse learning styles would include Minds in Motion®, activity-oriented instruction, centers, and technology integration. In addition, the study district held interventions in small groups and the various intervention programs used (e.g. Zoo Phonics®, Visualizing and Verbalizing®, and PASS Phonological Awareness Program®) provided a substantial amount of hands-on activities. Activity-oriented instruction provides a balance of auditory, visual, and kinesthetic interaction, which can be beneficial to boys, as well as girls (Conlin, M 2003; Gurian & Stevens, 2010; Jantz, G.L., 2014). This type of instruction may align better with more children’s learning style and help them maximize their own learning.

Statistical analysis confirmed that students in the low SES category and intervention co-occurred more frequently than expected by chance ($p < .001$, Fisher’s exact test). In other words, placement in Tier 2 was greater for lower poverty students than their counterparts. However, considering each cohort separately, only Cohort 2 exhibited statistically significance differences for SES and group inclusion. When comparing schools, School C did not demonstrate statistically significance differences, but School A and School B did show a
statistical significance when comparing differences of means ($p_A = .037$, $p_B = .027$, $p_C = .083$, Fisher’s exact test). Since SES is the single best predictor of average student performance (Jensen, 2009; Reardon, 2013), School C would appear to have reduced the stereotypical poverty relationship with academic struggles. The focus groups commented on the role that socio-economic status appears to play in student preparedness for school. They questioned how we, as educators and as a country, could intervene with children earlier.

Academic gains were subjected to a two-way analysis of variance having two levels of group inclusion and two levels of SES. No statistically significant differences were found between the groups. This means no interaction effect was found between the independent variables in relation to the academic gains. In other words, it can be said that there is no evidence that the academic gains varied with the level of SES in intervention or non-intervention. While much of the research base indicated that SES is the single best predictor of average student performance (Jensen, 2009; Reardon, 2013; Silvernail, Sloan, Cleelsea, Johnson, & Stump, 2014), in this study the results would indicate that this potential barrier of poverty can be reduced through the intervention process. This research is not without precedent. Several researchers in the past have found similar outcomes (Fuchs, et al., 2003; Marston, 2005; Shapiro, Zigmond, Wallace, & Marston, 2011; Telzrow, McNama, & Hollinger, 2000; Torgesen et al., 2001; VanDerHeyden, Witt, & Gilbertson, 2007).

The challenges that children living in poverty face in school can be overcome. Ching (2012) stated that low income should not mean low expectation or it can lead to a feeling of marginalization, a lack of educational achievement, and even dropping out. The unique challenges faced by lower-income students must be respected, while not becoming an excuse (Jensen, 2009). Poverty does not have to be a detrimental trait for a child. A study in Maine
found that while poverty level was a factor, other factors that could be controlled by the system included the type of school, teacher experience, teacher educational level, students’ connection to school, and personal connections with students (Silvernail et al., 2014). Reardon (2013) concluded that if the discrepancy was not rectified, schools would no longer be the great equalizer. Some schools are defying the factors for the benefit of students.

Lastly, I analyzed the achievement gains (NCE SRI: TERA-3) against various nested groups, such as entry grade. A statistically significant difference was found within the year of implementation and among the cohorts, but not among schools or entry grade. Specifically, a mean difference was found between the first and fourth year, as well as, the second and fourth year of implementation. The first and second year showed statistically higher gains than the fourth year. Next, Cohorts 1 and 2 were each shown to have a mean difference compared to both Cohorts 3 and 6. Similar to implementation year one and two, Cohort 1 and 2 resulted in statistically higher gains than Cohort 3 and 6. No other statistically significant mean differences were found by year of implementation or by cohort. Interestingly, the between-school analysis indicated no significant difference in academic gains, nor were any differences found by entry grade. Before considering the potential implications of these results, it is important to recall the research cohort structures, Cohort 1 was year one of implementation. Cohort 2 and 3 were part of implementation years two and three, respectively. Each of these cohorts began in kindergarten. Cohort 4, 5, and 6 were first grade cohorts from implementation years two, three, and four, respectively.

Interpreting these results indicates that the execution of the RTI framework among the schools was consistent based on no statistically significant differences in means. In other words, the schools all performed comparatively across the district and it did not appear to matter if
students were in intervention in kindergarten or first grade. Since the RTI framework structure was developed as a district protocol, this outcome would be indicative of a uniform program implementation across the district. For the purpose of my study, the evidence of uniform program implementation aids in reducing study limitations by essentially providing three parallel studies.

When examining implementation years (or cohorts), the statistically larger gains for the initial years, year one and two (Cohort 1, 2, and 4), over the later years, year three and four (Cohort 3, 5, and 6), may be indicative of program fidelity. It is noteworthy that the district administrators were more involved during the first and second years of implementation where gains were larger. The focus group discussed several key factors including implementation fidelity and administrators’ focus that may have influenced program outcomes. An assumption might be that program fidelity was stronger with a higher level of administrative oversight. Research does indicate that fidelity of delivery plays a strong role in maximizing program effectiveness (Foodman & Moats, 2004; Fuchs & Fuchs, 2006; Gresham et al., 2000; Instructional Research Group, 2006; NASDSE, 2006; Telzrow, McNamara, & Hollinger, 2000).

An interesting point to consider would be the lack of a statistically significant difference in academic gains between placement in kindergarten or first grade. Research on early intervention would indicate the earlier the intervention can occur the better (Flippo, 2001; Kovaleski & Prasse, 2004; Lopez, 2009; Pilulski, 1994; Vaugh et al., 2007; Ziolkowska, 2007). One point to consider would be the impact of kindergarten readiness, or background knowledge (Hintz, 2008; Leverett, 2006; Lunsford, 2015). In the case of the kindergarten cohorts, were some of the intervention students in need of remediation due to a lack of or poor background experiences? Whereas, were some of the first grade students better prepared initially and their
struggles were exposed as the curriculum became more challenging? Research would indicate that background knowledge has a high correlation (.66) to learning new materials (Marzano, 2004). In many ways, these questions support the need for the RTI framework. The framework advocates for intervention at the time of need.

**Section summary.** In my review of the nested group results, it was established that membership did play a role in placement, specifically for SES. That being said, by the end of the study, the achievement gains were slightly higher for the lower SES group than the higher SES group. In general, this evidence adds to the merits of other researchers that the social injustices can be overcome. In considering gender, positive results were seen for both the placement in intervention and academic gains. These outcomes support the converging evidence endorsing the benefits of the RTI framework, specifically in the primary grades (Fuchs & Fuchs, 2003; Fuchs, Fuchs, & Speece, 2002; Denton, 2012; Marston, 2005; Torgensen, 2004; Wanzek & Vaughn, 2007).

**Research Question 4: From the perspective of the focus group, what understandings result from the quantitative data analysis of the academic outcomes between and among the nested groups?**

The overall perception of the focus group participants was that the RTI process showed an impact on student achievement. One participant simply said, “The time and effort the staff and students are spending on the RTI process is working.” A student’s need for a different amount of time to learn and a program that meet the student’s individual learning style, unique needs, and is in the zone of proximal learning for that student supports two key learning tenants. Tomlinson’s (1995) work on differentiated instruction supports student’s individualized instructional needs based on time, program design, and learning style. Additionally, Wiggins
and McTighe’s (2000) Understanding by Design, UbD, framework promotes planning for students with the end goal in mind. The end goal for the students in this study would be reading on grade-level by the end of third grade. Just like the RTI process, UbD would promote providing instruction within the zone of proximal learning and planning based on knowledge of the students and their goal. Currently in the education world, educators are encouraged to differentiate instruction, to plan based on the end goal, to utilize formative assessment for decision-making, and to integrate a range of tools. The intent is to provide more rigorous, thought-providing work for students. One tool that would seem to support students learning would be the RTI process. With respect to the first research question (i.e. long-range differences on reading achievement), the focus group believed that the RTI process was helping the district to meet the goal of all students reading on grade level by the end of third grade.

Beyond the specifics of student results, a theme that emerged from the focus groups concerned logistics. The focus group understood that the logistical components of RTI related to the importance of a systemic approach. In this research study the district not only implemented RTI school-wide, but district-wide with support from district administrators, financial support, ongoing professional development, consistent progress monitoring tools, and additional staff. The focus groups concluded that the critical logistical components were the quality of Tier 1 instruction, intervention group size, organization of a program with student needs, and various staffing aspects. Each of these logistical components required a team approach to manage the organizational change, to implement effectively, and to maintain fidelity. According to research on program implementation, the nature of these logistical components merit a school-wide implementation (O’Connor, Harty, & Fulmer, 2005; Fuchs & Fuchs, 2006; Vaughn et al., 2009) as seen in the study district.
Quality instructional delivery in Tier 1. Fundamental to the integrity of the RTI process for identification of special education needs would be the quality of the instructional delivery in Tier 1. Consistent with research, the practice of high-quality, research-based instruction is one of the underlying principles of RTI. This principle has made RTI a more viable approach to identification for learning disabilities than the discrepancy model (Conner, Pasta, Fishmans, Glasney, Schatschneider, Crow, 2009; Denton, Vaughn & Fletcher, 2003; Foorman, Francis, Fletcher, Lyon et al., 2001; NASDSE, 2006; Sharpio et al., 2011; Snow, Burns, & Griffin, 1998). Several focus group participants talked about the critical nature of providing a guaranteed viable core instruction by a highly-qualified teacher.

The discussion lead to several aspects of the district’s programming that provided evidence of quality core instruction. The evidence mentioned comprised student scores on summative assessments, student growth data, teacher evaluation, and teacher qualifications. The focus groups perception is substantiated by the quantitative data on the cohort group. Pursuant to the literature, high-quality core instruction should support approximately 80% of students (Fuchs & Deshler, 2012; Fuchs & Fuchs, 2006; NASDSE, 2006; Shapiro, 2011). In this study, the intervention groups represented approximately 20% of the grade-level population. According to the research, the quality of the instructional delivery impacts the number of false positives yielded from poor teaching resulting in low achievement and causing misidentification of students with SLD (Donavan & Cross, 2002; Fletcher et al., 2004: Fuchs & Fuchs, 2006). Based on the literature, the district in this research provided quality Tier 1 instruction and therefore the false positives should be somewhat negated.
In addition, a viable core program is critical because a poor program would yield non-responders that would be easier to remediate in Tier 2 (Bean, 2008; Gersten & Dimino, 2006; Hill et al., 2012; Newmann, et al., 2011) and skew the results. The core instruction is the foundation for learning and as indicated by research, the teacher is the most significant predictor of student achievement (Allington, 2005; Darling-Hammond, 1999; Denton, Vaughn, & Fletcher, 2003; Sharpio et al., 2011). Measuring a student’s response to Tier 2 instruction can only be deemed reliable if the core instruction is presumed adequate (Stecher, Fuchs, & Fuchs, 2005).

In summation, the focus group participants proposed that strong Tier 1 instruction provided a solid foundation for the interpretation of student data for academic need for Tier 2 and later for any learning discrepancy that may lead to identification for a disability.

**Intervention group size.** The focus group members recognized that the structure of small groups was appropriate, as evidenced by the student growth outcomes mentioned previously. This belief is supported by the research on effective group size from Prime Time and other early interventions (Denton, 2012; Schwartz Schmitt, & Lose, 2012; Wanzek & Vaughn, 2007). The research indicates that the smaller the group the higher the gains. In addition, the focus group, along with prior studies, noted the importance of the group size being determined by the specifics of the program (Denton, 2012; Schwartz, Schmitt, & Lose, 2012; Wanzek & Vaughn, 2007). Since the overarching goal of RTI is student achievement, the focus group perspective, as well as the research, indicate that appropriate group size is critical to support quality programming. In closing, small intervention groups, while not specifically defined, are supported in the literature (Denton, 2012, Fuchs & Fuchs, 2006, Schwartz, Schmitt, & Lose, 2012, Wanzek & Vaughn, 2007).
**Organization of the program to meet student needs.** Turning to program organization (i.e. selection of students and placement of students in a program), the focus groups discussed the thoughtful process by which the RTI team (i.e. classroom teacher, special education teacher, paraprofessionals, reading coach, principal, and the school psychologist) assigned a student to an intervention group based on the student’s data, the teacher’s understanding of the student’s needs, and program options. Several participants mentioned the attention to detail based on program knowledge, other students included in the various groups for best personality fit among students, and the instructional aide’s (IA) teaching style. It is notable that the value of the reading coach’s leadership in this decision-making process was commented on several times.

Each of these aspects of organization are supported by research (Batsche, Elliot, Graden, Grimes, & Prasse, 2000; Denton, 2012; Fuchs, Fuchs, & Speece, 2002; Fuchs, Mock, Morgan, & Young, 2003; Strangeman, et al., 2006).

The implementation and decision-making process executed at this district was a hybrid protocol. A hybrid protocol blends problem-solving and standard treatment methods. In looking at the research on both aspects of a hybrid model, studies indicated that standard protocol provides easier fidelity and the problem-based protocol provided improved alignment to students’ needs. O’Connor, Harty, & Fulmer (2005) found that by using a standard protocol the outcome for intervention students as compared to historical peers exhibited improved reading achievement. In this study, the outcome for intervention students showed an 84.7% improvement in students’ reading levels. In looking at the growth in NCE of the intervention students as compared to the non-intervention students, the intervention students saw a 12 point growth over the non-intervention students. Other studies indicated that between 53 and 84% of students in standard treatment protocols improve reading scores (Marston, 2005; Vaughn, Linan-
Thompson, & Hickman, 2003; Vellutino, Scanion, Zhiang, & Schatschneider, 2008). The St. Croix River Education District Model study resulted in student reading achievements improving from 35% to 70% of students meeting the CBM cut scores after following a problem-solving protocol. The benefits provided by the hybrid model include a moderate level of quality control in the program structure with sensitivity to individual needs (Fuchs, Mock, Morgan, & Young, 2003). Through the thoughtful process, attention to details, and leadership, students thrived in the hybrid model. This model suited the district in this study as well.

Various staffing aspects. Appropriate alignment of staff members with student RTI Tier 2 groups, ongoing professional development for staff, and leadership involvement were noted as important aspects of the RTI logistics by the focus group participants. The student results from this research indicate a relationship between inclusion in intervention and academic growth. Accordingly, one may assume the program structures, such as these staffing aspects were appropriately aligned to benefit students. Each aspect provides insight into the development of a consistent program that provides fidelity and quality work.

Matching the staff member to the student Tier 2 intervention groups. All three teacher focus groups commented on the essential nature of aligning the instructional strengths of the IA or teacher to not only the instructional program for the Tier 2 group, but to the student intervention group. Research would indicate that a paraprofessionals (or IAs) can deliver effective intervention if a careful alignment of the program, interventionist, and group were considered (Denton, 2012). One example given was a specific group of intervention students, primarily boys, who needed a firm yet compassionate hand. In this incident, a new IA was placed with a group who was not seeing growth based on progress monitoring. Within a matter of a few weeks, the group enjoyed tremendous growth to the extent that the program was
modified to meet the students’ new needs. Aligning staff to the program and student group is supported by research on program adjustments (Fuchs, Fuchs, & Speece, 2002; Fuchs, Mock, Morgan, & Young, 2003).

**Ongoing professional development for Tier 2 instructors.** Participants concurred that since support tools have been implemented, the instructional aides have enjoyed greater efficacy. Support has come in many ways: professional development, reading coach guidance, weekly planning time, and strategies to promote positive behavior outcomes. A common view among focus group participants was that the paraprofessionals have taken ownership of their groups and the RTI process. A common view among focus group participants was one of ownership of their groups and the RTI process by the instructional aides. Greater efficacy had led to stronger voices, increased buy-in, and a school culture viewed as a positive learning environment.

Professional development, both formal workshops and ongoing support by a coach, has been shown to have a large significant effect size on the quality of instruction (Nelson-Walker et al., 2013; Nicholson, 2009) and should be extended to paraprofessionals (Hauerwas & Goessling, 2008). Tileston (2009) stated that a critical component of the RTI framework is ongoing support in the form of professional development for all staff. Bailey’s research (2010) on teachers’ perception of RTI provided clear evidence that districts must invest their resources into professional learning. The understanding of the focus groups related to the value of professional development and other forms of professional learning environment to increase efficacy, enhance buy-in, and improve instructional delivery. Leadership must understand that ongoing professional development not only benefits struggling students, but can benefit all students.
Leadership. Moreover, leadership is integral to a school-wide program implementation (Fullan, 2001; Mazano, Waters, & McNulty, 2005). The focus group found several layers of leadership within the RTI program implementation. Leadership that impacted the implementation of the RTI program included district administrators, building administrators, reading coach, and teachers who took ad hoc roles as leaders within their grade level. The value of the teacher and reading coach leadership were both declared in previous sections. It was suggested by the focus group that the administrative team supported the culture change necessary to implement a program of this magnitude and maintain fidelity to the structures. In two different meta-analysis studies, the correlation between academic achievement and building administrators is strong, second only to teachers (Leithwood, Louis, Anderson, Wahlstorm, 2004; Mazano, Waters, & McNulty, 2005). Marano, Waters, ad McNulty (2005) pointed out that strong collaborative leadership has been shown to be vital to initial implementation and maintaining integrity over time. As evidenced by the focus groups in this study and a number of recent studies (Fullan, 2001; Marzano, Waters, & McNulty, 2005; Silins, Mulford, & Zarins, 2002), each of these formal and ad hoc leaders contribute to a school community with their willing to take risks, make decisions, and provide quality instruction for students.

Summary perception. The perception of the focus groups was that the current implementation of RTI within the district was beneficial to students. Their understanding was that decisions should be deliberate and based on student data as well as staffing facets. Through data-based decision-making, students’ learning could be supported and appropriate identification of students with a specific learning disability can occur. Research on teacher perception of RTI found that teachers wanted to support their students and were willing to do whatever it took (Bailey, 2010). This perception is also supported through the research on RTI.
The research indicated the impact on the learning environment requires a cultural change that needs the support of the teachers and leadership as well as professional development (Greenwood et al., 2008). Each of these aspects is related to the need for a school-wide program.

**Conclusions**

Under the direction of the discrepancy model, gender, ethnicity, and socio-economic status impacted identification for a disability (Coffey & Obring, 2000; Fuchs & Fuchs, 2006). A specific learning disability (SLD) is when a student exhibits a deficit in at least one psychological processing area (LDA, 2010). In other words, a discrepancy or difference in a student’s ability and the student’s academic achievement exists. Based on the definition of the discrepancy model, this in turn meant that gender, ethnicity, and socio-economic status impacted academic achievement negatively. In this study, students who received intervention grew significantly more than their peers who scored just above the cut for intervention and did not receive intervention. This research adds to the literature indicating that the RTI framework supports students’ learning through early intervention and can reduce the previous discrepancies found based on group inclusion.

Even though a relationship was noted between initial placement in intervention and inclusion in lower poverty groups, the positive outcome was that by third grade the academic gap was substantially reduced when considering social economic status. While the intervention group did not surpass the non-intervention group, the students made substantial growth and were closing the gap. These results counter earlier research on disproportionate patterns for referrals for specific learning disabilities (SLD) under the discrepancy model (Becian, Beebe, Fuchs & Fuchs, 2006; Coffey & Obringer, 2000; Cortiella & Horowitz, 2014; MacMillan, & Gresham, 1999; Sharpio et al., 2011; Vaughn et al., 2009). As mentioned previously, the RTI framework
provides just-in-time interventions that support individual student’s academic needs. A plethora of research supports early intervention. Studies advocate not only for early intervention, but that intervention should be no later than first grade (Bursuck and Blanks, 2010; Velluntino, Scanion, Zang, & Schatschneider, 2008). As seen in this study, there were significant results by providing targeted, explicit instruction through ongoing evaluation by formative assessment.

One specific area of concern that I want to note is the issue of social justice relative to poverty. Educators need to be conscious of preventing a student’s socio-economic status from impacting his or her educational outcomes. The academic gains for lower poverty children in this study provided evidence that positive growth can be achieved. As educators, and as a nation, we need to continue to address this issue to make sure every child has the opportunity to access the knowledge and skills to improve life opportunities.

The original problem posed in this thesis discussed how our society is dependent on a literate population to drive the economy, as well as, social forces (CERI, 1992: EnGauge, 2003). Response to Intervention (RTI) represents the much needed change in the process of supporting struggling students (Fuchs & Fuchs, 2006; Gersten et al., 2009; Hintze, 2008; Lopez, 2009; Nelson-Walker, Fien, Kosty, Smolkowski, Smith, & Baker, 2013; Shapiro, Zigmond, Wallace, & Marston, 2011). Practitioners benefit from quality research to support implementation. In this study, the appropriate alignment of staff, ongoing professional development, and leadership were noted as important aspects of RTI. The hybrid model, problem-based structure with standard treatment protocols, appears to have worked well to increase student achievement. Though this study helps to fill the gap in research, further studies are necessary to provide educational leaders with a solid foundation for best practices to meet all children’s needs.
Implications for Action

The nature of this RTI implementation makes this study more difficult to replicate as compared to the well-defined procedures provided by a standard treatment protocol. However, the reality of the educational system is that a problem-solving model or a hybrid model is more relevant than the more traditional clinical procedures. The current landscape supports the need for ongoing research in this area (Strangeman, et al., 2006).

The following recommendations are offered as possible ways to improve this study:

1. The data was rich, but would have benefited from a longer time period.
2. The study explored implementation from both kindergarten and first grade cohorts. In retrospect, only one grade would have been prudent. A concern arose from the elimination of duplicates for the first grade cohorts because removal limited the data set size. By studying a single grade level, the duplications would have been avoided.
3. Additional special education details had been collected, such as method of initial request for special education placement or further screening data to differentiate the false positives and false negatives. It would have been valuable to have the referral method, RTI or parent, to aid in discerning the differences in identification specific to the RTI framework.
4. If state standardized test data would have been available, I would have been able to use additional inferential statistics beyond the pass/fail results. The normed data would have permitted better discrimination of students’ scores to allow for more detailed analysis and comparisons.
5. A standardized test, such as NWEA data to measure growth throughout the process as both pre- and post-test assessments would have permitted better
discrimination of students’ scores to allow for more detailed analysis and comparisons.

6. If other factors, such as intervention group size, the classification of the interventionist, or the intervention program, would have been included in the data additional nested groups could have been analyzed.

7. Instructional alignment between tiers had been studied. An underlying assumption of RTI is that all students received high-quality core instruction. To determine outcomes based on Tier 2, it is assumed that Tier 1 provided a solid instructional foundation. In this study, I assumed the strength of Tier 1 based on the pass rates for the post-tests. Further evaluation of Tier 1 would provide additional support for the study.

8. A focus group was not conducted with instructional aides. Instructional aides who facilitated the Tier 2 interventions may have added additional insight if they had been included in the focus group, most likely as a separate group to permit a comfortable single layer hierarchy conversation.

**Recommendations for Further Research**

The following recommendations are offered for future research:

1. Given that this study provides a basis for concluding that appropriate intervention can reduce the traditionally perceived barriers of gender and SES, defining the attributes that constitute delivery of a quality RTI framework would prove to be of value to the discipline. Attributes to consider studying would include the dosage (i.e. cycle duration, session frequency, and session length), leadership aspects that prove critical to maximize growth, and the appropriate group size that provides strong student growth, but is also feasible in the school setting. Such an
effort would enable educational leaders to derive protocols for RTI from a research base.

2. Another recommendation would be to conduct a similar study at more school districts, thus permitting further generalizability of the findings. Additional patterns may surface from a study that spanned several school districts in a region.

3. Following a similar research structure, the addition of questions similar to VanDerHeyden, Witt, & Gilberton (2007) would be useful. The primary question they studied concerned the effect of RTI on the total number of referrals and the percentage of qualifying evaluations.

4. More research is needed on the factors impacting the execution of the RTI framework, as different criteria are likely to result in different student growth outcomes. My recommendation would be to initiate field studies rather than clinical studies. Shapiro, Zigmond, and Marston (2011) concurred that the core of RTI has a plethora of research; however, there is a gap in the research on real-life implementation of RTI. Practitioners would benefit if this gap would be filled with quality research.

5. Given the federal and state mandate for high-quality research-based interventions and the lack of progress under the reign of the discrepancy model, a series of longitudinal studies would improve identification of a specific learning disability (SLD). These studies, based on this model, would consider the definition of adequate and inadequate response, or growth, to intervention. Additional research would lead to improved guidance for practitioners, a much needed aspect. The most critical consequence of improved execution would be increases in student
achievement and improvement in identification of students with SLD.

**Recommendations for District Practitioners**

The following recommendations are offered for practitioners, specifically at the study district, but more generally for educators at-large. A subgroup of students (Table ___ $n = 127$) was found in this study. This subgroup included students who have participated in intervention as least two years prior to third grade. Using the database, an investigation of the outcomes for this subgroup would be merited to understand if a higher percentage resulted in an IEP or if two years of intervention supported the students' needs. Based on the feedback from the focus groups, the district should investigate (1) methods by which to encourage increased retention of instructional aides, (2) which interventions appear to be providing the maximum gain, and (3) the longevity of the intervention’s impact. Most important, a summation of this study should be shared with all stakeholders as recommended by the focus group participants.

**Summary**

In general, this study showed the RTI framework demonstrated positive outcomes for student achievement. Students in the intervention group grew substantially more than their counterparts. The gap may not be closed, but it has decreased substantially. There are indications that the identification of specific learning disabilities was positively impacted. It appeared that with quality interventions in place, students were able to make gains in the area of reading.

A strong RTI implementation can support all students. Through high-quality instruction and interventions matched to each student’s needs, the underlying principles of Response to Intervention are achieved. All children can learn. The disproportionalities of the past might be eradicated, leading to appropriate identification for specific learning disabilities.
Together these results suggest that students do benefit from the RTI framework implemented in this district. Given the results, my perspective on RTI during the primary years of school in the area of reading is one of cautious optimism that the assess-react-assess approach of the RTI framework helps to meet the needs of individual students.
References


Bean, R. M. (2009). *The reading specialist: Leadership for the classroom, school, and*


Abstract retrieved from ProQuest Dissertations and Thesis database. (UMI No: 3351235)


Brewlow, J.M. (2012). Dropout nation: By the numbers: Dropping out of high school. PBS


Coyne, M. D., Kame'enui, E. J., Simmons, D. C., & Harn, B. A. (2004). Beginning reading intervention as inoculation or insulin: First-grade reading performance of strong


EnGauge 21st century skills: Literacy in the digital age. (2003). Naperville, IL: North Central Regional Educational Laboratory, Metiri Group & NCREL.


Fashola O., & Slavin, R. (1996). Effective and replicable programs for students placed at risk in elementary and middle school. *Show me the evidence!*


first-grade longitudinal reading study of the national research center on learning
disabilities. Reading and Writing: An Inter-disciplinary Journal, 21(4), 413-436.


and evaluation on pedagogy, student achievement, and student awareness of learning.


the identification of learning disabilities. Learning Disabilities Research & Practice, 13,
204-219.

policymakers, and parents. Teaching Exceptional Children, 38(1), 57-61.

Teaching Exceptional Children 39(5), 14-20.


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


21-26.


Indiana Department of Education and CTB/McGraw-Hill (2014). *Indiana Statewide Testing for*


Evaluation-Criteria-for-SLD.pdf.


Lopez, J. T. (2009). *Reading outcomes of pre-literate 3rd-grade students after two years of combined reading classroom instruction and individualized intervention or classroom instruction alone*. University of Nebraska, Omaha, Nebraska. (UMI No: 3330574)

Approaches, 5(2), 269-283.


Intervention.


Tilly, W. D. (2003). How many tiers are needed for successful prevention and early intervention?
Heartland Area Education Agency’s evolution from four to three tiers. Paper presented at
the National Research Center of Learning Disabilities Responsiveness-to-Intervention
Symposium, Kansas City, MO.

Thurlow, M. L., Moen, R. E., Liu, K. K., Scullin, S., Hausmann, K. E., & Shyyan, V. (March
2009). Disabilities and reading: Understanding the effects of disabilities and their
relationship to reading instruction and assessment. Minneapolis, MN: University of
Minnesota, Partnership for Accessible Reading Assessment.

Intensive remedial instruction for children with severe reading disabilities: Immediate
and long-term outcomes from two instructional approaches. Journal of Learning
Disabilities, 34, 33-58.

Torgensen, J. K. (2004). Lessons learned from research on interventions for students who have
difficulty learning to read. In Learning and Teaching Reading, 89-103. Retrieved from


cross-disciplinary examination of the prevalence of mixed methods in educational

Urban Schools Services: School Dependent Status 16.3 | Association of Christian Schools
International. (n.d.) Retrieved from https://www.acsi.org/resources/cse-magazine/urban-
school-services-school-dependent-students-163.

http://dibels.uoregon.edu/digelinfo.php


Vaughn (Eds.). *Evidence-based reading practices for response to intervention*, (pp. 11-29). Baltimore: Brookes.


Vaughn, S., & Fuchs,

Appendix A

School Demographics: Student Population

<table>
<thead>
<tr>
<th></th>
<th>School A</th>
<th>School B</th>
<th>School C(^\text{13})</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>School grade range</strong></td>
<td>K-6</td>
<td>K-6</td>
<td>K-6</td>
</tr>
<tr>
<td><strong>School population (Average class size)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008-2009 +</td>
<td>623</td>
<td>886</td>
<td>872</td>
</tr>
<tr>
<td>2009-2010</td>
<td>626</td>
<td>647</td>
<td>1118</td>
</tr>
<tr>
<td>2010-2011</td>
<td>685</td>
<td>647</td>
<td>1137</td>
</tr>
<tr>
<td>2011-2012 ++</td>
<td>749</td>
<td>658</td>
<td>1094</td>
</tr>
<tr>
<td>2012-2013 *</td>
<td></td>
<td></td>
<td>930</td>
</tr>
<tr>
<td>2013-2014</td>
<td>776 (22.8)</td>
<td>702 (23.4)</td>
<td>940 (21.9)</td>
</tr>
</tbody>
</table>

* Data from IDOE with the exception of 2012-2013 which is from district record due to IDOE data error
+ Redistrict students, and staff, due to overcrowding between 2008-2009 and 2009-2010
++ Redistrict students, and staff, due to overcrowding between 2011-2012 and 2012-2013

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Caucasian</td>
<td>95.8%</td>
<td>95.4%</td>
<td>90.5%</td>
<td>89.2%</td>
<td>96.3%</td>
<td>94.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.0%</td>
<td>0.6%</td>
<td>2.3%</td>
<td>3.8%</td>
<td>1.5%</td>
<td>1.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.5%</td>
<td>1.3%</td>
<td>2.3%</td>
<td>1.7%</td>
<td>0.3%</td>
<td>1.2%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>1.1%</td>
<td>0.6%</td>
<td>0.5%</td>
<td>1.1%</td>
<td>0.0%</td>
<td>1.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiracial</td>
<td>2.4%</td>
<td>2.1%</td>
<td>4.1%</td>
<td>4.0%</td>
<td>1.5%</td>
<td>1.4%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.2%</td>
<td>0.5%</td>
<td>1.6%</td>
<td>1.1%</td>
<td>0.2%</td>
<td>0.9%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Free/Reduced Lunch</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2008-2009</td>
<td>23.8%</td>
<td>32.6%</td>
<td>23.8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009-2010</td>
<td>29.8%</td>
<td>36.0%</td>
<td>29.4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-2011</td>
<td>30.7%</td>
<td>39.6%</td>
<td>32.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011-2012</td>
<td>29.2%</td>
<td>36.8%</td>
<td>37.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012-2013</td>
<td>28.1%</td>
<td>35.3%</td>
<td>33.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013-2014</td>
<td>31.1%</td>
<td>39.6%</td>
<td>31.4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Special Education</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2008-2009</td>
<td>19.7%</td>
<td>15.9%</td>
<td>26.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009-2010</td>
<td>15.0%</td>
<td>12.2%</td>
<td>22.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-2011</td>
<td>14.0%</td>
<td>14.2%</td>
<td>24.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011-2012</td>
<td>15.2%</td>
<td>18.8%</td>
<td>20.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012-2013</td>
<td>15.0%</td>
<td>17.7%</td>
<td>19.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013-2014</td>
<td>13.0%</td>
<td>16.7%</td>
<td>17.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{13}\) School C is the home for the Functional Academic Classes (FAC), as such students with specialized needs attend.
Appendix B

School Demographics: Assessment Results

<table>
<thead>
<tr>
<th></th>
<th>School A</th>
<th>School B</th>
<th>School C</th>
</tr>
</thead>
<tbody>
<tr>
<td>School grade range</td>
<td>K-6</td>
<td>K-6</td>
<td>K-6</td>
</tr>
<tr>
<td>ISTEP+ ELA, Grade 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008-2009</td>
<td>81.1%</td>
<td>86.3%</td>
<td>80.0%</td>
</tr>
<tr>
<td>2009-2010</td>
<td>85.1%</td>
<td>80.0%</td>
<td>8.9%</td>
</tr>
<tr>
<td>2010-2011</td>
<td>94.9%</td>
<td>80.8%</td>
<td>86.7%</td>
</tr>
<tr>
<td>2011-2012</td>
<td>90.0%</td>
<td>85.4%</td>
<td>91.4%</td>
</tr>
<tr>
<td>2012-2013</td>
<td>92.4%</td>
<td>84.1%</td>
<td>86.5%</td>
</tr>
<tr>
<td>2013-2014</td>
<td>93.9%</td>
<td>89.9%</td>
<td>87.1%</td>
</tr>
<tr>
<td>IREAD-3, Grade 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011-2012</td>
<td>90.9%</td>
<td>93.3%</td>
<td>93.5%</td>
</tr>
<tr>
<td>2012-2013</td>
<td>98.3%</td>
<td>95.2%</td>
<td>91.5%</td>
</tr>
<tr>
<td>2013-2014</td>
<td>99.1%</td>
<td>98.0%</td>
<td>92.9%</td>
</tr>
<tr>
<td>TERA-3 Reading Quotient Averages, Kindergarten (BOY)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008-2009</td>
<td>91.3</td>
<td>86.8</td>
<td>90.0</td>
</tr>
<tr>
<td>2009-2010</td>
<td>89.2</td>
<td>90.1</td>
<td>87.3</td>
</tr>
<tr>
<td>2010-2011</td>
<td>96.2</td>
<td>95.1</td>
<td>90.5</td>
</tr>
</tbody>
</table>

Note: Indiana ISTEP+ ELA, Grade 3 averages were 76.8%, 81.7%, 84.2%, 86.2%, 85.2%, and 83.6% respectively.

The Indiana IREAD-3 state averages were 85.7%, 91.4%, 90.8% respectively.
Appendix C

School Demographics: Faculty

<table>
<thead>
<tr>
<th>School grade range</th>
<th>School A</th>
<th>School B</th>
<th>School C&lt;sup&gt;14&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K-6</td>
<td>K-6</td>
<td>K-6</td>
</tr>
</tbody>
</table>

Staff demographics:

(2013-14 as representative year)

<table>
<thead>
<tr>
<th></th>
<th>School A</th>
<th>School B</th>
<th>School C</th>
</tr>
</thead>
<tbody>
<tr>
<td># Administrators</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td># Teachers (Certified and highly qualified)</td>
<td>34</td>
<td>30</td>
<td>43</td>
</tr>
<tr>
<td># Instructional assistance (Certified)</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td># Instructional assistance (Classified)</td>
<td>10</td>
<td>14</td>
<td>14</td>
</tr>
</tbody>
</table>

Average teaching experience (years): District average in 2013 was 13 years.

% Teachers Highly Qualified (Title II Regulations): 100% 100% 100%

% Teachers with advanced degrees<sup>15</sup>: District had 22.3% of elementary teachers with a master degree or above.

Note: The number of teachers listed above includes classroom (i.e. homeroom) teachers only. In addition to the classroom teachers, each school had special area teacher, a guidance counselor, and a reading coach. The instructional aides listed include those aides that support the classroom. Special education aides are not included.

<sup>14</sup> School C expanded to two buildings prior to the 2009-2010 school year. At this time, the campus administration increased to include two principals and one assistant principal. The second assistant principal was in charge of the grade 5-6 building.

<sup>15</sup> The number of advanced degrees is low as compared nationally due to the changes in license requirements in Indiana. A requirement for continued license renewal is ongoing professional development. Advanced degrees are no longer an Indiana requirement for a teaching license.
Appendix D

Focus Group Consent Form

This form will be used by the faculty for their consent in participating in the focus group.

**Study Title:** Response to intervention: A mixed method longitudinal study of the impact on reading achievement within a school district

**Study Purpose and Rationale:** The purpose of this study was to examine the impact of the RTI process on student achievement with the intent of adding to the research base. Likewise, the data provided evidence to the district to guide program decisions. The study investigated whether students identified as lower achieving readers in kindergarten or first grade, who received the deliberate intervention in reading, had higher levels of reading achievement by the end of third grade than students who did not receive the intervention. The independent variable in this study was the use of the intervention program in an attempt to increase reading achievement. Academic achievement in reading was the dependent variable and was measured by reading assessment results and placement in special education.

**Inclusion/Exclusion Criteria:** To be eligible to participate in this study, you must be an employee of the district from 2008 through 2014, have knowledge of the district's RTI protocols, involved with the primary grade classrooms, and willing to participate in the study.

**Participation Procedures and Duration:** Participation in the focus group will be voluntary. If more than one staff member meets the above criteria to fill a focus group seat, seniority will be used for selection purposes. Interviews will last approximately 1 to 1 1/2 hours and will take place at the central office in a private meeting room.

**Data Confidentiality:** All interview responses will be confidential.

**Audio Recording:** To ensure accuracy, with your permission, the interviews will be audio recorded. Once the interview is concluded, a transcription of the interview will be created. The transcription will be created by Transcription Plus (2850 Canterbury Lane, Indianapolis, IN 46220). Transcription Plus has been creating transcriptions for the medical, insurance, and business industries since 1991. Customers include individuals to Fortune 500 companies, such as Roche. All files are transmitted through a secure FTP server to ensure privacy and documents are secured with 256-bit encryption, the same level of encryption used by NASA, the FBI, and the Department of Defense. Any names used on the audiotape will be changed to pseudonyms when the recordings are transcribed. Following the transcription, only the primary investigator will have access to the audio file. Following the transcription, the primary investigator will code and analyze the transcript in terms of the themes that emerge.

**Data Confidentiality:** As explained above, the interview data will be assigned a pseudonym so the data remains confidential. The researcher therefore protects participants’ identifies from being associated with the data. The dissertation data will focus on findings from the theme analysis. If quotes are used from the interviews, the pseudonym will be used and all precautions will be taken to avoid confidentiality issues. No individual teacher or administrator will be identified.
Storage of Data: The narrative data will be stored on a password protected laptop that is locked in the primary investigator's office at night. Paper data will be stored in a locked file cabinet in the researchers' office. Digital and paper data will be stored for three years and then deleted. Only members of the research team will have access to the data.

Risks or Discomfort: There are no anticipated risks or discomforts with this study.

Benefits: There are no anticipated benefits to the focus group participant with this study.

Voluntary Participation: Agreeing to be interviewed in this study is completely voluntary and you are free to withdraw your permission at any time for any reason without penalty or prejudice from the investigator. Please feel free to ask questions of the investigator before beginning the interview.

IRB Contact Information: For your rights as a research subject, you may contact the following: Director, Office of Research Integrity, Ball State University, Muncie, IN 47306, (765) 285-7070 or at irb@bsu.edu.

Study Title: Response to intervention: A mixed method longitudinal study of the impact on reading achievement within a school district

Consent: I, ______________________________, agree to be interviewed for this research project entitled, "Response to intervention: A mixed method longitudinal study of the impact on reading achievement within a school district." I have had the study explained to me and my questions have been answered to my satisfaction. I have read the description of this project and give my consent to participate in the interview. I understand that I will receive a copy of this informed consent form to keep for future reference.

I understand that the interview will be recorded and give my consent. My initials in the following box indicate my permission to audiotape the interview: __________________

To the best of my knowledge, I meet the inclusion/exclusion criteria for participation (described on the previous page) in this study.

________________________________________________________________________  ________________________
Participant's Signature                                    Date

Researcher Contact Information

Principal Investigator:  Faculty Supervisor:
Laura M. Miller, Graduate Student  Dr. Marilynn Quick
Department of Educational Leadership  Department of Educational Leadership
Ball State University  Ball State University
Muncie, IN 47306  Muncie, IN 47306
Telephone: (765) 285-3287  Telephone: (765) 285-3287
Email: lmmiller2@bsu.edu  Email: mquick@bsu.edu
Appendix E

Interview Protocol for Focus Group Discussions

**Background:** At the onset of the focus group discussion, the data were shared with the participants along with explanations of the various statistical tests. Care was taken to avoid influencing the participants by providing an explanation of the statistics without the researcher's interpretation of the data. Participants were asked if any further explanation of the study or the statistics was necessary and provided a hard copy for the discussion of the data.

The following open-ended questions, along with questions aiming to probe deeper into participants' comments, were asked during the focus group discussions interviews (i.e. 1 teacher group per school and 1 administrator group from the district) with school faculty.

**Introduction:** Thank you for agreeing to assist in the analysis of the RTI data and the district's implementation over the past several years. The discussion will be recorded and coded into emerging themes. All copies of the discussion will remain in a secure file and your identification will remain confidential.

First, the goal of this research is to gain insight into student achievement outcomes of the district's RTI implementation. I will begin with an explanation of data and results. During the discussion portion, I will ask questions regarding the data and your experience with the process. This process is intended to deepen the understanding of the complex process of student achievement in regards to the RTI process by encapsulating statistical data with a more detailed understanding of the phenomena. Now, I would like to turn your attention to the presentation on the data and statistics. At any time during our discussion, please feel free to ask for clarification.
Focus Group Discussion Questions: Reflecting on the data, what differences do you notice between and among the nested groups (e.g. by school, by year of implementation, by gender, between treatment and control groups)? Please be as specific as possible and provide evidence within the results where possible.

Pending the direction of the group discussion, this question was repeated to seek additional information on each combination within the nested groups or probing questions were included to enhance the reflection by the group. One such question was, throughout the implementation, have you seen student outcomes (e.g. reading ability, assessment outcomes, and special education identification) change over the time? If so, what do you see contributing to this change (e.g. staff knowledge of the processes, teacher involvement in the progress monitoring, improved leadership in the implementation)?

1. Reflecting on the statistical results, what inferences can you make regarding the implementation of RTI?

2. From your professional knowledge of the district's implementation of the RTI process as well as the statistical results we have just discussed, how can the RTI process be improved in the future?

3. To conclude this discussion, do you have any additional comments regarding your overall understanding of the long-range impact of participating in RTI or do you have any final comments on the implementation of RTI that you would like to express?

Thank you for assisting me with my dissertation research. As part of the educational community at the district, the findings will be accessible to you and my hope is the findings prove beneficial to you and your students.
Appendix F

**Focus Group Initial Coding**

<table>
<thead>
<tr>
<th>THEMES</th>
<th>Research Question(s)</th>
<th>Administrator (School A)</th>
<th>Teacher (School B)</th>
<th>Teacher (School B)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Distribution by gender</strong> was closer than anticipated</td>
<td>RQ3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SES</strong> is a big part of academic struggles</td>
<td>RQ3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Program refinements</strong> impact the program</td>
<td>RQ4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Media attention</strong> (print, television, and social media) impacts parent interaction</td>
<td>RQ4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Programming appears to be working</strong></td>
<td>RQ1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A range of logistics</strong> impact the program</td>
<td>RQ4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Special Education</strong> identification seems to be on target, but the numbers do seem high.</td>
<td>RQ2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Between buildings</strong> - each building has its own personality which leads to slight differences</td>
<td>RQ3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Role of IA</strong> - efficacy, training, buy-in is critical</td>
<td>RQ4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RTI: DIFFERENCES IN READING ACHIEVEMENT 239