Closing the Reading Achievement Gap: The Impact of a Multitiered Response to Intervention Framework in Kindergarten

Victoria Borsella
St. John Fisher College, victoriaborsella@gmail.com

Follow this and additional works at: https://fisherpub.sjfc.edu/education_etd

Part of the Education Commons

How has open access to Fisher Digital Publications benefited you?

Recommended Citation

Please note that the Recommended Citation provides general citation information and may not be appropriate for your discipline. To receive help in creating a citation based on your discipline, please visit http://libguides.sjfc.edu/citations.

This document is posted at https://fisherpub.sjfc.edu/education_etd/366 and is brought to you for free and open access by Fisher Digital Publications at St. John Fisher College. For more information, please contact fisherpub@sjfc.edu.
Closing the Reading Achievement Gap: The Impact of a Multitiered Response to Intervention Framework in Kindergarten

Abstract
The purpose of this quantitative, quasi-experimental, ex post facto study was to examine the impact the Response to Intervention (RTI) framework had on kindergarten reading scores as measured by the standardized test for the achievement of reading early literacy assessment (STAR-ELA). This study employed an ex post facto design utilizing a retrospective cohort. Two 2 × 2 mixed-subjects factorial analysis of variances (ANOVAs) were used in addressing the study's three research questions. First, the study evaluated the combined effectiveness of Tier 1 and Tier 2 reading interventions on nonclassified kindergarten students’ reading achievement to determine if the implementation of RTI impacted the reading scores of kindergarten students. Findings reveal that Tier 1 and Tier 2 RTIs were effective in enhancing students’ reading scores. Second, the study compared the efficacy of Tier 1 and Tier 2 interventions separately to determine if there was any relationship between the type of RTI employed and the resulting improvements in reading achievement. Findings show that Tier 2 students exhibited significantly greater improvement than Tier 1 students. Finally, this study investigated the potential role of gender in moderating the effect of Tier 2 interventions on reading achievement. The data examined displayed that both males and females in Tier 2 RTI showed approximately equal and statistically significant gains in STAR-ELA scores. Recommendations are that further studies should include students who are classified with an individualized education plan, in addition to English language learners, to reveal how interventions can be effective with different populations. In addition, further research should be conducted on Tier 3 interventions.

Document Type
Dissertation

Degree Name
Doctor of Education (EdD)

Department
Executive Leadership

First Supervisor
Jeff Wallis

Second Supervisor
Winsome Gregory

Subject Categories
Education

This dissertation is available at Fisher Digital Publications: https://fisherpub.sjfc.edu/education_etd/366
Closing the Reading Achievement Gap: The Impact of a Multitiered Response to Intervention Framework in Kindergarten

By

Victoria Borsella

Submitted in partial fulfillment of the requirements for the degree Ed.D. in Executive Leadership

Supervised by

Dr. W. Jeff Wallis

Committee Member

Dr. Winsome Gregory

Ralph C. Wilson, Jr. School of Education
St. John Fisher College

August 2018
Dedication

There have been many people who have walked beside me throughout this journey and who have guided, inspired, and motivated me. First, I would like to thank my dissertation chair, Dr. Wallis, and committee member, Dr. Gregory, for your support and leadership throughout this process; you have motivated and encouraged me more than you can imagine. I would like to thank my team, Dr. Ed.D., who has become my lifelong new family members. Each one of you has a special place in my heart. Thank you for helping me and pushing me to complete this journey. I would like to thank my daughter, Alexandra. I love you for the little girl that you once were and for the amazing woman you have become; I am so proud of you.

I would like to thank my best friend, my rock, Toniann. You have shown me how to be strong, stronger than I ever thought I could be. You encouraged and supported me from the beginning to keep moving forward and never give up. You believed in me, even when I did not believe in myself. You have taught me how to be a better woman, mother, and friend. Thank you for giving me the courage to follow my dreams.

I would like to dedicate this dissertation to the memory of my father, Victor Borsella, and brother, Richie Borsella. I have become a more resilient person because of your influence and love. I know you watch me from Heaven and have guided my every step. “My precious child, I love you, and I would never leave you. During your times of trial and suffering, when you see only one set of footprints, it was then that I carried you.”
Biographical Sketch

Victoria Borsella is currently the District Associate Administrator for Pre-K-2nd Grade Early Literacy at the Port Washington School District in Long Island, New York. Ms. Borsella attended Iona College from 1994 to 1998 and graduated with a Bachelor of Science degree in Elementary Education in 1998. She attended the College of New Rochelle from 2003 to 2016 and graduated with three Master of Science degrees in Reading and Literacy, Special Education, and School Building and School District Leadership. She came to St. John Fisher College in the summer of 2016 and began doctoral studies in the Ed.D. Program in Executive Leadership. Ms. Borsella pursued her research in Response to Intervention under the direction of Dr. Jeff Wallis and Dr. Winsome Gregory and received the Ed.D. degree in 2018.
Abstract

The purpose of this quantitative, quasi-experimental, ex post facto study was to examine the impact the Response to Intervention (RTI) framework had on kindergarten reading scores as measured by the standardized test for the achievement of reading early literacy assessment (STAR-ELA). This study employed an ex post facto design utilizing a retrospective cohort. Two $2 \times 2$ mixed-subjects factorial analysis of variances (ANOVAs) were used in addressing the study’s three research questions. First, the study evaluated the combined effectiveness of Tier 1 and Tier 2 reading interventions on nonclassified kindergarten students’ reading achievement to determine if the implementation of RTI impacted the reading scores of kindergarten students. Findings reveal that Tier 1 and Tier 2 RTIs were effective in enhancing students’ reading scores.

Second, the study compared the efficacy of Tier 1 and Tier 2 interventions separately to determine if there was any relationship between the type of RTI employed and the resulting improvements in reading achievement. Findings show that Tier 2 students exhibited significantly greater improvement than Tier 1 students. Finally, this study investigated the potential role of gender in moderating the effect of Tier 2 interventions on reading achievement. The data examined displayed that both males and females in Tier 2 RTI showed approximately equal and statistically significant gains in STAR-ELA scores. Recommendations are that further studies should include students who are classified with an individualized education plan, in addition to English language
learners, to reveal how interventions can be effective with different populations. In addition, further research should be conducted on Tier 3 interventions.
# Table of Contents

Dedication .......................................................................................................................... iii

Biographical Sketch ........................................................................................................... iv

Abstract ............................................................................................................................... v

Table of Contents .............................................................................................................. vii

List of Tables ...................................................................................................................... x

List of Figures .................................................................................................................... xi

Chapter 1: Introduction ....................................................................................................... 1

   Problem Statement .......................................................................................................... 3

   Theoretical Rationale ...................................................................................................... 3

   Statement of Purpose ...................................................................................................... 6

   Research Questions ......................................................................................................... 6

   Null Hypothesis .............................................................................................................. 7

   Alternative Hypothesis ................................................................................................... 7

   Potential Significance of the Study ................................................................................ 8

   Definitions of Terms ....................................................................................................... 8

   Chapter Summary ......................................................................................................... 11

Chapter 2: Review of the Literature .................................................................................. 12

   Introduction and Purpose .............................................................................................. 12

   Public Law 94-142 ........................................................................................................ 13

   IDEA ............................................................................................................................. 16
## List of Tables

<table>
<thead>
<tr>
<th>Item</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 3.1</td>
<td>District Student Enrollment Breakdown 2016-2017</td>
<td>57</td>
</tr>
<tr>
<td>Table 3.2</td>
<td>District Demographics for the Student Population 2016-2017</td>
<td>58</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>Descriptive Statistics on the STAR-ELA Reading Achievement Test at Pretest and Posttest for Tier 1 and Tier 2 Students</td>
<td>72</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>Results of Shapiro-Wilk Tests of Normality of Distributions for Tier 1 and Tier 2 Students at Pretest and Posttest</td>
<td>72</td>
</tr>
<tr>
<td>Table 4.3</td>
<td>Summary Table for 2 (RTI Tiers) × 2 (Time) Mixed-Subjects Factorial ANOVA</td>
<td>75</td>
</tr>
<tr>
<td>Table 4.4</td>
<td>Descriptive Statistics on the STAR-ELA Reading Achievement Test at Pretest and Posttest for Male and Female Tier 2 Students</td>
<td>80</td>
</tr>
<tr>
<td>Table 4.5</td>
<td>Results of Shapiro-Wilk Tests of Normality of Distributions for Male and Female Tier 2 Students at Pretest and Posttest</td>
<td>81</td>
</tr>
<tr>
<td>Table 4.6</td>
<td>Summary Table for 2 (Gender) x 2 (Time) Mixed-Subjects Factorial ANOVA</td>
<td>82</td>
</tr>
</tbody>
</table>
## List of Figures

<table>
<thead>
<tr>
<th>Item</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 3.1</td>
<td>Response to Intervention Model of Tiered Instruction</td>
<td>54</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Frequency Histograms of Pretest and Posttest STAR-ELA Reading Achievement Test Scores for Tier 1 Students and Tier 2 Students</td>
<td>71</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>Mean Performance on the STAR-ELA Reading Assessment Among Tier 1 and Tier 2 Students at Fall Pretest and Spring Posttest.</td>
<td>75</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>Frequency Histograms of Pretest and Posttest STAR-ELA Reading Achievement Test Scores for Male and Female Tier 2 Students</td>
<td>80</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>Mean Performance on the STAR-ELA Reading Assessment Among Male and Female Tier 2 Students at Fall Pretest and Spring Posttest</td>
<td>83</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

Many primary-aged children struggle to develop reading skills in the first few years of school (Reynolds, Wheldall, & Madelaine, 2011). Before third grade, 20% of all children have difficulty with reading. This equates to 10 million children in the United States (National Center for Education Statistics [NCES], 2015). Researchers continually indicate that students who do not learn to read in the early grades will experience negative effects to their later reading success (Fuchs & Fuchs, 2009; Vellutino, Scanlon, Zhang, & Schatschneider, 2008; Wanzek, Roberts, Otaiba, & Kent, 2014). Studies have shown the positive effects of early intervention with students who have difficulty reading (Otaiba et al., 2011; Vellutino et al., 2008).

Many changes have occurred in reading instruction within the last 20 years (Clarke et al., 2011). Researchers state that children who do not acquire basic reading skills in the primary grades remain poor readers throughout their academic career (Clarke et al., 2011; Clay, 1985; Fuchs & Fuchs, 2009; Fuchs & Vaughn, 2012; Juel, 1988; Otaiba et al., 2011; Vaughn & Fuchs, 2006; Vellutino et al., 2008; Wanzek et al., 2014). Studies indicate that young children who struggle with reading in the early grades and do not receive intervention will perform below their peers (Fuchs, Fuchs, & Compton, 2012; Otaiba & Torgesen, 2007; Vellutino et al., 2008; Wanzek et al., 2014). Researchers have focused on the achievement gaps that exist between groups of young students (Mokhtari, Neel, Kaiser, & Le, 2015). In addition, researchers, educators, and practitioners have begun emphasizing the importance of prevention, rather than correction, of reading.
difficulties (Fuchs & Vaughn, 2012). Furthermore, researchers have stated that kindergarten is the critical window in which educators can prevent reading failure (O’Connor, Fulmer, Harty, & Bell, 2005).

The Individuals with Disabilities Education Improvement Act (IDEA, 2004) mandates that schools implement a tiered structure to prevent reading failure. Response to Intervention (RTI) is a tiered instructional model that is used for the prevention of reading failure in all 50 of the United States (Otaiba, Wagner, & Miller, 2014). RTI emphasizes research-based instruction with increasing levels of intensity to support students who are struggling with reading (Weiss & Friesen, 2014). Variations within RTI models exist, and schools have the flexibility to use different research-based methods to identify struggling readers (Gersten et al., 2009). Within an RTI tiered approach, some schools may incorporate two tiers of increasing instruction, while other schools incorporate seven tiers (Berkeley, Bender, Peaster, & Saunders, 2009). Further still, one school’s Tier 2 approach might be similar to another school’s Tier 5 (Fuchs & Fuchs, 2009). According to Hoover and Love (2011), a three-tiered RTI model that provides early support and progress monitoring to young students is most effective in closing the achievement gap in reading.

For RTI to effectively prevent reading failure, interventions must be executed properly and monitored often (Gilbertson, Witt, Singletary, & VanDerHeyden, 2007; Hill, King, Lemons, & Partanen, 2012). In addition, the accurate identification of students requiring intensive intervention is a crucial factor in the prevention of reading difficulties (Otaiba et al., 2011). However, within school districts, there are still many
unanswered questions about the fidelity of the school-wide implementation of RTI (Bradshaw, Reinke, Brown, Bevans, & Leaf, 2008).

**Problem Statement**

Of the students who enter kindergarten, 60% are not ready for school (U.S. Department of Education [USDOE], 2015). Kindergarten readiness is the ability to demonstrate skills that enable a child to participate and succeed in school (Cooke, Kretlow, & Helf, 2010). Reading serves as a basis for learning, and a deficit in reading causes lifelong challenges (Brynner, 2008). Students who begin school with a weakness in reading skills have a hard time catching up to their peers. In addition, Brynner (2008) stated that poor reading skills in adolescent and adulthood have been found to have negative effects on employment, and poor reading skills contribute to social segregation. Early identification of poor readers is fundamental in solving this problem (Wanzek et al., 2014). Research suggests that reading failure can be reduced with a high-quality, core reading-intervention program (Case et al., 2014). Systematic implementation of RTI can possibly be an effective tool to preventing reading failure (Stahl, 2016). According to Dallas (2017), RTI is used to “provide interventions early and proactively for students who show risk of developing academic deficiencies” (p. 1).

**Theoretical Rationale**

Developmental lag theory (DL) is characterized by a slower progression in cognitive tasks (Stanovich, Nathan, & Vala-Rossi, 1986). The DL theory further implies that as the brain develops and matures, children will catch up to their peers. The DL model has been used in the area of reading difficulties for many years (Stanovich et al., 1986). Many practitioners wait for children who are *late bloomers* to catch up to their
peers (Lyon et al., 2001). Torgesen et al. (1999) stated that a reading disability is most commonly caused by a lack of ability. Two longitudinal studies by Juel (1988) and Shaywitz et al. (1999) show that it is not a developmental lag but a skill deficit that prevents children from reading proficiently (Torgesen, 2000). According to Torgesen (1998), students will not close learning gaps and will not catch up to their peers without intensive intervention. Therefore, intervention is necessary to prevent reading failure.

Fletcher and Vaughn (2009) defined RTI as a framework for increasing student outcomes that involves the use of a universal screener, progress monitoring, increasing tiers of intervention, and data-based decision making.

**Problem Solving Model.** According to Fuchs and Fuchs (2016), most practitioners implementing RTI use a problem-solving model (PSM). This model emphasizes individual interventions that focus on the student, environment, and curriculum (Tilly, 2006). Within the PSM, interventionists determine the issue or gap in learning, the team members meet to analyze and implement instruction, they refine the intervention, and they monitor and assess the progress (Fuchs & Fuchs, 2016). Throughout the PSM and within all tiers, “data about a student’s responsiveness to intervention becomes the driving force” (Grimes, 2002, p. 4). Within this model, teams of practitioners collaborate to review student data (Little et al., 2012).

**Zone of proximal development.** Cognitive psychologist, Lev Vygotsky (1978), discussed the concept of how growth occurs in children. Vygotsky stated that to comprehend the relationship between development and learning, one must differentiate between two developmental levels: the actual developmental level (ADL) and the potential level of development (PLD) (Wang, Bruce, & Hughes, 2011). According to
Vygotsky (1978), the ADL refers to what a child can demonstrate independently, and it involves different mental functions that are completed through various developmental stages. The PLD involves what a child can only achieve with the assistance of others (Vygotsky, 1978). The difference between both levels is the zone of proximal development (ZPD) (Levykh, 2008).

Vygotsky (1978) defined the ZPD as “the distance between the actual developmental level as determined by independent problem-solving and the level of potential development as determined through problem-solving under adult guidance or in collaboration with a more capable peer” (p. 86). In relationship to education, Vygotsky (1978) summarized a child’s ZPD as:

The distance between the level of his actual development, determined with the help of a learning task performed independently, and the level of a child’s potential development, determined with the help of learning tasks performed by the child under the guidance of adults and in collaboration with his smarter classmates. (p. 42)

Vygotsky (1978) stated that the interaction between the teacher and student is a crucial aspect of stimulating a child’s cognitive development (Schcolnik, Kol, & Abarbanel, 2006). The ZPD is the target zone where a child can learn a new skill with the assistance of an adult (Schrader, 2015). Instruction that is too difficult or easy may bring about insufficient progress (Vygotsky, 1978). Vygotsky’s (1978) work is important in the implementation of RTI, because Vygotsky believed that instruction should occur within the ZPD (Ellis, Larkin, & Worthington, n.d.). This belief supports the multitiered instructional levels of RTI. Within the RTI framework, the universal screener can be
used to determine the ZPD for each student. Students are instructed within their ZPD and frequent progress monitoring can provide continual instructional information (Grigorenko, 2009).

**Statement of Purpose**

The purpose of this quantitative study was to examine the impact a three-tiered RTI model had on nonclassified kindergarten students’ reading achievement. Nonclassified kindergarten students received Tier 1 intervention (the core curriculum) from their classroom teacher. Support staff implemented intensive small-group instruction to students in Tier 2 and Tier 3. This study explored the impact RTI had on student reading scores before and after the intervention, implementing a baseline assessment, using ex post facto data and a retrospective cohort. In addition, this study investigated if gender had a significant impact on reading scores.

**Research Questions**

According to Creswell (2014), quantitative research questions investigate the relationship between independent and dependent variables. In addition, “research questions or hypotheses specifically focus the purpose of the study” (Creswell, 2014, p. 143). RTI provides early intervention to struggling readers as a remedy to prevent reading failure. Early intervention is critical in preventing reading failure so that the reading gap does not widen between peers (Noltemeyer, Boone, & Sansosti, 2014). Most researchers promote the implementation of a three-tiered framework, which is utilized in RTI, and it can be used to prevent reading failure for at-risk students (Noltemeyer et al., 2014). This study investigated with the following questions:
1. To what extent does the implementation of Response to Intervention (RTI) impact nonclassified kindergarten students’ reading ability as measured by scores on the STAR Early Literacy Assessment (STAR-ELA)?

2. Is there a relationship between improvement in reading scores on the STAR-ELA and the types of RTI provided to nonclassified struggling kindergarten readers?

3. To what extent, if any, does gender moderate the impact of Tier 2 RTI on reading scores of nonclassified struggling kindergarten readers as measured by scores on the STAR-ELA?

**Null Hypothesis**

H1o: There is no difference between the scores of nonclassified kindergarten students who received Tier 2 RTI and those who did not.

H2o: There is no relationship between improvement in reading scores and the type of RTI provided to at-risk kindergarten students reading at a below-grade level.

H3o: Gender does not moderate the effectiveness of a Tier 2 RTI on the reading scores of nonclassified struggling kindergarten readers as measured by scores on the STAR-ELA.

**Alternative Hypothesis**

H1o: There is a difference between the scores of nonclassified kindergarten students who received Tier 2 RTI and those who did not.

H2o: There is a significant relationship between improvements in reading scores and the type of RTI provided to at-risk kindergarten students reading below-grade level.
H3o: Gender significantly moderates the effectiveness of Tier 2 RTI on the reading scores of nonclassified struggling kindergarten readers as measured by scores on the STAR-ELA.

**Potential Significance of the Study**

The number of children who are at risk of reading failure continues to increase in the United States (NCES, 2015). Current research supports that early interventions in the primary grades prevents reading failure (Stahl, 2016). Without intervention in the early grades, students who are struggling with reading will continue to struggle in the future. Researchers have stated that kindergarten is a critical time for intervention (O’Connor et al., 2005; Stahl, 2016; Wanzek & Vaughn, 2007). Data from this study may indicate the effectiveness of the RTI process and the impact on kindergarten reading scores. In addition, this study may be helpful to teachers and administrators who seek to increase student reading scores by implementing RTI in kindergarten.

**Definitions of Terms**

*Assessment* – a tool that is used to monitor student progress.

*At-Risk Students* – low-achieving learners who have reading difficulties, reading one or more grade levels below standard (Wanzek, Wexler, Vaughn, & Cuillo, 2010). For this study, at-risk students are considered those who score below the cutoff points of developing reading failure.

*Core Curriculum* – high-quality classroom reading instruction that is based on methods that are scientifically validated.

*Criterion-Referenced Assessment* – a measure of student performance relative to a standard.
Differentiated Instruction – a teaching method that provides modifications and alterations to the curriculum to account for different learning needs.

Direct Instruction – a teaching model that explicitly teaches carefully prescribed and planned lessons (Robinson, Lambert, Towner, & Caros, 2016).

IQ Discrepancy Model – an approach to identifying children with a learning disability who might need to access special education services. To access services, students’ testing must show a severe discrepancy between their IQ and their academic ability (Restori, Katz, & Lee, 2009).

Individualized Education Plan (IEP) – an instrument that identifies, describes, and prescribes an instructional protocol for children who are identified as in need of special educational services.

Nonclassified Students – learners who do not have an IEP.

Nonresponder Students – learners who display little improvement from effective literacy instruction (Torgesen, 2000).

Practitioners – instructional staff who provide support services to struggling students. Staff includes kindergarten classroom teachers, reading specialists, speech and language pathologists, special education teachers, teaching assistants, academic intervention support teachers, and school principals.

Progress Monitoring – a process in which a practitioner collects reading data on a student’s progress and performance (Otaiba et al., 2014).

Reading Failure – a student who performs significantly below his or her grade level in reading.
Response to Intervention (RTI) – a research-based, multitiered scientific instructional framework for the early identification and support of students with learning needs (Mellard, McKnight, & Woods, 2009). The RTI framework referenced in this study is a three-tiered model, with increasing levels of intensity based on a student’s instructional need (Clarke et al., 2011).

Standardized Test for the Assessment of Reading (STAR) – a means to measure early literacy and numeracy skills in eight domains: print concepts, phonemic awareness, phonics, fluency, vocabulary, counting, operations, and measurement and data.

Tier 1 Intervention – a universal core instruction implemented by a classroom teacher that assures all students receive research-based quality instruction (Bornstein, 2015).

Tier 2 Intervention – instruction that is implemented by a teacher to students who need more intensive additional support and who do not respond to the core instruction (Fuchs & Fuchs, 2007).

Tier 3 Intervention – instruction that is implemented by a teacher to students who are nonresponders to Tier 1 and Tier 2 interventions. Instruction by the teacher is more frequent and has a greater emphasis and intensity that is used to target specific reading strategies (Wanzek & Vaughn, 2010).

Universal Assessment – a curriculum-based tool that is used to monitor and measure student progress in the major components of literacy instruction, that is, STAR-ELA (Otaiba, 2005).
Chapter Summary

Research suggests that possessing strong beginning reading skills are critical for students’ future school success. When students are not identified early as struggling readers, they may never catch up to their peers. RTI is a research-based, multitiered framework that has the potential to support 95% of students who are struggling with reading (Burns & Gibbons, 2012). To effectively help struggling readers, Fuchs and Vaughn (2012) opined that schools need to identify which model of RTI is most successful in preventing reading failure.

This research paper has five chapters. The first chapter reviewed the research problem, the purpose of the study, the research questions, and the potential significance of this study. A review of the literature is presented in Chapter 2. The research design, methodology, and analysis is discussed in Chapter 3. Chapter 4 presents a detailed analysis of the results and findings, and Chapter 5 discusses the findings, implications, and recommendations for future research and practice.
Chapter 2: Review of the Literature

This chapter provides an analysis and synthesis of RTI literature. Of the children who are not reading proficiently by the end of third grade, 83% will be at risk of dropping out of high school (Fiester, 2013). RTI is a framework that is implemented in all 50 United States to prevent reading failure (Stahl, 2016).

Introduction and Purpose

The purpose of this study was to examine the impact RTI had on nonclassified kindergarten students’ reading scores as measured by the STAR-ELA. In addition, this study explored if a relationship exists between the type of RTI provided to students and their reading scores. Further, this study investigated if gender had an impact on reading scores.

A review of the literature revealed that early intervention is imperative to prevent reading failure (Cunningham & Stanovich, 1997; Fuchs & Vaughn, 2012; Juel, 1988). Moreover, Denton (2012) believed that struggling readers should be identified in kindergarten and provided with interventions. Wanzek et al. (2014) opined that if struggling readers do not receive intervention before third grade, it is almost impossible, without intensive intervention in place, for struggling students to catch up to their peers.

Analysis of the literature shows that RTI is a multitiered framework that can prevent reading failure (Stahl, 2016). In addition, there are two approaches to RTI: the standard treatment approach, and the problem-solving approach. Both approaches can
prevent reading failure; however, schools must determine which model is most appropriate for their population (King & Coughlin, 2016).

**Public Law 94-142**

Public Law 94-142, The Education for All Handicapped Children Act, was approved in 1975, and it mandated that states and local education agencies provide for the education of all children with disabilities (Smith, 2005). According to Smith (2005), there were difficulties in the way disabilities were being defined. In 1997, the law was reauthorized and called the Individuals with Disabilities Education Act (IDEA). Within IDEA, the IQ discrepancy model was implemented to identify specific learning disabilities (McGill, Styck, Palomares, & Hass, 2016.) The discrepancy model established a difference between cognitive and intellectual ability and a student’s academic achievement (Restori et al., 2009).

**IQ discrepancy model.** Four criteria need to be met in order to identify a child as having a learning disability using the IQ discrepancy model (IQDM). This is done by establishing a discrepancy between cognitive ability and academic performance, identifying a cognitive processing deficit, determining educational needs with or without special education, and making exclusionary considerations (Restori et al., 2009). According to Gresham, Restori, and Cooke (2008), there are issues associated with the IQSM. First, it is difficult to make an early identification because it employs a wait-to-fail approach, where students are not provided with services in a timely manner. Second, there is a lack of scientific-based evidence because there is no evidence of this model being valid or reliable for identifying students with learning disabilities. Next, there are issues involving the manner in which practitioners apply this approach, and many
children are being over identified, which results in an increase in the special education population. Finally, students who have academic achievement problems may not be identified because of their average IQ score.

**IQ discrepancy model studies.** There is a substantial body of research that states that ability achievement models do not accurately identify students with a learning disability (Fletcher et al., 2002; Francis et al., 2005; Vellutino, Scanlon, & Lyon, 2000). One longitudinal study, conducted by Shaywitz, Fletcher, Holahan, and Shaywitz (1992), followed 415 children from Connecticut. The sample consisted of 85% Caucasian, 11% African American, 2% Hispanic, and 2% unknown ethnicity. The children were followed longitudinally from kindergarten through ninth grade, but the data reported are through fifth grade. The purpose of this study was to compare two definitions of reading disability: a discrepancy-based model and a low-reading-achievement model. Children were identified and divided into two groups in the second grade based on the discrepancy and low-achieving formula. Data were collected using parent surveys, teacher surveys, and student-based measures.

The findings of the Shaywitz et al. (1992) study show that both groups of children had similar performance on assessments in terms of language, gross motor activities, visual perception, and behavior in kindergarten. However, there was one major difference between the groups. The discrepancy group’s performance was superior to the low-achievement group in both the second and fifth grades. Analysis of the data indicated that IQ accounted for almost all the variance between the groups. These findings support the notion that a reading disability should be defined based on reading alone (Siegel, 1998).
Similarly, Fletcher, Francis, Rourke, Shaywitz, and Shaywitz (1992) conducted a study to address the validity of discriminating between children with reading disabilities, according to the absence or presence of a discrepancy between intelligence test scores and academic performance. Included in the sample were 1,069 children from Windsor, Ontario who were referred for evaluation of a learning disability. The children ranged from 9-14-years old, and they achieved a full-scale IQ of 79 or above. The sample was predominately White, and no significant differences between age, gender, or race were noted when grouping. The children were sorted into one of five categories: (a) met a discrepancy-based definition uncorrected for the correlation of IQ and achievement, (b) met a discrepancy-based definition correcting for the correlation of IQ and achievement, (c) met a low-achievement definition with no discrepancy, (d) met the criteria for both a and b, and (e) met none of the criteria, and the child did not have a reading disability. The groups were compared using 10 neuropsychological tests. Results indicate small group differences that resulted in little differences among the groups. The researchers particularly noted that there was “the failure to find large differences between children who are low-achieving relative to children who meet a variety of discrepancy-based definitions; these comparisons call into question the value of the concept of discrepancy” (Fletcher et al., 1992, p. 560). Limitations of this study included one population, and it could not be generalized to different populations. In addition, comparisons were limited to a set of criteria that was based on one assessment.

In synthesizing both studies, it is apparent that the use of the discrepancy formula lacked validity in the classification of a learning disability and reading instruction. According to Siegel (1989), low IQ scores do not necessarily indicate a reading
deficiency, “therefore children with low IQ scores who are reading disabled can
legitimately be considered reading disabled and not slow learners who are poor readers
due to lack of ability” (Siegel, 1989, p. 212). This supports the notion that IQ-
achievement discrepancy does not reliably distinguish between readers who are disabled
and nondisabled readers (Vellutino et al., 2000).

**IDEA**

In contrast to Public-Law 94-142, IDEA (2004) allows states to select the
discrepancy model and other alternatives to identify students with learning disabilities;
however, the states “must not require the use of severe discrepancy between intellectual
ability and achievement” (IDEA, 34 C.F.R. 300.307(a)(1)). Further, IDEA states, that the
states “must permit the use of a process based on the child’s response to scientific,
research-based intervention” (IDEA, 34 C.F.R. 300.307(a)(2)). According to McGill et
al. (2016), the paradigm shift to RTI was the result of research suggesting that students
were able to benefit from academic interventions in order to prevent learning disabilities.
In light of all the problems with the IQDM, RTI was implemented (Restori et al., 2009).
RTI models aid in determining if a child may have a learning disability or needs
academic intervention grounded in scientific-based evidence (Tran, Sanchez, Arellano, &
Swanson, 2011).

**Early Intervention**

There is an abundance of literature that identifies early intervention as an
important factor in preventing reading failure (Catts, Neilsen, Bridges, & Lu, 2016; Fuchs
& Fuchs, 2006; Fuchs, Fuchs, & Vaughn, 2014; Lam & McMaster, 2014; Otaiba et al.,
2014; Sparks, Patton, & Murdoch, 2014). Similarly, there are many researchers who
have stressed the importance of preventing reading failure rather than correcting it (Denton, Ciancio, & Fletcher, 2006a; Hunter et al., 2015; Otaiba et al., 2014). There is an ample amount of research that shows reading failure is preventable for all but a very small percentage of children (Hunter et al., 2015). The early exposure to print can lay a more successful foundation for reading and a greater desire to read (Sparks et al., 2014).

A longitudinal study conducted by Cunningham and Stanovich (1997) examined the influence of print exposure to explain various measures of reading achievement. Although this study is dated, it is significant because the students were followed over 10 years and were assessed at various periods throughout their school journey. In addition, this study extended previous work from which their colleagues had identified print exposure to be a powerful variable even after controlling differences in cognitive ability (Stanovich & Cunningham, 1992, 1993; Stanovich, West & Harrison, 1995; West & Stanovich, 1991). The researchers attempted to close a gap in the literature by examining early reading acquisition and its relationship to reading ability in adolescence.

Cunningham and Stanovich (1997) followed 56 first-grade students from two classrooms in a middle-class elementary school (32 boys, 24 girls). Ten years later, 27 students remained in the school district for follow-up testing (15 boys, 12 girls). The students were administered standardized reading tasks each year. A correlational study was conducted, and the findings indicate that early reading acquisition in first grade could predict 11th-grade reading comprehension—even after accounting for variance in comprehension ability. Therefore, the Cunningham and Stanovich study shows that first-grade reading comprehension is a significant predictor of 11th-grade reading achievement and the importance of early reading success for later reading ability. In addition, a
pattern is established as early as first grade in which students with strong reading skills engage in reading more often than their nonskilled peers (Cunningham & Stanovich, 1997). However, there were limitations to this study, one being the small sample size.

Sparks et al. (2014) duplicated Cunningham and Stanovich’s (1997) study with a larger number of students. Over 10 years, 54 students were tracked and given assessments each year throughout the study. The findings of the study are similar to Cunningham and Stanovich’s (1997) study, where exposure to print in early years was a significant predictor of 10th-grade reading comprehension, even after the IQ variable had been controlled. There were some noteworthy differences between the two studies. Cunningham and Stanovich (1997) found that first-grade reading ability predicted a variance in print exposure in 11th grade. In addition, third- and fifth-grade reading abilities were stronger predictors of print exposure than first grade. In contrast, the Sparks et al. (2014) study found that only second- and third-grade reading ability predicted variance in print exposure in 10th grade. Both studies stressed the need for an early reading start, which is critical for establishing a successful school journey (Cunningham & Stanovich, 1997; Sparks et al., 2014).

Using an historical control group in two different schools, O’Connor et al. (2005) conducted a longitudinal study of 400 students from kindergarten to third grade. The study took place in an urban setting within the Northeast United States. Ethnically, students were 83% European American, 12% African American, 2% Hispanic, and 3% Native American. The purpose of the O’Connor et al. study was to identify the effects of RTI on student reading outcomes. Students were divided into groups that were constructed from a baseline assessment. Students who were nonresponsive to Tier 1
instructions received small-group intervention in the classroom three times a week for 25 minutes. Students nonresponsive to Tier 2 instruction received one-on-one instruction 5 days a week for 30 minutes.

The findings of the O’Connor et al. (2005) study show that 29% of the students who received intervention did not need subsequent intervention for the next 3 years. Only seven students who were identified as struggling readers in kindergarten received additional intervention for all 4 years. The study reveals that students within intervention groups score significantly higher on various reading tasks when compared to an historical control group from the same school. One limitation of the O’Connor et al. study was that university researchers implemented the Tier 1 and Tier 2 interventions. This might account for some bias when administering or grading the assessment.

In a ground-breaking longitudinal study, Juel (1988) showed the importance of not using a wait-to-fail model on struggling readers. The study was ground breaking in that Juel (1988) revealed that it is almost always a skill deficit that prevents a child from reading, and it is not a developmental lag because as a child’s brain matures, difficulties with learning to read would diminish. The Juel study was one of the first to show that a wait-to-fail model is not effective with struggling readers.

Juel (1988) tracked 54 students from a school in Texas, from first to fourth grade, examining whether poor readers caught up to their average-reading peers. Juel used standardized tests of reading and listening comprehension, phonemic awareness, word recognition, and decoding. At the end of first grade, Juel (1988) divided the students into two groups based on their standardized test scores. Throughout this 3-year study, the poor readers did not catch up to their average counterparts. The findings indicate that out
of 24 students who were poor readers in first grade, 21 of them were still at least 6 months behind their average counterparts in fourth grade. These findings showed that students who are reading below grade level at the end of first grade continue to be poor readers at the end of fourth grade.

In addition, Juel (1988) found two factors that poor readers lacked: phonemic awareness and the ability to decode words. Juel also found evidence that early writing skills were not a predictor of later writing abilities, and there were no correlations between the two. Findings from the Juel (1988) study stressed the important need for the RTI process.

As shown in the studies of Cunningham and Stanovich (1997), Juel (1988), O’Connor et al. (2005), and Sparks et al. (2014), the need for early reading success to prevent ready failure is evident. In addition, as seen in these studies, the research is clear that targeted assessments in kindergarten and first grade can often foretell reading difficulties in later grades. Furthermore, the gap between poor and strong readers broadens over the elementary school years, and after third grade, the gap becomes almost impossible to close without intensive intervention (Otaiba, 2005). Likewise, Reynolds et al. (2011) suggested that practitioners need to intervene as early as kindergarten to possibly prevent detrimental effects on future school performance.

**Importance of Preschool**

Developmental studies indicate that children begin to learn earlier than formally thought (Kuhl, 2012). Brain studies show that children’s abilities measured as early as infancy can predict later learning and functioning (Kuhl, 2012). In addition, Kuhl (2012) stated that children’s language abilities from 1-2 years of age are foretelling of their
preliteracy skills at age 5 years. A meta-analysis by Camilli, Vargas, Ryan, and Barnett (2010) proposed that children who spend a year in developmentally appropriate preschool programs would improve in early language and literacy skills. Further still, six longitudinal studies associated entry-level reading skills in kindergarten as the strongest predictor of future academic success (Duncan et al., 2007). Young et al. (2008) suggested, “There is evidence that key early literacy skills are predictive of subsequent literacy achievement in kindergarten and first grade that can be taught to preschool age children” (p. 14).

According to Lonigan and Phillips (2016), there is a growing acknowledgement of the importance of preschool years for the development of later reading skills. In addition, research suggests that intervention as early as preschool can help influence reading abilities in later years (Lonigan & Cunningham, 2013). According to Ferguson (2009), closing early childhood achievement gaps is imperative because “lost developmental time in a child’s life cannot be reclaimed” (p. 2). Missall et al. (2007) expressed that exposure to literacy in preschool can function as an early intervention to help students build crucial foundational skills to prevent later reading failure.

**Emergence of Kindergarten**

Friedrich Froebel, a German educationalist, established the first kindergarten in 1837 in Germany (Lee, Burkam, Ready, Honigman, & Meisels, 2006). Froebel was guided by the philosophies of Rousseau and Pestalozzi who believed children should “develop their mental, social, and emotional faculties through play, music, movement, interaction with the outdoors, and opportunities to engage in independent creative pursuits” (Lee et al., 2006, p. 166). Froebel believed that kindergarten students should
experience differentiation based on their developmental characteristics (Chung & Walsh, 2000). Margarethe Schurz, a student of Froebel, opened the first kindergarten in the United States in 1857. Self-directed play, in addition to the development of the whole child, remained the focus of kindergarten in the United States (Fromberg, 2006). In 1970, research on cognitive growth emerged, and kindergarten changed from a play-based curriculum to an academic curriculum (Lee et al., 2006). However, educators continue to debate the goals and purposes of kindergarten (Votruba-Drzal, Li-Grining, & Maldonado-Carreno, 2008). Some educators believe that a play-based, or traditional model, of kindergarten is appropriate, while others believe that kindergarten should incorporate academics and interventions (Lee et al., 2006). Both models are contradictory of each other. According to Lee et al. (2006):

> Those who advocate for the “developmentally appropriate” kindergarten are criticized for underestimating children’s capacity to acquire a wide variety of skills and concepts. Those who advocate for a “formal kindergarten” are criticized for narrowing the curriculum, ignoring children’s social and emotional needs, and dampening young children’s natural curiosity and enthusiasm to learn. (p. 167)

States require kindergarten teachers to incorporate reading skills into the curriculum (Cooke et al., 2010). The Common Core State Standards (CCSS) shifted teaching in the early grades to include complex thinking, collaboration, curriculum, assessment, and accountability (Nicholson, Bauer, & Woolley, 2016).
Kindergarten Reading

According to Rasinski (2017), many children still struggle with reading in the United States. The Eunice Kennedy Shriver National Institute of Child Health and Human Development, (2000) identified precise competencies that affect beginning reading including: (a) phonics, (b) phonemic awareness, (c) word decoding, (d) fluency, (e) comprehension, (f) alphabetic principle, and (g) vocabulary. In addition, studies indicate that there are benefits to using a scripted supplemental intervention (Helf, Cooke, & Konrad, 2014; Simmons et al., 2011). Most kindergarten students who struggle with reading are deficit in phonological skills (O’Connor, 2000). Simmons et al. (2011) suggested that schools provide systematic additional intervention in kindergarten, because interventions that start early are more effective than those started in later years.

Students who experience reading problems need explicit and systematic instruction in kindergarten (Wanzek, Roberts, & Otaiba, 2013). A study conducted by Scanlon, Vellutino, Small, Fanuele, and Sweeney (2005) found that small-group instruction, 3 days a week, focusing on phonemic awareness resulted in significant reductions in reading difficulties at the end of first grade.

According to Kent, Wanzek, and Otaiba (2012), intensifying the effectiveness of teaching strategies by the classroom teacher can improve reading achievement. Some strategies include: (a) explicit instruction, (b) expanding time spent on reading instruction, (c) differentiation, (d) whole- and small-group instruction, and (e) meaningful text experiences (Juel & Minden-Cupp, 2000; Otaiba et al., 2008; Pressley et al., 2001). Furthermore, it has been found that student engagement and occasions to directly apply skills have increased student reading outcomes (Kent et al., 2012).
Response to Intervention

In 2004, President George W. Bush reauthorized into law the IDEA. This reauthorization allowed schools to use an RTI model instead of an IQDM to prevent reading failure and identify children with learning disabilities (Stahl, 2016). In addition, the New York State Education Department implemented learning standards which stated that each student must be instructed by a core curriculum. A core curriculum provides research-based differentiated instruction, which is aligned to the learning standards.

An RTI model is a tiered instructional framework that emphasizes research-based instruction with increasing tiers or levels of intensity based on a student’s instructional need (Clarke et al., 2011). Tier 1 intervention is designed to deliver high-quality, research-based instruction to all students (Burns & Gibbons, 2012). Tier 2 interventions are designed to provide additional intensive supports for students who do not make progress within Tier 1 (Fuchs et al., 2014). Tier 3 is the most intensive tier of support, and it is reserved for students who fall significantly below their peers. Typically, a special education evaluation may be needed in Tier 3 (Denton, Fletcher, Anthony, & Francis, 2006b).

The purpose of RTI is to provide additional supplementary support to struggling readers (Stahl, 2016). However, there are various definitions and descriptions of RTI (Fuchs et al., 2012). According to Burns and Gibbons (2012), RTI is a multiliter prevention framework. To most efficiently allocate resources to teach all students, RTI requires the systematic use of data to make instructional decisions. In a multiliter framework, the instruction is intensified for struggling readers who are not showing growth with less-intensive instruction (Turse & Albrecht, 2015). As a student moves
through the tiers, differentiated instruction is provided. Each tier of academic intervention becomes more intensive as students move across the tiers (Fuchs & Fuchs, 2006).

There are still some arguments by researchers as to how many tiers are appropriate to deliver instruction to struggling readers; however, most models employ a three-tiered approach (Fuchs & Fuchs, 2006) (Appendix A). Each state can develop its own interpretation of the framework, leading to various models (Fuchs, Fuchs, & Stecker, 2010). The Institute of Education Sciences published a practical guide for educators to effectively implement RTI (Otaiba et al., 2014). Included in the guide are five essential components for effective implementation. These essential components include a high-quality core curriculum, universal screening, increasingly intensive tiers, progress monitoring, and consistency of implementation (Otaiba et al., 2014).

One challenge in the RTI process is to identify which components are most effective for increasing student achievement (Little et al., 2012). Wanzek and Vaughn, (2007) synthesized research identifying reading interventions that were most effective with struggling readers. Wanzek and Vaughn (2007) examined 18 studies in their synthesis. They used specific criteria to determine which studies were important. These criteria included interventions that were provided for a significant amount of time (20 weeks or more), kindergarten through third-grade students, and children with reading difficulties. The studies were run from 1995 to 2005. The data results suggest three important factors that were effective with struggling readers: the duration of the intervention, the instructional group size, and the grade level of intervention. The findings reveal positive outcomes for students with reading difficulties who participated
in extensive interventions. The limitations to the synthesis included students in the
sample who had not yet received effective reading instruction within Tier 1. These
students may not be representative of the students who had the most significant reading
difficulties (Wanzek & Vaughn, 2007).

Recently, researchers evaluated the effectiveness of RTI, core curriculums, and
classroom interventions that are most appropriate to use with struggling readers (Stahl,
2016). Balu et al. (2015) released a large-scale study of the RTI process involving over
1,200 elementary schools, across 13 states, using students in Grades 1-3. Balu et al.
(2015) gathered data from the 2011-2012 school year, which consisted of surveys
completed by staff and student records on reading assessments. This study compared
RTI intervention practices between schools. Balu et al. (2015) found that 86% of the
impacted schools reported that RTI was fully implemented within their school, compared
to 56% in the sample school. In addition, 67% of the schools replaced core instruction
for students needing supplemental support, instead of adding additional instruction. To
prevent reading failure, Balu et al. (2015) found that teachers must provide 90 minutes of
core instruction in addition to targeted intensive instruction for struggling readers. As
well, students in all schools who were struggling with reading did not receive an
intervention that was significantly different from the students who were above grade
level. The findings of the Balu et al. study indicate that there were differences in how
schools were implementing RTI and the recommended effective RTI framework.
Likewise, these findings suggest a thorough evaluation of the effectiveness of RTI within
schools. One limitation of the Balu et al. study was that it used a regression discontinuity
(RD) design that demonstrated a causal relationship. In both the Wanzek and Vaughn
(2007) and Balu et al. (2015) studies, it is evident that many factors can influence the effectiveness of RTI. However, when RTI is implemented correctly, positive reading outcomes occur. In addition, school districts need to examine their RTI process to ensure all components are being implemented with fidelity to prevent reading failure (Stahl, 2016).

An experimental study conducted by Case et al. (2010) examined the impact of a short-term (11 weeks) reading intervention with 30 struggling first-grade students. Students were randomly assigned to two groups. The intervention group received 30 minutes of small-group supplemental tutoring 3 days a week, and the control group remained in their classroom with the usual instruction. The intervention group received instruction in vocabulary, fluency, sight-word recognition, comprehension, letter-sound correspondence, and spelling. The findings of this study reveal that students who receive an intervention have significantly better growth on decoding, spelling, and fluency. These results suggest that students who receive Tier 2 interventions were quickly remediated in decoding, spelling, and fluency. Limitations to this study included a small population of 30 participants. However, due to the experimental design of this study, the researchers were able to point to the medium-to-large effect-size estimates that the control group children were unable to match the intervention group’s responsiveness (Case et al., 2010).

Likewise, Kerins, Trotter, and Schoenbrodt (2010) found comparable results to Case et al. (2010) in a Tier 2 instruction of 23 first graders. The study included students who were identified as having below-average reading abilities and attended the same suburban public school in southern Maryland. The purpose of the Kerins et al. (2010)
study was to determine if children would benefit from 17 weeks of additional, supplemental scripted instruction in segmenting, blending, and phonics, using multisensory techniques, compared to children receiving classroom instruction. Students were randomly assigned to two groups. The two groups consisted of an intervention group, which received additional instruction that was provided by a speech pathologist and a special education teacher, and the control group children remained in their classroom and did not receive additional support. Results from the study reveal that the students within the intervention group exhibited growth on phoneme blending, segmentation, and running records when compared to the control group. However, an analysis of the results revealed that there were no group differences on any norm-referenced tests of word reading. Limitations to this study included a small population size, which cannot be generalized to all first-grade students. In addition, the Tier 1 instruction was not monitored. Although the school required teachers to allocate for literacy instruction, the Kerins et al. (2010) study could not ensure the instruction took place with fidelity.

In analyzing the Case et al. (2010) and Kerins et al. (2010) studies, it is noteworthy to mention that although there were no differences in standardized test scores, the intervention groups significantly improved in reading, segmenting, sight-word recognition, and blending during the interventions. Wanzek and Vaughn, (2007) hypothesized that if students are given a longer duration of intervention, perhaps the results would be different because the effectiveness of long-term reading interventions indicates a significant increase in reading scores. In addition, both the Case et al. (2010) and Kerins et al. (2010) studies incorporated a research-based reading program within
their Tier 1 instruction; however, monitoring to ensure the fidelity of the intervention did not occur.

**Tier 1.** In RTI, Tier 1 is a universal core instruction implemented by classroom teachers that assures all students receive research-based quality instruction (Bornstein, 2015). This core instruction is one element in preventing reading failure (Wanzek et al., 2014). Within Tier 1, there is also a universal assessment that provides baseline data for each student (Burke et al., 2012). During Tier 1 instruction, the classroom teacher delivers a high-quality, research-based core curriculum that meets the needs of 80-90% of the students in the classroom (Stahl, 2016). In addition, general education teachers must make informed decisions based on the data to provide necessary interventions to struggling readers (Jones, Yssel, & Grant, 2012). Schools continue to struggle with selecting an appropriate intervention, which is imperative to effectively implementing RTI (Fuchs & Vaughn, 2012).

Otaiba et al. (2014) stated that researchers and practitioners have concerns about the lack of research when implementing a multitiered model. In a study conducted by Otaiba et al. (2014), the researchers compared the efficacy and impact of two RTI models on reading outcomes. They used dynamic RTI and typical RTI. This study took place in the Southeast United States, within seven school districts in their first year of RTI implementation. There were 522 first-grade students who participated in the study; less than 3% of the children were limited in English proficiency. The researchers used a randomized, controlled experiment to assign students into the two different RTI groups. The first group was the dynamic RTI, which quickly refers Grade 1 students who are lacking the most skills to the most intensive tier for support. The second group, the
typical RTI group, started all students in Tier 1, and over time, the students were moved throughout the tiers.

The findings of the Otaiba et al. (2014) study reveal that students in the dynamic RTI group had significantly higher reading scores compared to students in the typical RTI group. In addition, the students who were immediately eligible for Tier 3 intervention in the dynamic RTI group achieved higher reading scores compared to the Tier 3 students in the typical RTI—students who had to wait longer to be eligible for that tier. Within this experimental design, findings show that the students with the least amount of skills, who received the most amount of intervention, performed significantly better in reading compared to the typical group. One limitation of this study was that the findings may not be generalized to older students because only first-grade students participated. Another limitation was that the assessors were not blind to the students’ conditions.

Gilbert et al. (2013) conducted a randomized control study involving 649 struggling first-grade readers to improve their reading scores using a multi-tiered model. Out of 11 schools, students were recruited for two consecutive years. The district where the study took place had a 95% disadvantaged student population, and the district mandated that the study be conducted with kindergarten students. Gilbert et al. (2013) examined what proportion of the at-risk students, who received additional interventions, had achieved reading performance in the normal range in Grades 1-3. The problem the researchers were trying to solve was whether the RTI models achieved the preventive intent of reading failure. The findings of Gilbert et al. (2013) study were similar to Otaiba et al. (2014) in that the students who received immediate intervention, because of
the unresponsiveness to Tier 1 instruction, benefited from the intervention compared to students who remained within Tier 1. However, the findings of the Gilbert et al. (2013) study also show students who participated in Tier 2 interventions were not prevented from future reading failure. Of the students tested at the end of first grade, 41% did not score in the normal range on reading. One limitation of the study was the lack of progress monitoring within Tier 1 intervention. According to Otaiba et al. (2014), students should be frequently monitored for progress to prevent reading failure. Another limitation of the Gilbert et al. (2013) study was that graduate students (noncertified) delivered the tiered interventions rather than certified staff.

In synthesizing both the Gilbert et al. (2013) and the Otaiba et al. (2014) studies, it is important to note that an experimental control group was used in both cases. Otaiba et al. (2014) had certified trained teachers delivering the instruction to students who struggled with reading. In contrast, Gilbert et al. (2013) did not use certified staff. This might account for the discrepancy in the results of student achievement. However, it is clear in both studies that the students who received the intervention immediately had an increase in reading performance.

Mathes et al. (2005) evaluated two approaches to small-group instruction and investigated the effectiveness of combining Tier 1 and Tier 2 interventions for struggling readers in first grade. The study was conducted in six high-performing urban school districts in Texas. The first approach, called Proactive Reading (PR) emphasized positive feedback, fluency, and direct instruction. The second approach, Responsive Reading (RR) had less structured lessons. Both interventions provided phonemic awareness, phonics, comprehension, and fluency, but they differed in that PR was implemented
using scripted, systematic, direct instruction. Practitioners in the PR group selected lessons based on the student needs that were determined by diagnostic screening. The students in the RR group spent time reading and responding to text. The findings indicate that students in both intervention groups scored significantly higher on reading indicators than the students who did not receive any intervention; however, students within the PR group achieved higher scores on reading outcomes.

The findings from all these studies emphasize the importance of an RTI model. Fuchs and Fuchs (2006) expressed that the best approach for struggling readers is to provide scientific research-based instruction with a high-quality core curriculum. As shown in these studies, Tier 1 instruction can enhance the learning of all students, especially students who are at risk for reading failure. In addition, these studies show that direct instruction and explicit RTI intervention can have a significant impact on the improvement of reading performance.

**Tier 2.** Tier 2 interventions are designed for students who do not respond to the core curriculum and need a more intensive level of intervention (Fuchs & Fuchs, 2007). Progress monitoring, small group size, as well as length and frequency of the intervention, must be incorporated into Tier 2 for it to be effective (Fuchs, Mock, Morgan, & Young, 2003; Otaiba, 2005). Tier 2 interventions should occur three to five times a week, and progress should be frequently monitored (Gersten et al., 2009; Otaiba et al., 2014). The literature suggests that 15% of students will benefit from Tier 2 interventions (Abbott & Wills, 2012).

Wanzek et al. (2016) conducted a meta-analysis that extended the previous work of Wanzek and Vaughn (2007), which studied reading interventions in Tier 3. However,
Wanzek et al. (2016) examined less-intensive Tier 2 interventions for students with reading difficulties in kindergarten through third grade. The Wanzek et al. (2016) study was the first meta-analysis that explored the overall effects of Tier 2 interventions on students’ reading ability as well as intervention features that could be associated with improved achievement in reading. The researchers analysed 72 studies using specific criteria. The studies that met the criteria were peer reviewed and addressed research questions about Tier 2 reading interventions in the early elementary grades, and they included reading outcome data. In addition, the studies included a treatment comparison and experimental or quasi-experimental design.

The findings of the Wanzek et al. (2016) analysis demonstrate a positive effect in Tier 2 interventions on foundational reading skills such as decoding, phonemic awareness, fluency, and word identification. Similar to the findings of Wanzek and Vaughn (2007), this meta-analysis found that a variety of reading interventions had improved reading outcomes in the kindergarten through third grade. In addition, the findings indicate that there were no differences in immediate effects related to the type of intervention a child received because the outcomes were not statistically different between intervention types. Similarly, group size did not affect the reading outcomes. Students who were provided one-on-one interventions, and students in groups of two to five achieved similar reading outcomes. The Wanzek et al. (2016) meta-analysis was unable to determine whether there were differences in student outcomes based on the qualifications of the practitioner implementing the intervention.

Case et al. (2014) investigated the immediate and long-term effects of brief Tier 2 interventions for first grade struggling readers in a parochial school near Atlantic City,
New Jersey. This study extended their previous study and replicated the Tier 2 interventions with a larger group of students (462) and (123). The students were selected using a screening paradigm that included teacher ratings and two direct measures of reading. There were 61 students assigned to the intervention group and 62 students assigned to the control group. The interventions were implemented three times a week for 40 minutes over a 3-month period. Group size varied from two to four students. Interventions focused on phonemic awareness skills, spelling, vocabulary, oral reading fluency, and comprehension. The control group received interventions from their classroom teacher while the intervention group received instruction from tutors using scripted lessons that were developed by evidence-based instructional methods. Analyses were conducted using pretest and posttest measures.

The findings of the Case et al. (2014) study showed that the intervention group significantly improved on decoding, fluency, and word spelling. In addition, the researchers investigated the long-term effects of Tier 2 interventions and followed students until the end of second grade. The findings reveal no significant differences between the control and intervention groups. However, data show that students who were identified as responders at the end of first grade outperformed students who identified as nonresponders. There were no other significant findings of growth measures. Case et al. (2014) surmised one possibility that might explain these results is when the intervention group of children was pulled out of their classroom, this reduced the number of children for the teacher to instruct, therefore the teacher may have been able to give more attention to individual students’ needs. In addition, the Case et al. (2014) study examined the brief instruction of Tier 2 interventions.
One limitation of the Case et al. (2014) study was that the control group did not have data on what type of reading intervention occurred by the classroom teacher (Tier 1); this might have skewed the results. In addition, the intervention used may have limited the students’ abilities to maximize their potential, if the intervention used was too easy. Lastly, the pretest was given in September but the intervention did not take place until January. Developmentally, students may have changed during that period of time (Case et al., 2014).

Similarly, Denton et al. (2011) compared the effects of Tier 2 interventions provided to first-grade students who were at risk for reading failure. This study was conducted in nine schools located in the Southwestern United States. The population within the schools was 51% African American, 35% Hispanic, 9% White, and 5% other ethnicities. Using the Texas Primary Reading Inventory, 680 first-grade students were screened, and 273 students met the at-risk criteria, with a final number of 192 students participating in the study after all variables were accounted for. The students were randomly assigned to three treatment groups: extended, concentrated, and distributed interventions. The groups differed based on the length and duration of the intervention.

All students received Tier 1 research-based instruction from their classroom teacher (Denton et al., 2011). Within Tier 2 interventions, all students received the same intervention that provided systematic, explicit, and direct instruction. Students were assessed using a pretest and posttest. Findings from the Denton et al. (2011) study state that there were no significant differences in any reading outcomes between the groups of students who received interventions on three different schedules. The first-grade students who received 16 hours of small-group intervention scored equally as well as those who
received 32 hours. Limitations of this study included the lack of a no-treatment comparison group. In addition, the results of the Denton et al. (2011) study may not be generalizable to other grade levels. Also, this study did not indicate the level of differentiation used within Tier 1 instruction, which could have skewed the results of the Tier 2 instruction (Denton et al., 2011).

In synthesizing these Tier 2 studies, it is apparent that the type of Tier 1 instruction given to students was imperative to their progress within Tier 2. In all three studies (Case et al. 2014; Denton et al. 2011; Wanzek et al. 2016), the researchers indicated their lack of knowledge of the intensity or differentiation that occurred within Tier 1; this might account for the findings of no difference between the groups. Data suggest that a systematic differentiated Tier 1 approach could have a significant impact on reading outcomes (Jefferson, Grant, & Sander, 2017).

**Tier 3.** Tier 3 interventions are reserved for students who fail to respond to the first two tiers of intervention, or Tier 3 is for students who make minimal progress, which is about 2-5% of students (Vaughn, Linan-Thompson, & Hickman, 2003). According to Wanzek and Vaughn (2010), Tier 3 students who make minimal gains after being taught through a core curriculum and receive supplementary interventions through Tier 2, are considered to be *nonresponders*. Nonresponders are students who unsatisfactorily respond to an intervention. Tier 3 interventions incorporate greater reading intensity for students with severe reading difficulties (Wanzek & Vaughn, 2010). Some models of intervention incorporate special education services in Tier 3; however, individual school districts can determine how to incorporate special education into their RTI model (Denton et al., 2010). Tier 3 interventions require a more individualized approach and
There have been very few studies that have examined the response to multitier models involving Tier 3 interventions (Greulich et al., 2014). Denton, Tolar, Fletcher, Barth, and Francis (2013) conducted a randomized control trial involving second-grade students, who demonstrated an insufficient response to Tier 1 and Tier 2 interventions, that evaluated the effects of a Tier 3 reading intervention. The quantitative study was conducted in the Southwestern United States throughout 10 elementary schools. Four schools were in a small city, and the other six schools were in a large urban district, with 89% of the students qualifying for free lunches. The participants in this study consisted of 72 students who received Tier 1 and Tier 2 interventions during first grade. The research question asked was: Do a larger proportion of students, who receive the intervention, meet benchmarks relative to students who receive typical instruction? Participants who were randomly assigned to Tier 3 were provided with intensive small-group instruction targeting specific reading indicators. An ANOVA was used to evaluate reading gains and factors that influenced the gains. One finding of this study suggested that the children might have improved in word reading and comprehension, but they may have had delays in fluency. The results of this study show that an explicit, intensive reading intervention can be effective for students who need Tier 3 interventions, particularly in remediating word reading and phonemic awareness.

A limitation of the Denton et al. (2013) study was the small sample size. When conducting a study within Tier 3 interventions, ensuring a large enough sample size is imperative. Due to the small sample size, results of the study cannot be generalized to a
larger population. One unique feature of this study was the implication for educational practice, which stated students who did not respond to Tier 1 and 2 interventions required systematic, intensive instruction to make gains within Tier 3.

Similarly, O’Connor et al. (2005) conducted a study that included Tier 3 interventions and the movement of students within tiers based on data. This longitudinal study involved 400 students ranging from kindergarten through Grade 3 in two different schools. Of the 400 students, 22 received Tier 2 and 3 interventions. Students who were nonresponsive to Tier 2 interventions received one-on-one tutoring 5 days a week. The findings of the O’Connor et al. (2005) study indicate that students who received Tier 2 and 3 interventions demonstrated growth on all reading measures when compared to a control group. All students reached grade-level standards with only 10 students who required Tier 3 interventions. However, of these 10 students, only four were able to read at grade level by the end of third grade. The results from this study are consistent with the number of students needing Tier 3 instruction, which was 2-5%.

As shown in all three studies, the Tier 3 intervention was intended for the few students who were nonresponders to the Tier 1 (the core curriculum) and Tier 2 interventions. In addition, within Tier 3, all studies emphasized the importance of direct, explicit instruction. In O’Connor et al. (2005), only six children were not able to read at grade level after the intervention was implemented. Furthermore, as seen in Denton et al. (2013), interventions that were shorter but more frequent and conducted with fidelity were most beneficial to the Tier 3 struggling readers. Implementing more of the same intervention may work for some students; however, Tier 3 students may require a different approach to learning (Denton et al., 2013).
**Fidelity of Implementation**

One study by Bailey (2014) investigated how seven rural school districts in Montana implemented four essential components of RTI. This study was significant in that very little empirically based studies have examined RTI in rural settings (Dexter, Hughes, & Farmer, 2008). Although 74% of the schools in this study collected data for progress monitoring, only 53% of the schools used the data to identify students who might have been at risk for reading failure. Most schools in the Bailey (2014) study reported that “data-based decision making as the most difficult component to implement” (p. 50). In addition, only three schools received professional development in RTI. Almost all schools in the study focused their instruction on Tier 2 interventions, and one school provided effective intervention on Tier 3. Findings from the Bailey study specify an absence of direction and insufficient RTI support. One limitation of the study was the lack of information provided to the researchers regarding RTI process within the school district. When data were used to inform decision making as part of the RTI process, gaps between at-risk students and typically developing peers decreased (Alonzo, Tindal, & Robinson, 2008).

Sharp, Sanders, Noltemeyer, Hoffman, and Boone (2016) examined the relationship of the school-wide implementation of RTI and reading scores. They used the RTI Implementation Scale for Reading (RTIS-R), which is designed for professionals in a school setting, to indicate their perception of RTI implementation school wide on a Likert scale. The study was significant in that it was the first quantitative study to examine the relationship between school-wide implementation of RTI and reading score outcomes. The participants in the Sharp et al. (2016) study consisted of 65 principals and
school psychologists from urban, suburban, and rural Pre-K through Grade 6 schools in Ohio. Findings from the study indicate that data-based decision-making was important in all tiers of RTI, and it significantly predicted student reading outcomes. One limitation of this study was the self-reported data, which relied on participants accurately answering questions. Another limitation was the use of convenience sampling, which might not have been representative of other areas in the United States.

Hill et al. (2012) reviewed 22 studies that examined the usefulness of elementary reading interventions within an RTI framework. The studies included in this review were an experimental or a quasi-experimental group design that focused on Tier 2 interventions in a small group or a one-on-one setting. Inter-rater reliability was calculated at 92.3%. The researchers found that most studies did not report information on Tier 1 implementation. In addition, Hill et al. were not clear whether Tier 1 instruction targeted the same skills as those in Tier 2. For RTI to be effective and prevent reading failure, a supplemental, additional instruction is needed in Tier 2 that targets the same skills as Tier 1 (O’Connor, Bocian, Beach, Sanchez, & Flynn, 2013). According to Hill et al. (2012), the effectiveness of Tier 2 interventions depends upon the alignment and quality of the Tier 1 instruction. However, all studies reported the fidelity of Tier 2 instruction. One limitation of this study was that the researchers only reported the fidelity of implementation, and they did not report student outcomes related to fidelity.

As shown in the studies of Bailey (2014), Hill et al. (2012), and Sharp et al. (2016), all indicate the importance of RTI implementation to prevent reading failure. Data used from the assessments can help guide practitioners in developing strategies for at-risk readers (Sharp et al., 2016). Although RTI was associated with improved
performance for struggling readers, it is apparent from the studies shown that there was a gap between what was being researched and what was being practiced (Fletcher & Vaughn, 2009).

Models of Response to Intervention

One aspect of a multitiered RTI approach is the type of intervention offered within Tier 2. Schools offer two basic models of intervention within an RTI framework: the standard treatment protocol (STP) and the problem-solving approach (PSA) (Fuchs & Fuchs, 2006). However, King and Coughlin (2016) recommended that schools need to determine which approach is best to use within their districts (King & Coughlin, 2016). Both models utilize interventions that increase with intensity, and their purpose was to provide effective instruction for struggling readers (Fuchs et al., 2003). However, the models differed in terms of who delivered the intervention and implementation techniques (Fuchs & Fuchs, 2006).

Problem-solving model. The problem-solving model utilized a team approach to implementing strategies based on data regarding student responsiveness to intervention (Fuchs & Fuchs, 2006). Carney and Stiefel (2008) added to the research based on a problem-solving approach that examined the long-term outcomes of success on 43 students in kindergarten through Grade 5 who attended a Midwestern United States elementary school. The ethnicity of this group of students was 88% Caucasian, 7% African American, and 5% Asian. The classroom teacher referred general education students to an instructional support team, due to academic or behavioral concerns. Of the 32 students referred, 28 students were identified as having academic concerns: 10 for behavioral issues and five for unspecified issues. Carney and Stiefel (2008) employed a
problem-solving approach that was individualized for every student. Each student received interventions until the problem was resolved, the year ended, or the student was referred for special education services. Students were followed for 3 years and data were taken in the form of records.

Based on their academic grades and discipline referrals, the results of the study Carney and Stiefel (2008) showed that 19 of the 32 students were successful after 3.5 years and 13 students were identified at moderate risk in year 4. In addition, 33% of the students were re-referred for Tier 2 services throughout the study. Some limitations to the study were the small sample size, no pretest data, and the lack of progress monitoring within Tier 2. Carney and Stiefel (2008) stated that the results of the study show the need for further research into the intervention given within the problem-solving approach.

Similarly, a study conducted by Ross and Begeny (2015), in a Southeast United States school district, examined the effects of evidence-based reading interventions, including the same instructional components but differing in treatment durations and student-teacher ratios. Four second-grade students who were identified as being at risk for reading failure, were randomly assigned to intervention groups. Of the four participants, two were African American, two Caucasian, with two males and two females. The study included a nonintervention control group and four intervention conditions that improved reading fluency. The conditions only differed in duration and instructional context. Findings indicate that all students profited from at least one of the treatment conditions when compared to the control group. The findings suggest, also, that the longer intervention conditions had more of a positive impact on the students’ reading rather than the shorter interventions. One significant finding: there was no
difference between the one-on-one treatment interventions compared to the small-group treatments. Ross and Begeny (2015) posited that, for educators, small-group instruction for reading fluency can be as effective as one-on-one instruction. One limitation of the study was the small sample size of four students. With such a small sample size, the results cannot be generalized to other populations. Ross and Begeny (2015) offered that further research in this area needs to be conducted because many students in schools today are performing below basic levels of reading competency.

Slavin, Lakes, Davis, and Madden (2009) conducted a meta-analysis and reviewed research on the achievement outcomes of students who were struggling readers in kindergarten through Grade 5, using six different reading intervention approaches, including: small-group instruction, computer-assisted instruction, one-on-one tutoring by teachers, one-on-one tutoring by paraprofessionals, classroom instructional process approaches with and without tutoring, and computer-assisted instruction. The analysis included 96 studies that incorporated a duration of 12 or more weeks of intervention; a randomized control group; valid and reliable measures used for assessment, such as standardized reading measures; and there were 15 or more students involved in each study.

Key findings from this meta-analysis revealed that one-on-one tutoring involving phonics and phonemic awareness were extremely successful in improving reading achievement for struggling readers. One-on-one tutoring conducted by teachers was more effective than the paraprofessionals delivering the instruction. Computer-assisted technology had little or no effect on reading achievement. Cooperative learning and structured phonetic models of instruction had strong effects on the struggling readers.
One limitation of this study was that it did not emphasize a correlation that would have added understanding to the effectiveness of the reading programs (Slavin et al., 2009).

In synthesizing these studies, it was apparent that schools needed to examine the interventions given within Tier 2. Tier 2 is the pivotal point in which instructional decisions made by practitioners most influence the general education students (King & Coughlin, 2016). Ross and Begeny (2015) found no significant difference in reading outcomes for students who participated in small group instruction, compared to one-on-one instruction. In contrast, the meta-analysis conducted by Slavin et al. (2009) found one-to-one instruction to be more effective than small group instruction. However, it is important to mention that any additional instruction had a positive effect on reading outcomes. One limitation to the use of a PSA was that it relied on teacher skills, and educators must implement a range of interventions based on student need (King & Coughlin, 2016).

**Standard treatment model.** Most school districts employ a standard treatment protocol (STP) approach to RTI for two reasons: it is more cost effective to train teachers in one strategy than it is to train them in multiple strategies, and more students can be serviced with less staff (Fuchs & Fuchs, 2006). According to Gilbert et al. (2013), there has been a scarcity of research directly comparing the two approaches and, therefore, one approach cannot be deemed better than the other.

Gilbert et al. (2013) conducted a randomized control trial that examined the efficacy of the standard protocol approach within a multitiered RTI model. This study incorporated 232 first-grade students from 11 schools of whom were identified as unresponsive to Tier 1 interventions. All students received kindergarten instruction. The
students were divided into two groups: Tier 1 continued with instruction, and Tier 2 received supplemental instruction. The students received interventions for 7 weeks, and progress monitoring was conducted weekly.

The posttesting conducted in the Gilbert et al. (2014) study revealed that all students made gains on all measures, and the interventions were effective in increasing the development of first-grade readers with an effect size of .19. Results also reveal that the students who participated in Tier 2 interventions were not precluded from future reading difficulties. Of the students who participated in additional interventions, 41% failed to score in the normal range on reading assessments compared to 47% of the Tier 1 students. These findings are consistent with studies conducted by Case et al. (2010) and Mathes et al. (2005) that challenged the preventive intent of the short-term standard protocol RTI approach. One limitation of the Gilbert et al. (2014) study was the monitoring of the Tier 1 instruction. When Tier 1 instruction was not monitored, the researchers could not say with certainty that the students received effective Tier 1 instruction.

Likewise, Wanzek and Vaughn (2008) conducted two studies examining first-grade students’ responses to varying amounts of time in reading intervention. Both studies took place in one Southwestern United States school district using two different elementary schools that had a high percentage of minority students and students living in poverty. The students were assigned to one of three groups that received either: (a) only one dose of intervention, (b) a double dose of intervention, or (c) no intervention (Wanzek & Vaughn, 2008). Both studies were conducted in successive school years with nonoverlapping samples of students. The participants in Study 1 and Study 2 included 50
first graders and 40 first graders, respectively, who were identified as at risk for reading difficulties. Findings of this study indicated that students within the RTI groups (single dose and double dose) revealed few differences. In addition, increasing the intensity of the intervention did not increase the number of students responding to the intervention. Therefore, having more test participants was not necessarily better. However, the results indicate that students within the RTI groups demonstrated gains on the pretest and posttest measures compared to the control group. Limitations of this study included the variability between the two schools with the amount of additional instruction of the comparison group, which was shortened because of budget cuts. In addition, the amount of time the treatment group received intervention was controlled, but the treatment time for the comparison group was not (Wanzek & Vaughn, 2008).

In synthesizing the studies, it was evident that using an STP might be cost effective for some school districts; however, as evident in Gilbert et al. (2013), the short-term effects of the standardized approach might not be the best intervention method for struggling readers. Gilbert et al. (2013) indicated that when students are nonresponsive to RTI, an additional strategy must be implemented.

**Response to Intervention in Kindergarten**

Research has shown that assessments in kindergarten can predict reading achievement in future grades (Catts et al., 2016; Catts, Fey, Zhang, & Tomblin, 2001; Juel, 1988; Schatschneider, Fletcher, Francis, Carlson, & Foorman, 2004). Fuchs and Fuchs (2006) implemented the RTI framework to improve early identification of struggling readers. Catts et al. (2016) found that implementing RTI as early as kindergarten can correctly identify children at risk. In addition, the Catts et al. (2016)
study investigated whether the response to language intervention in kindergarten was a predictor of third-grade reading comprehension. In the study, 366 kindergarten children were followed from the beginning of kindergarten until the end of third grade. Ethnically, the sample was diverse including students who were 63% Caucasian, 11% African American, 6% Hispanic, 7% American Indian, 6% Asian, and 7% multiracial; 35% of the children received free or reduced-fee lunches. All children were screened at the beginning of kindergarten using the standardized Dynamic Indicators of Basic Early Literacy Skills (DIBELS) assessment. Based on the DIBELS results, the children who were considered at risk scored one standard deviation below the mean or more. A control group and an at-risk intervention group were formed. The intervention group received phonological awareness, vocabulary and letter knowledge, and narrative comprehension every day for 30 minutes. The control group received some supplemental intervention, which was mostly focused on alphabetic principles. The results of the Catts et al. (2016) study showed that RTI in language assessed in the beginning of kindergarten predicted reading comprehension at the end of third grade. Specifically, in the study, there was evidence that the intervention of vocabulary instruction predicted increased reading outcomes. One limitation of the Catts et al. (2016) study was that the intervention and assessment protocol used was extremely lengthy to implement in a classroom without additional staff.

Scanlon et al. (2005) conducted a study near Albany, New York, involving 1,373 kindergarten students who were at risk for reading difficulties. The students (98% Caucasian) were followed through third grade. The purpose of the study was to explore the development of RTI indicators to help identify kindergarten students who might have
been at risk for reading difficulties after first grade and beyond. There were 462 kindergarteners who were identified as at risk from the universal screening. These students scored in the 30th percentile or less on the Letter Identification subtest of the Woodcock Reading Mastery Tests–Revised (WRMTR). Students were randomly assigned into two groups: an intervention group and a typical comparison group. The intervention group received additional instruction in a small group two times a week for 30 minutes, and data were based on each student’s needs. The typical control group received remedial services from a standard curriculum that was available in their school. Each group of students was assessed at three different points throughout the school year. Students within the intervention group continued treatment in first grade if their composite scores on the letter identification, word identification, and work attack subtests were below the benchmark midpoint (Scanlon et al., 2005).

The findings of the Scanlon et al. (2005) study suggest that children at risk for reading difficulties should be identified at the beginning of kindergarten, and they should be provided with strategic interventions to prevent reading failure. Of the students who received interventions at the beginning of kindergarten, 84% were meeting grade-level expectations at the end of third grade. Only 16% of the students demonstrated the need for further intervention. The data from this study indicate that the use of a universal screening at the beginning of kindergarten could significantly reduce the number of children who are at greater risk of having reading difficulties (Scanlon et al., 2005).

One limitation of the Scanlon et al. (2005) study was that the preventative approach to RTI instruction was not described in detail. Denton (2012) noted that research reviews and meta-analyses have identified that students with reading difficulties
benefit from instruction that is purposeful, that targets specific instruction and objectives, where mastery of skills are carefully monitored, and where various opportunities are given for practice (Denton, 2012).

Similarly, another study, Simmons et al. (2011), examined if the identification of at-risk kindergarten students who were then given RTI would improve and maintain their grade-level performance by third grade. The researchers investigated the performance of 41 students across a period of 4 years from kindergarten to Grade 3. At-risk criteria were determined by implementing the DIBELS assessment, and students who fell below the 30th percentile were considered at risk. All students received intensive, small-group intervention in kindergarten that focused on phonemic awareness, decoding, oral reading fluency, and comprehension. Furthermore, students continuing to perform below the 30th percentile at the beginning of first grade continued to receive RTI. All students’ progress was monitored in the fall of second and third grade, and the students were provided with intervention if their scores fell below the 30th percentile on the DIBELS. Simmons et al. (2011) study found that 93% of the students who received intervention at the beginning of kindergarten, read at or above grade level by the end of third grade. The findings of the study were limited to a small sample number of children. In addition, progress monitoring was only conducted once throughout the year except for the first grade in which students’ progress were monitored twice, which may have led to false positives (Fuchs & Vaughn, 2012).

In all these studies, the findings were consistent with the notion that reading difficulties should be identified early in the kindergarten year, and RTI instruction should be provided to students considered at risk for reading failure (Denton, 2012; Fuchs &
Fuchs, 2006). The early grades provide a unique chance to implement RTI instruction in which reading failure can be prevented (O’Connor et al., 2005). In addition, it was evident from these studies that students who were struggling with reading and received intervention in timely, small-group formats, where instruction was informed by monitoring their progress, could reduce reading failure significantly.

**Gender Differences**

Deficits in reading that occur early in a child’s school experience can lead to many negative outcomes (Allan, Joye, & Lonigan, 2017). Research states that children entering school without having “basic learning related skills are at increased risk of poor academic achievement” (Walker & Berthelsen, 2017, p. 70). Ayers (1909) articulated, as far back as 1909, that males enter school with a deficit in reading. In addition, preschool girls have had higher levels of emergent literacy skills than preschool boys (Justice, Invernizzi, Geller, Sullivan, & Welsch, 2005).

A study conducted by Chatterji (2006) compared test scores of 2,296 male and female kindergarten and first-grade students in 184 schools from the Early Childhood Longitudinal Study (ECLS). This study reveals that males perform below females in the following categories: (a) letter recognition, (b) beginning and ending sounds, (c) rhyming sounds, (d) print familiarity, (e) receptive vocabulary, and (f) vocabulary. Furthermore, the Chatterji (2006) study found that the size of the male discrepancy increased from 0.17 in kindergarten to 0.31 in first grade. The study indicates that gender differences exist between male and female children when entering school.

Similarly, a study conducted by Below, Skinner, Fearrington, and Sorrell (2010) assessed for gender differences in the reading skills of 1,218 kindergarten through
Grade 5 students. The study included 606 male students and 612 female students from three elementary schools, which were in a rural Southeastern United States public school district. All schools used a research-based core curriculum; however, two schools used the Scott Foresman reading curriculum, and one school used the Wilson Reading Foundation curriculum. The study utilized the DIBELS assessment as the benchmark assessment. The DIBELS is a standardized valid and reliable assessment that measures the performance of early literacy skills including phonemic awareness, phonics, orthography, and fluency (Below et al., 2010). The school district collected five standardized assessments from the DIBELS including: (a) initial sound fluency, (b) letter name fluency, (c) phoneme segmentation fluency, (d) nonsense word fluency, and (e) oral reading fluency. The assessments took place three times a year in the fall, winter, and spring. School psychologists, classroom teachers, and graduate students were trained and administered the assessments.

A two-way repeated ANOVA revealed that both male and female students showed improvement; however, the results show that females scored higher than males on initial sound fluency, letter naming fluency, phoneme segmentation fluency, and nonsense word fluency. In Grades 1-3 and 5, there were no significant gender differences with oral reading fluency (Below et al., 2010). In Grade 4, oral reading fluency scores of females were significantly higher than boys. The kindergarten girls scored significantly higher on all four literacy skills, which supports the notion that girls enter school with stronger literacy skills than boys (Justice et al., 2005).

A study conducted by Mohr and Price (2017) examined gender differences of 372 children who were learning to read. The children ranged in age from 4 to 5 years of age.
This study took place in England and incorporated a mix of 16 schools located in rural and urban areas. The study design incorporated three different interventions: (a) nonphonetic decodable vocabulary, (b) phonetic decodable vocabulary, and (c) synthetic phonics and decodable vocabulary. All schools used standardized assessments, such as the British Picture Vocabulary Scale and the York Assessment of Reading for Comprehension. The study utilized an independent $t$ test, which analyzed the pretest and posttest results. For students who participated in the synthetic phonics intervention, females scored approximately 7 months ahead of males on both the pre- and posttest assessments of the British Picture Vocabulary Scale. The gender gap widened for students who received interventions in nonphonetic decodable vocabulary compared to the other two intervention methods (Mohr & Price, 2017). The nonphonetic decodable vocabulary intervention showed the most gains for both males and females; however, females scored higher than males.

**Chapter Summary**

As seen from the studies, the importance of early intervention is evident. Even though a research-based, multitiered approach is mandated by IDEA (2004), there is an abundance of inconsistency in reading instruction within tiers. When tiers do not offer increasingly intense instruction, students do not improve in reading achievement (Mellard, McKnight, & Jordan, 2010). Speece and Walker (2007) posited that to prevent reading failure, schools need to identify which model or combinations of models are most effective for struggling readers. In addition, in order to effectively support students, Balu et al. (2015) stated that schools must evaluate the efficacy of RTI in their district.
Chapter 3: Research Design Methodology

The purpose of this chapter is to explain the research method the researcher used in this study. The researcher used a quantitative approach, implementing an ex post facto design and utilizing a retrospective cohort. To analyze the data, the researcher used an ANOVA.

General Perspective

The Individuals with Disabilities Education Improvement Act (IDEA, 2004) mandates that schools implement a tiered structure to prevent reading failure. RTI is a tiered instructional model that is used for the prevention of reading failure in all 50 United States (Otaiba et al., 2014). RTI emphasizes research-based instruction with increasing levels of intensity to support students who are struggling with reading (Weiss & Friesen, 2014). According to Hoover and Love (2011), a three-tiered RTI model that provides early support and progress monitoring to young students is most effective in closing the achievement gap in reading (Figure 3.1).

Reading serves as a basis for learning, and a deficit in reading causes lifelong challenges (Brynner, 2008). Students who begin school with a weakness in reading skills have a hard time catching up to their peers. Early identification of poor readers is fundamental in solving this problem (Wanzek et al., 2014). Of all students entering kindergarten, 60% are not ready for school (USDOE, 2015). Kindergarten readiness is
the ability to demonstrate skills that enable a child to participate and succeed in school (Cooke et al., 2010). Foundational early literacy skills in kindergarten have been discovered to be predictive of later reading achievement (O’Connor, 2000; O’Connor et al., 2005; Scanlon et al., 2005). Therefore, kindergarten is a critical window of opportunity in which educators can prevent reading failure (O’Connor et al., 2005). Research suggests that reading failure can be reduced with a quality core reading intervention program (Case et al., 2014). Systematic implementation of RTI can possibly be an effective tool to preventing reading failure (Stahl, 2016).

The purpose of this quantitative study was to examine the impact a three-tiered RTI model had on nonclassified kindergarten students’ reading achievement. This study used an ex post facto design utilizing a retrospective cohort. The researcher examined if the independent variables, RTI tier and time, had an impact or effect on the dependent

Figure 3.1. Response to Intervention Model of Tiered Instruction (NYSED, 2014).
variable, reading achievement. Reading achievement was defined as the increase in score on the standardized test for the achievement of reading early literacy assessment (STAR-ELA). RTI is defined as the “school-wide framework for which students at risk for reading difficulty are identified and provided with evidence-based and data-informed instruction” (Denton, 2012, p.1). At the time of this research, the school district in this study had implemented an RTI model including: (a) a core instructional curriculum, (b) a universal screener, (c) data analysis, (d) three tiers of intensity, and (e) progress monitoring (Appendix A). The hope is that the data from this study can be used to guide school districts in the effectiveness of an RTI model on student reading scores, as well as to add to the body of research involving RTI strategies and kindergarten reading scores.

Therefore, this study examined the following questions:

1. To what extent does the implementation of Response to Intervention (RTI) impact nonclassified kindergarten students’ reading ability as measured by scores on the STAR Early Literacy Assessment (STAR-ELA)?

2. Is there a relationship between improvement in reading scores on the STAR-ELA and the types of RTI provided to nonclassified struggling kindergarten readers?

3. To what extent, if any, does gender moderate the impact of Tier 2 RTI on reading scores of nonclassified struggling kindergarten readers as measured by scores on the STAR-ELA?

According to the research, the gold standard for evaluating interventions is the random control trial in which participants are randomly assigned to a treatment and control group (Smith, 2014). However, in education there are some ethical concerns with
this type of intervention, which is having subjects excluded from treatment when the participants might need it most (Cappelleri & Trochim, 2015). Therefore, this study implemented an ex post facto retrospective design using pretest and posttest archival data to address the research questions (Gravetter & Forzano, 2016). An ex post facto design was used because the researcher did not manipulate any of the variables.

A quasi-experimental study was chosen because the participants were not randomly assigned (Leedy & Ormrod, 2010). Assignment of the students was based on the universal screener, STAR-ELA. A retrospective design was utilized because the researcher did not manipulate any of the variables; the intervention or action had previously been accomplished. According to Schenker and Rumrill (2004), for ethical reasons, the variables in a retrospective ex post facto design cannot be experimentally manipulated. The researcher’s goal was to determine whether the independent variables (RTI tier and time) affected the outcome, or the dependent variable (reading achievement or reading scores), by examining pretest and posttest scores. The researcher examined student performance on the STAR-ELA assessment in the 2016-2017 school year. The RTI framework incorporates three tiers of interventions; however, in this study, only two tiers were analyzed.

Research Context

The fundamental rule of a research study is that the research questions dictate the research design (Kerlinger & Lee, 2000). This quantitative quasi-experimental study examined the impact RTI had on kindergarten reading scores; therefore, it was appropriate to use an ex post facto design. By its design, ex post facto research explores occurrences that have previously happened (Johnson & Christensen, 2008). In addition,
this study investigated the relationship between the dependent variable (reading achievement/scores) and the independent variables (RTI tier and time) when manipulation of the independent variable was not feasible (Ary, Jacobs, & Sorensen, 2010). The research objectives described in the purpose statement required a quantitative approach. The impact of the interventions, as shown by academic standardized test scores, warranted a numerical analysis.

**Research Participants**

Data used for this study consisted of archival records collected during the 2016-2017 school year by one suburban public school district located within the Hudson Valley area of New York State. At the time of this research, the district contained five schools, 245 teachers, and 3,213 students (New York State Education Department [NYSED], 2016). Tables 3.1 and 3.2 present the enrollment and district demographics.

Table 3.1

*District Student Enrollment Breakdown 2016-2017*

<table>
<thead>
<tr>
<th>Grade</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-1</td>
<td>377</td>
</tr>
<tr>
<td>2-5</td>
<td>1,042</td>
</tr>
<tr>
<td>6-8</td>
<td>769</td>
</tr>
<tr>
<td>9-12</td>
<td>925</td>
</tr>
<tr>
<td>Total Student Enrollment</td>
<td>3,113</td>
</tr>
</tbody>
</table>
Table 3.2

District Demographics for the Student Population 2016-2017

<table>
<thead>
<tr>
<th>Variable</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>70</td>
</tr>
<tr>
<td>Asian</td>
<td>13</td>
</tr>
<tr>
<td>Hispanic</td>
<td>11</td>
</tr>
<tr>
<td>Multi-Racial</td>
<td>4</td>
</tr>
<tr>
<td>African American</td>
<td>2</td>
</tr>
<tr>
<td>Graduation Rate</td>
<td>96</td>
</tr>
<tr>
<td>Attendance</td>
<td>96</td>
</tr>
<tr>
<td>Economically Disadvantaged</td>
<td>0</td>
</tr>
<tr>
<td>English Language Learners</td>
<td>6</td>
</tr>
</tbody>
</table>

Participants in this quantitative study included students in kindergarten that ranged from 4- through 6-years of age. The researcher examined and analyzed data for student reading achievement from one school that implemented a three-tiered RTI model (Appendix B). The total elementary school population consisted of 377 students of which 180 were kindergarten students. All 180 kindergarten students were assessed using the STAR-ELA; however, there were 39 identified students with an individualized education plan (IEP) who were excluded from the study in addition to a co-teaching class. Therefore, data from 141 participants were examined. Consent was obtained from the school district to analyze the data, but individual consent from the students was not necessary because the data was archival (Appendix C and D). To protect the identity of the participants, student names were not given to the researcher. Identification numbers were assigned to each student, and the numbers were matched on the pretest and posttest data. Assessment data were kept online in a password-protected file by the researcher. In this study, the school district administrators had previously established the teachers, classrooms, core-reading program, RTI framework, and reading intervention programs.
Therefore, this study utilized ex post facto data because the cause and effect of the intervention had already occurred, and the researcher did not manipulate any of the variables. The school district collected benchmark data on kindergarten students as part of the RTI process. The data consisted of 180 kindergarten students scores collected from the STAR-ELA. The quantitative data collection occurred at two times during the school year, fall and spring of the 2016-2017 school year. It is noteworthy to state that all the children in this study had attended a preschool prior to entering kindergarten. It is also important to state that this school district implemented the Wonders Reading Program (Wonders). Wonders is a comprehensive research-based K-6 English Language Arts program that is designed to prepare all students for college and career readiness in the 21st century (McGraw-Hill Education, 2018). In addition, Wonders improves student literacy by providing students with a core literacy-based foundation (McGraw-Hill Education, 2018).

**Instruments Used in Data Collection**

The STAR-ELA is a criterion-referenced diagnostic assessment implemented on a computer to measure the reading skills of students in Pre-K through Grade 3 (Renaissance Learning, 2017). The STAR-ELA is one of the first assessments to be rated as highly reliable and highly valid by the National Center on Response to Intervention (NCRTI, 2010). The reliability of the STAR-ELA for internal consistency is 0.97, which is very high, and 0.79 for retest consistency (Renaissance Learning, 2013). The validity of the STAR-ELA is 0.87, which is considered a strong correlation (Renaissance Learning, 2013). The National Center on Response to Intervention at the American
Institutes for Research described the STAR-ELA as accurate, reliable, valid, and generalizable (NCRTI, 2014).

Data were collected throughout the 2016-2017 school year during the fall (September) and spring (May). Students received a scaled score and a percentile rank for each of the seven domains, which is a “criterion-referenced score that indicates a student’s percent of mastery of skill within that domain” (Renaissance Learning, 2017, p. 2). Default benchmark scores were assigned to each assessment period. Benchmarks are “minimum performance levels students are expected to reach by certain points of the year in order to meet end-of-year performance goals” (Renaissance Learning, 2017, p. 4). Default benchmarks were set at the 10th, 25th, and 40th percentile for each grade based on 2015 norms (Appendix E). The STAR-ELA data report contained seven subdomain categories, covering 41 skill sets, containing 145 separate literacy skills. The literacy domain of the STAR-ELA assessed: (a) alphabetic principle, (b) concept of word, (c) visual discrimination, (d) phonemic awareness, (e) phonics, (f) structural analysis, and (g) vocabulary (Appendix F).

During the 2016-2017 school-year, kindergarten students were given the STAR-ELA once in the fall and then, again, in the spring using an iPad and headphones. Support staff assisted the students with the login information as the STAR-ELA utilized a unique code number assigned to each student. As the students answered questions correctly, the difficulty of the questions increased. When students answered four questions incorrectly and consecutively, the assessment reverted to easier questions, and it ultimately stopped. Data were automatically entered into the computer and staff had immediate access to assessment results.
Procedures for Data Analysis

The researcher analyzed the data with the Statistical Package for the Social Sciences (SPSS) and entered the pretest and posttest data scores for each student into the SPSS program. An ANOVA was used to analyze the data. An ANOVA is a statistical method used to test differences between two or more means (Trochim, 2006). A simple $t$ test could be used to analyze this data. However, using a $t$ test would not be reliable because there were more than two samples being compared. Doing so would increase the chance of a Type 1 error, which is the incorrect rejection of the null hypothesis, a false positive (Siegel, 1956). An ANOVA was most effective in comparing the means of the data set.

A factorial ANOVA was used to analyze the data. Factorial ANOVAs are used in research when one wants to analyze differences on a continuous dependent variable between two or more independent, discrete grouping variables (Trochim, 2006). In this analysis, reading achievement was compared by both time and tier. The time variable had two groups (pretest fall and posttest spring). The tier had two groups (Tier 1 and Tier 2). The ANOVA uses the $F$ test, which allows researchers to make the overall comparison on whether a group means differ. The $F$ test is the ratio of two independent variance estimates of the same population variance. Considering an alpha of 0.05, if the calculated $F$ value is larger than the critical $F$ value, after accounting for degrees of freedom, the null hypothesis ($H_0$) would be rejected and the alternative hypothesis ($H_a$) would be accepted. $F$-test degrees of freedom are calculated between groups ($K - 1$) and within groups ($N - K - 1$) where $K$ equals the number of groups. The results of the factorial ANOVA are presented in the form of main effects and the interactions among
the study variables. Post hoc analyses were conducted consisting of a series of independent \( t \) tests.

The assumptions of normality and homogeneity of variance were assessed. Normality assumes that the scores are normally distributed (symmetrical bell shape) and were assessed using the one sample Kolmogorov Smirnov (KS) test. Homogeneity of variance assumes that both groups have equal error variances and homogeneity was assessed using a Levene’s test.

**Chapter Summary**

This quantitative study implemented an ex post facto design that utilized archival data from a retrospective cohort. Data were collected from the STAR-ELA in the 2016-2017 school year. Data were sorted based on a baseline benchmark score indicated on the STAR-ELA. Based on student performance, relative to the benchmark score, the students’ data were assigned to one of three RTI categories. The participants included 180 nonclassified kindergarten students within one suburban school district. An ANOVA was used to analyze pretest and posttest data. The researcher used SPSS to calculate the statistical analysis. The results of the analysis, here, are presented in the form of tables, charts, and reports. The researcher used these data to answer the research questions and reject the null hypothesis. The goal of this study was to analyze the impact of the RTI framework and interventions on the reading scores of nonclassified kindergarten students. This study adds to the body of research on RTIs.
Chapter 4: Results

The overarching goal of this retrospective, quantitative, quasi-experimental, ex post facto, pretest-posttest research was to evaluate the impact an RTI framework had on nonclassified kindergarten students’ reading achievement. There were three specific goals. First, the study was to evaluate the combined effectiveness of Tier 1 and Tier 2 reading interventions on nonclassified kindergarten students’ reading achievement to determine if the implementation of these RTIs impacted the reading ability of kindergarten students. Second, the study goal was to compare the efficacy of the Tier 1 and Tier 2 interventions, separately, to determine if there was any relationship between the type of RTI employed, that is, Tier 1 or Tier 2, and to gather the resulting improvements in reading achievement. Finally, the goal was to examine the potential role gender had in moderating the effect of Tier 2 interventions on reading achievement, to see if boys and girls benefitted equally or unequally from the intervention. Reading achievement, the study’s dependent variable, was measured using scaled scores on the STAR-ELA. This instrument was used in measuring reading achievement both before and after the RTI to enable pretest-posttest comparisons.

The nonclassified kindergarten students who participated in this study were assigned to RTI tiers based on their benchmark scores on the STAR-ELA. Students who performed above the cutoff score of 496 on the STAR-ELA were assigned to Tier 1, while students assigned to Tier 2 scored at or below 496. Tier 1 students received a core curriculum instruction, the Wonders Reading Program, from their classroom teachers.
Tier 2 students received additional staff-implemented intensive small-group instruction. Data were gathered from a total of 141 students. Of these, 104 were assigned to Tier 1 (50 males and 54 females), and 37 were assigned to Tier 2 (14 males and 23 females). All students were between 4 and 6 years of age.

The data analyzed in this study were archival, and they were collected during the 2016-2017 school year from a suburban elementary school located in the Hudson Valley area of New York State. Reading achievement at pretest was evaluated in fall 2016, and the posttest evaluation was performed in spring 2017. There were no missing data.

**Research Questions**

The three research questions that were addressed in this study are listed below along with their corresponding null and alternative hypotheses:

1. To what extent does the implementation of Response to Intervention (RTI) impact nonclassified kindergarten students’ reading ability as measured by scores on the STAR Early Literacy Assessment (STAR-ELA)?
   
   \( H_{1A} \): RTI significantly improves the reading ability of nonclassified kindergarten students.

2. Is there a relationship between improvement in reading scores on the STAR-ELA and the types of RTI provided to nonclassified struggling kindergarten readers?
   
   \( H_{20} \): There is no relationship between the types of RTI provided to nonclassified struggling kindergarten readers and the amount of improvement in their reading ability.
H2A: There is a significant relationship between the types of RTI provided to nonclassified struggling kindergarten readers and the amount of improvement in their reading ability.

3. To what extent, if any, does gender moderate the impact of Tier 2 RTI on reading scores of nonclassified struggling kindergarten readers as measured by scores on the STAR-ELA?

H30: Gender does not moderate the effectiveness of Tier 2 RTI on the reading scores of nonclassified struggling kindergarten readers as measured by scores on the STAR-ELA.

H3A: Gender significantly moderates the effectiveness of Tier 2 RTI on the reading scores of nonclassified struggling kindergarten readers as measured by scores on the STAR-ELA.

Data Analysis and Findings

All data manipulations and statistical analyses for this study were performed using IBM SPSS 24.0, G*Power, or they were performed with a hand calculator. Two 2 × 2 mixed-subjects factorial ANOVAs were used in addressing the study’s three research questions. Different portions of the first ANOVA were relevant to both Research Question 1 and Research Question 2. In the type of ANOVA used, the between-subjects factor was the RTI Tier, with two levels (Tier 1 and Tier 2); the within-subjects factor was Time, with two levels (pretest and posttest); and the dependent variable, Reading Achievement/Scores, was measured by the STAR-ELA. The main effect of Time addressed Research Question 1 by evaluating improvements in reading achievement from the pretest to posttest, which were collapsed across the RTI tiers. In this same factorial
ANOVA, the RTI Tier × Time interaction effect addressed the second research question by evaluating the degree to which Tier 1 and Tier 2 interventions were differentially effective, that is, whether the type of RTI was associated with the magnitude of reading improvement from pretest to posttest. The second factorial ANOVA was used to analyze data only from the students who received Tier 2 interventions and to address Research Question 3. In that ANOVA, the between-subjects factor was Gender, with two levels (male and female); the within-subjects factor was Time, with two levels (pretest and posttest); and the scores on the STAR-ELA, again, served as the dependent measure of reading achievement. The Gender × Time interaction effect from the second ANOVA was of primary interest in evaluating whether gender moderated the effectiveness of the Tier 2 intervention.

In addition to these factorial ANOVAs, the study made use of Bonferroni-adjusted post hoc comparisons of cell means to explore sources of significant effects identified by some of the ANOVA $F$ tests. The Bonferroni correction is used to adjust probability values because of an increased risk of Type 1 errors when making multiple statistical tests (Armstrong, 2014). The Bonferroni correction bypasses the problem of Type I errors that conclude there is a significant difference present when there is not (Armstrong, 2014). Post-hoc comparisons are performed after an ANOVA has revealed significant main and/or interaction effects. The effects that the ANOVA detects as significant often involve differences among several group means, but the ANOVA does not indicate which specific group means differ to produce the significant main and/or interaction effects. The post-hoc comparisons are used after the fact, to identify exactly which group means differed significantly and resulted in the significant ANOVA $F$ test.
Pairwise comparisons of even a small number of means (i.e., comparing two means at a time) can result in a substantial number of significance tests. With \( k \) means, there are \( \frac{k^2-k}{2} \) possible pairwise comparisons. The significance tests associated with each of these comparisons carries some probability of a Type I error, which is equal to the chosen level of statistical significance. Over a series of several pairwise comparisons, the total number of Type I errors to be expected is equal to the number of comparisons, multiplied by the chosen significance level. For instance, in a series of 20 comparisons, each using the .05 level of significance, one test would be expected to yield significance in the absence of a true difference in the population: \( .05 \times 20 = 1 \). In order to hold the probability of a Type I error at a reasonable level (e.g., .05) across a series of post-hoc comparisons, a Bonferroni adjustment is commonly used (Armstrong, 2014). Using the Bonferroni procedure, the level of significance selected for each of a series of significance tests is determined by dividing the desired family-wise Type I error rate by the number of tests. Accordingly, if one wishes to maintain a Type I error rate of .05 across the series of \( c \) post-hoc comparisons, each of those comparisons is evaluated using a significance level of \( .05/c \). Alternatively, the level of significance that is associated with a post-hoc comparison is adjusted prior to reporting by multiplying the exact significance by the number of comparisons in the series. That is the approach to Bonferroni-adjusted post-hoc comparisons that is taken by SPSS.

**Research Question 1.** To what extent does the implementation of Responses to Intervention (RTI) impact nonclassified kindergarten students’ reading ability as measured by scores on the STAR Early Literacy Assessment (STAR-ELA)?
The first research question did not differentiate between the tiers of RTI to which the students were exposed, but rather, it sought to determine if, across those RTI tiers, students showed improvement in their reading achievement. The research hypothesis (H1A) was that the combined data of all students in the study, that is, both tiers, would show significant improvement in reading achievement. The corresponding null hypothesis (H10) was that the RTI approach would have no significant effect in improving reading achievement.

**RQ1 analyses.** Research Question 1 was addressed using a $2 \times 2$ mixed-subjects factorial ANOVA. The between-subjects factor was RTI tier, with two levels, Tier 1 and Tier 2. The within-subjects factor was time, with two levels, pretest and posttest. The dependent variable was reading achievement/score measured at pretest and posttest with the STAR-ELA. The main effect of the time factor was the focus of this analysis because a significant $F$ test of that main effect would indicate that, collapsed across tiers, students scored significantly differently at the pretest and posttest. It was hypothesized that the time main effect would be significant, and that the mean posttest reading achievement scores would be higher than the mean pretest reading achievement scores. The time main-effect $F$ test and an inspection of the pretest and posttest means to determine the direction of the change from pretest to posttest was sufficient to answer Research Question 1. However, what the $F$ test for the main effect of time did not indicate was whether the change from the pretest to posttest was observed in each of the separate Tier 1 and Tier 2 groups. While the first research question did not require an answer to that question, curiosity motivated separate comparisons of the pretest to posttest improvements of the Tier 1 and Tier 2 students. Those comparisons were performed
using Bonferroni-adjusted post-hoc comparisons of: (a) Tier 1 pretest versus Tier 1 posttest performance, and (b) Tier 2 pretest versus Tier 2 posttest performance.

The results of any statistical analysis are only valid to the degree that the data in the analysis display certain characteristics. Therefore, before performing the $2 \times 2$ mixed-subjects factorial ANOVA, each of several statistical assumptions was tested. First, each cell of the factorial design was checked for outliers. Outliers exert a disproportionate effect on statistical outcomes, and they are statistically aberrant and unrepresentative of the rest of the sample. Outliers were screened by standardizing the STAR-ELA scores within each cell of the factorial design in search of $z$-scores exceeding $\pm 3.30 (p < .001$ in a normal distribution). One Tier 2 posttest outlier was identified with a STAR-ELA score of 434 ($z = -3.54$). All data from this outlier were excluded from the first ANOVA, leaving 140 cases for the analysis, 104 in Tier 1 and 36 in Tier 2.

The ANOVA also assumes that the data are normally distributed within each cell of the factorial design. Violations of that assumption can distort the exact significance levels ($p$-values) for the $F$ tests of the main and interaction effects. Data normality was evaluated both visually, by examining frequency histograms of the STAR-ELA scores, and statistically, by calculating measures of skewness and kurtosis for each distribution and also by using the Shapiro-Wilk test of normality. Figure 4.1 shows frequency histograms for the STAR-ELA scores for each cell of the $2 \times 2$ factorial design. Table 4.1 provides descriptive statistics as a function of RTI tier and time, and Table 4.2 summarizes the results of the Shapiro-Wilk tests for normality. George and Mallery (2003) suggested that distributions showing skewness and kurtosis measures exceeding $\pm 1.0$ should be considered not normal, and Meyers, Gamst, and Guarino (2013)
recommended that the Shapiro-Wilk test of normality should be evaluated using a stringent significance level \( p < .001 \) to mitigate that test’s sensitivity to trivial departures from normality, especially when sample sizes are greater than 50. Visual inspection of the frequency histograms showed that all distributions showed some deviations from normality, but these were particularly pronounced in the distributions of the pretest scores of both tier groups, where skewness values exceeded ±1.00, and the Shapiro-Wilk tests were significant at or beyond the .001 level. When the assumption of data normality is violated, it is sometimes possible to use a data transform (e.g., log10 or square root) to reshape the distribution toward a more normal shape (Warner, 2008). However, these data transforms are only effective when the same transform can be applied to all cells in the factorial design. In this study, that was not possible because Tier 1 pretest scores were negatively skewed, and Tier 2 pretest scores were positively skewed. The Tier 1 pretest scores showed little skewness, but they were leptokurtic, and the Tier 2 posttest scores were reasonably normal. Another alternative that is available to the data analyst when the normality or some other assumption is violated is to use a statistical alternative that is more robust with respect to the violations of the violated statistical assumptions. However, there is no such nonparametric alternative to the mixed-subjects factorial ANOVA (Kirk, 2013). Tabachnick and Fidell (2013) offered some support for this choice when sample sizes are reasonably large, noting that:

For grouped data, it is the sampling distribution of the means of variables that are to be normally distributed. The Central Limit Theorem reassures us that, with sufficiently large sample sizes, sampling distributions of means are normally distributed regardless of the distributions of variables. For example, if there are at
least 20 degrees of freedom for error in a univariate ANOVA, the $F$ test is said to be robust to violations of normality of variables, provided there are no outliers. (pp. 78-79)

In this factorial ANOVA, Tier 1 and Tier 2 were both of sufficient size that the error term degrees of freedom for the main and interaction effects were 138.

![Figure 4.1. Frequency histograms of pretest and posttest STAR-ELA reading achievement test scores for Tier 1 students (top row) and Tier 2 students (bottom row).](image)
Table 4.1

*Descriptive Statistics on the STAR-ELA Reading Achievement Test at Pretest and Posttest for Tier 1 and Tier 2 Students*

<table>
<thead>
<tr>
<th>Tier</th>
<th>Fall Pretest</th>
<th></th>
<th></th>
<th></th>
<th>Spring Posttest</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>$M$</td>
<td>$SD$</td>
<td>Skew</td>
<td>Kurtosis</td>
<td>$n$</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>1</td>
<td>104</td>
<td>600.9</td>
<td>78.4</td>
<td>1.04</td>
<td>0.46</td>
<td>10</td>
<td>733.5</td>
<td>79.3</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>454.1</td>
<td>36.2</td>
<td>1.03</td>
<td>0.20</td>
<td>36</td>
<td>693.5</td>
<td>57.8</td>
</tr>
</tbody>
</table>

Table 4.2

*Results of Shapiro-Wilk Tests of Normality of Distributions for Tier 1 and Tier 2 Students at Pretest and Posttest*

<table>
<thead>
<tr>
<th>Tier</th>
<th>Time</th>
<th>Statistic</th>
<th>df</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pretest</td>
<td>0.903</td>
<td>104</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Posttest</td>
<td>0.959</td>
<td>104</td>
<td>.003</td>
</tr>
<tr>
<td>2</td>
<td>Pretest</td>
<td>0.880</td>
<td>36</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>Posttest</td>
<td>0.959</td>
<td>36</td>
<td>.206</td>
</tr>
</tbody>
</table>

The remaining two statistical assumptions of the ANOVA—homogeneity of between-group variances and covariances—were evaluated using a preliminary run of the ANOVA in order to take advantage of some of the diagnostic tools available in the output. The mixed-subjects factorial ANOVA is based in part on the assumption that the variability of the dependent variable (STAR-ELA scores) is approximately equal between the groups that form the between-subjects factor (RTI tier). Violation of the homogeneity
of variance assumption can distort the exact significance levels reported from the ANOVA. The homogeneity of variances was tested using Levene’s test, which found that score variances of the Tier 1 and Tier 2 students were significantly unequal, both at pretest, $F(1, 138) = 16.01, p < .001$, and at posttest, $F(1, 138) = 10.58, p = .001$. It was concluded that the homogeneity of variance assumption was violated. The mixed-subjects factorial ANOVA also assumed that the scores on the within-subjects factor show similar covariances across all groups that form the between-subjects factor. In terms of this study, homogeneous covariances would mean that correlations between the pretest and posttest scores are approximately equal for the Tier 1 and Tier 2 students. That assumption was evaluated using Box’s M test in this study. The test was significant, Box’s $M = 26.13, F(3, 71741.07) = 8.52, p < .001$, indicating a violation of the assumption. The pretest-posttest correlation among Tier 1 students was $r(102) = .44, p < .001$ (two-tailed), while in Tier 2 the pretest-posttest correlation was considerably lower, $r(34) = .15, p = .40$ (two-tailed).

In summary, several of the statistical assumptions upon which the mixed-subjects factorial ANOVA were based were violated by the characteristics of the data in this study—normality, homogeneity of variance, and homogeneity of covariance. The effect of those violations distorted the exact significance levels ($p$-values) associated with the ANOVA $F$ tests of the main and interaction effects. To mitigate against the possibility that those distortions might have caused some effects to appear to be statistically significant ($p < .05$), when the true significance level was $p > .05$, Meyers et al. (2013) recommended using a more stringent level of significance in evaluating the $F$ tests. Consistent with this recommendation, the ANOVA was used to address Research
Question 1 and the .001 level of significance rather than the more common .05 significance level (Dattalo, 2008).

**RQ1 findings.** The results of the $2 \times 2$ mixed-subjects factorial ANOVA are summarized in Table 4.3, and the cell means are plotted in Figure 4.2. The cell means and other descriptive statistics were presented previously in Table 4.1. It is the main effect of time that is of particular note in addressing Research Question 1. The main effect of time was extremely strong and statistically significant, $F(1, 138) = 594.26, p < .001$, partial $\eta^2 = .812$. That main effect and the inspection of the means, provided in Table 4.1 and plotted in Figure 4.2, indicate that students in this study scored significantly higher at the spring posttest than at the fall pretest. The null hypothesis that RTIs have no effect on the reading ability of nonclassified kindergarten students, was rejected; and the research hypothesis that RTI significantly improves the reading ability of nonclassified kindergarten students, was accepted.

As noted previously, what the $F$ test for the main effect of time did not establish is whether the pretest-posttest improvements in reading achievement were present in both Tier 1 and Tier 2 students. Bonferroni-adjusted post-hoc comparisons of the two groups’ pretest and posttest means shed further light on that question. Both Tier 1 and Tier 2 students showed significant improvement from the fall pretest to the spring posttest. The Tier 1 students improved an average of 132.59 points ($SE = 7.74, p < .001$), and the Tier 2 students improved an average of 239.39 points ($SE = 13.15, p < .001$). Not only did students who were exposed to the RTI framework show significant improvement in reading achievement, the students in both tiers of the RTI, who were examined in this study, showed significant improvement.
Table 4.3

Summary Table for 2 (RTI Tiers) × 2 (Time) Mixed-Subjects Factorial ANOVA

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tests of Between-Subjects Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RTI Tier</td>
<td>466885.83</td>
<td>1</td>
<td>466885.83</td>
<td>63.44</td>
<td>&lt;.00</td>
<td>.315</td>
</tr>
<tr>
<td>Error</td>
<td>101544.24</td>
<td>138</td>
<td>7359.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tests of Within-Subjects Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>1850147.32</td>
<td>1</td>
<td>1850147.3</td>
<td>594.26</td>
<td>&lt;.00</td>
<td>.812</td>
</tr>
<tr>
<td>RTI Tier × Time Interaction</td>
<td>152524.44</td>
<td>1</td>
<td>152524.44</td>
<td>48.99</td>
<td>&lt;.00</td>
<td>.262</td>
</tr>
<tr>
<td>Error</td>
<td>429641.89</td>
<td>138</td>
<td>3113.35</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.2. Mean performance on the STAR-ELA reading assessment among Tier 1 and Tier 2 students at fall and spring posttest.
**Research Question 2.** Is there a relationship between improvement in reading scores on the STAR-ELA and the types of RTI provided to nonclassified struggling kindergarten readers?

As described previously, both Tier 1 and Tier 2 students showed significant improvement in reading achievement from the pretest to the posttest assessments. What Research Question 2 focused on was whether or not these improvements were equal or differential. In other words, while recognizing that both groups improved significantly, was the improvement significantly greater for one group over the other?

*RQ2 analyses.* Research Question 2 was addressed with the same $2 \times 2$ mixed-subjects factorial ANOVA that was used in evaluating Research Question 1. To repeat, the between-subjects factor was RTI tier, with two levels (Tier 1 and Tier 2); and the within-subjects factor was time, with two levels (pretest and posttest). The dependent variable was reading achievement score, which was measured at both the pretest and posttest using the STAR-ELA assessment. Within this factorial, it was the RTI Tier $\times$ Time interaction effect that was relevant to Research Question 2. A significant interaction between the factors would indicate that the significant pretest-posttest improvements in performance, seen in both Tier 1 and Tier 2 students, which was already established in addressing Research Question 1, were significantly greater for one tier over the other. Inspection of the group means at pretest and posttest would then determine which group improved significantly more.

Tests of the statistical assumptions of the $2 \times 2$ within-subjects factorial ANOVA used in evaluating Research Question 2 were previously tested and reported above in conjunction with Research Question 1. The results of those tests do not need to be
repeated here except to note that it was determined that violations of some of the statistical assumptions of the factorial ANOVA did occur. To mitigate against possible distortions in the exact significance levels for the ANOVA $F$ tests caused by these violations, all effects were evaluated using a stringent level of significance ($p < .001$). That conservative standard was applied to the $F$ test of the RTI Tier $\times$ Time interaction effect that was the focus in this instance.

**RQ2 findings.** The RTI Tier $\times$ Time interaction effect was relatively strong and statistically significant, $F(1, 138) = 48.99, p < .002$, partial $\eta^2 = .262$. The means listed in Table 4.1 and plotted in Figure 4.2 support the following interpretation of this significant interaction. Although both Tier 1 and Tier 2 students improved significantly from the pretest to posttest, which was established previously in Research Question 1, the improvement shown by the Tier 2 students was significantly greater than that seen among the Tier 1 students. The null hypothesis associated with Research Question 2—that there was no relationship between the types of RTI and the amount improvement in reading ability—was rejected. The alternative hypothesis—that there was a significant relationship between the types of RTI and the amount of improvement in reading ability—was accepted.

Although no additional analyses were required to answer Research Question 2, the use of Bonferroni-adjusted post-hoc comparisons was motivated by curiosity to elaborate and expand upon the answer. How can the pattern of pretest-posttest changes, which are depicted in Figure 4.2 for Tier 1 and Tier 2 students, best be described?

Post-hoc comparisons showed that the Tier 2 students scored significantly lower than the Tier 1 students on the fall pretest (146.93 points lower, on average, $SE = 13.57,$
However, the difference between the two tiers shrank by the time of the posttest to only 40.03 points, on average ($SE = 17.59$). The difference between the tiers at posttest was not statistically significant at the $p < .001$ level of significance, which was adopted for use in this study to mitigate violations of the statistical assumptions ($p = .006$).

**Research Question 3.** To what extent, if any, does gender moderate the impact of Tier 2 RTI on the reading ability of nonclassified struggling kindergarten readers as measured by scores on the STAR-ELA?

It was established previously that the reading achievement performance of both the Tier 1 and Tier 2 students showed significant improvement from pretest to posttest (Research Question 1), and that the Tier 2 students showed significantly greater improvement than the Tier 1 students (Research Question 2). Research Question 3 focused on the struggling Tier 2 students to learn if their gender might have moderated the effectiveness of the Tier 2 intervention, that is, did Tier 2 RTI work better for boys or girls? The research hypothesis ($H_{3A}$) was that gender would influence how much improvement was seen from the pretest to posttest, although no prediction was made as to whether males or females would show the greater improvement. The null hypothesis ($H_{30}$) was that the pretest-posttest improvements would be about the same in magnitude among males and females.

**RQ3 analysis.** A $2 \times 2$ mixed-subjects factorial ANOVA was performed to address the study’s third research question. This ANOVA analyzed data only from the 37 students who were assigned to the Tier 2 intervention; no Tier 1 students were included in the analysis. The between-subjects factor was Gender, with two levels, male
and female. There were 14 males and 23 females in the analysis. The within-subjects factor was Time, with two levels, pretest and posttest. Scores on the STAR-ELA provided a measure of the dependent variable, Reading Achievement Scores. The $F$ test for the Gender $\times$ Time interaction effect was the focus of Research Question 3, because a significant interaction would indicate that the pretest-to-posttest improvement in reading achievement scores seen in Tier 2 students, which were established in both Research Questions 1 and 2, was of significantly different magnitudes for male and female Tier 2 students.

This second ANOVA was, again, preceded by tests of the statistical assumptions upon which the procedure was based. First, all cells of the factorial design were screened for outliers by standardizing scores within each cell and search for $z$-scores exceeding $\pm 3.30$. No outliers were identified. Next, the normality assumption was tested. Frequency histograms were generated for the STAR-ELA data in all cells of the $2 \times 2$ factorial design, measures of skewness and kurtosis were calculated, and the Shapiro-Wilk tests of normality were performed. Figure 4.3 shows the frequency the histograms, Table 4.4 provides the descriptive statistics, and Table 4.5 summarizes the results of the Shapiro-Wilk tests. Data distributions in all cells were somewhat negatively skewed, with values of skewness exceeding the $\pm 1.00$ values, which were used to identify seriously not-normal distributions in two cells—male pretest scores and female posttest scores. The data in these two cells also showed extreme leptokurtosis, with kurtosis values exceeding $+1.0$ in both cells. Despite the presence of two strongly skewed, leptokurtic data distributions, none of the Shapiro-Wilk tests of normality reached significance at the stringent .001 level.
Figure 4.3. Frequency histograms of pretest and posttest STAR-ELA reading achievement test scores for male (top row) and female (bottom row) Tier 2 students.

Table 4.4

_Descriptive Statistics on the STAR-ELA Reading Achievement Test at Pretest and Posttest for Male and Female Tier 2 Students_

<table>
<thead>
<tr>
<th>Gender</th>
<th>( n )</th>
<th>( M )</th>
<th>( SD )</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>( n )</th>
<th>( M )</th>
<th>( SD )</th>
<th>Skew</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>14</td>
<td>458.36</td>
<td>34.10</td>
<td>−1.79</td>
<td>3.84</td>
<td>1</td>
<td>705.5</td>
<td>0</td>
<td>−0.08</td>
<td>−0.24</td>
</tr>
<tr>
<td>Females</td>
<td>23</td>
<td>453.04</td>
<td>37.95</td>
<td>−0.80</td>
<td>−0.57</td>
<td>2</td>
<td>674.9</td>
<td>9</td>
<td>−1.31</td>
<td>3.11</td>
</tr>
</tbody>
</table>
Table 4.5

*Results of Shapiro-Wilk Tests of Normality of Distributions for Male and Female Tier 2 Students at Pretest and Posttest*

<table>
<thead>
<tr>
<th>Gender</th>
<th>Time</th>
<th>Statistic</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>Pretest</td>
<td>0.831</td>
<td>14</td>
<td>.013</td>
</tr>
<tr>
<td></td>
<td>Posttest</td>
<td>0.940</td>
<td>14</td>
<td>.419</td>
</tr>
<tr>
<td>Females</td>
<td>Pretest</td>
<td>0.872</td>
<td>23</td>
<td>.007</td>
</tr>
<tr>
<td></td>
<td>Posttest</td>
<td>0.895</td>
<td>23</td>
<td>.020</td>
</tr>
</tbody>
</table>

Although sample sizes were relatively small in this analysis, there were still 35 degrees of freedom for all error terms in the ANOVA. Consequently, as suggested by Tabachnick and Fidell (2013), the central limit theorem should be expected to provide some immunity with respect to the violation of the assumption of normality. This fact, and the failure of both the log10 and square-root data, transforms to provide appreciably improved normalization of the distributions, which led to the decision to proceed with an analysis of the raw data using a stringent significance level \( p < .001 \) to mitigate against possible distortions in the reported significance levels caused by not-normal data.

The remaining two statistical assumptions of the ANOVA, homogeneity of between-group variances and covariances, were evaluated using a preliminary run of the ANOVA to take advantage of some of the diagnostic tools available in the output. The assumption of homogeneity of variances was tested using Levene’s test. The males and females showed approximately equal variances in their STAR-ELA scores at both pretest, \( F(1, 35) = 1.47, p = .233 \) and at posttest, \( F(1, 35) = 0.52, p = .472 \); and it was concluded
that the homogeneity of variances assumption was satisfied. The homogeneity of
covariances assumption was tested with Box’s M test, Box’s $M = 1.57, F(3, 26588.66) =
0.49, p = .690$, which showed that correlations between the pretest and posttest scores
were approximately equal for male, $r(12) = -.08, p = .783$, and female students, $r(21) =
.04, p = .857$. Thus, the homogeneity of covariances assumption was satisfied.

**RQ3 findings.** The results of the $2 \times 2$ mixed-subjects factorial ANOVA are
summarized in Table 4.6, and the cell means are plotted in Figure 4.4. The cell means
were presented previously in Table 4.4. The effect of the chief interest in this ANOVA
was the Gender $\times$ Time interaction effect. The interaction effect was not significant, $F(1,
35) = 0.88, p = .897$, indicating that Tier 2 males and females showed approximately
equal improvement in reading achievement from the pretest to posttest. This finding
directly addressed Research Question 3, which asked if gender moderated the impact of
Tier 2 intervention on the reading achievement of struggling kindergarten students. The
data gathered in this study did not support the conclusion that gender influenced students’
responses to the Tier 2 intervention and it was concluded that there was insufficient

<table>
<thead>
<tr>
<th>Table 4.6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Summary Table for 2 (Gender) x 2 (Time) Mixed-Subjects Factorial ANOVA</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tests of Between-Subjects Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>5608.26</td>
<td>1</td>
<td>5608.26</td>
<td>1.77</td>
<td>.193</td>
<td>.048</td>
</tr>
<tr>
<td>Error</td>
<td>111206.34</td>
<td>35</td>
<td>3177.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tests of Within-Subjects Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>957178.30</td>
<td>1</td>
<td>957178.30</td>
<td>303.88</td>
<td>&lt;.001</td>
<td>.897</td>
</tr>
<tr>
<td>Gender $\times$ Time Interaction</td>
<td>2779.38</td>
<td>1</td>
<td>2779.38</td>
<td>0.88</td>
<td>.354</td>
<td>.025</td>
</tr>
</tbody>
</table>
evidence to reject the null hypothesis that gender does not moderate the effectiveness of a Tier 2 RTI on reading achievement scores. No further analyses were necessary in order to address Research Question 3. However, the ANOVA included additional output, which provided answers to some unasked questions. Those results are reviewed here, motivated, again, by curiosity. The main effect of gender was nonsignificant, $F(1, 35) = 1.77, p = .193$, indicating that, collapsing across the pretest and posttest assessments, male and female students in Tier 2 performed at approximately equal levels. There was a strong and statistically significant main effect of time, $F(1, 35) = 303.88, p < .001$, partial $\eta^2 = .897$, indicating that the students performed significantly higher at the spring posttest than at fall pretest. That improvement was already noted, however, in connection with Research Question 1 and Research Question 2.

Figure 4.4. Mean performance on the STAR-ELA reading assessment among male and female Tier 2 students at fall pretest and spring posttest.
**Post-hoc power analysis.** When a finding that was expected in a study fails to materialize in the form of a statistically significant effect, the question is whether that failure reflects the real absence of the sought-after effect in the population or if the failure was due to a Type II error. Type II errors occur when the expected effect actually exists in the population, but it failed to appear in the particular sample that was chosen for the study because of a sampling error, that is, it might be an idiosyncratic sample. Samples can be peculiar, and small samples are especially likely to fail to capture the characteristics of their parent populations due to sampling error. In this study, it is suspected that the efficacy of the Tier 2 reading intervention might be different for males and females, that is, that gender moderated the efficacy of the Tier 2 intervention. However, the Gender × Time interaction effect that would have demonstrated the moderating effect of gender failed to reach statistical significance. Consequently, it was important to evaluate the likelihood that the nonsignificant interaction effect in this study resulted from a Type II error. This evaluation was accomplished using a post-hoc power analysis performed using the G*Power software. Parameters input to the analysis were as follows. The strength of the interaction effect in the population, as measured by Cohen’s (1992) $f$ statistic, was set at $f = 0.25$, which represents an effect of medium strength. Cohen (1992) described medium effects as those that “would likely be visible to the naked eye of a careful observer” (p. 156). The probability of a Type II error was set at the standard level of $\alpha = .05$, despite the fact that a more stringent significance level was used in the ANOVA to mitigate the violation of the normality assumption. This decision was based on the reality that had the interaction effect achieved the .05 level of significance, but had not reached .001, it would most certainly have been highlighted.
anyway. The total sample size was $N = 37$; the number of groups was two; the number of measurements was two; the correlation between the pretest and posttest scores, which were collapsed across genders, was $r = .02$; and the nonsphericity correction factor was left at the default value of $\varepsilon = 1$, because the assumptions of homogeneity of variances and covariances in that ANOVA were met. The statistical power $(1 - \beta)$ for the test of the Gender $\times$ Time interaction effect was estimated by the G*Power analysis to be $0.56$. Thus, with the sample size available in this study, the likelihood was $56\%$ that a medium-sized Gender $\times$ Time interaction effect in the population would be detected as a statistically significant interaction effect in the sample ANOVA—not much better odds than the toss of a coin. Given that the Type II error rate was $\beta$ and the statistical power was $1 - \beta$, it follows that there was a $44\%$ chance that the nonsignificant Gender $\times$ Time interaction effect observed in this study was caused by a Type II error. It should be noted that the estimate of statistical power from G*Power is undoubtedly overly generous and that the estimated Type II error probability is an underestimate of the real value. G*Power assumes that the total sample size, which is specified, has been evenly divided between the groups in a study (the tiers in this study). If that is not the case, then the actual available statistical power $(1 - \beta)$ decreases, and the probability of a Type I error $(\beta)$ increases. It is fair to say that the test of the Gender $\times$ Time interaction effect in this study was insufficiently supported by the available sample size.
Summary of Results

This study used an ex post facto, retrospective, pretest-posttest research design to evaluate improvements in reading achievement among kindergarten students who were exposed to two tiers of an RTI model. In addition to evaluating the effectiveness of each RTI tier, the study examined if the two tiers were differentially effective in improving reading achievement. Finally, the study sought to determine if the effectiveness of the Tier 2 intervention was moderated by student gender. Reading Achievement, the dependent variable, was measured at pretest and posttest using the STAR-ELA.

**Major finding 1.** The first major finding is related to Research Question 1: To what extent does the implementation of Response to Intervention (RTI) impact nonclassified kindergarten students’ reading ability as measured by scores on the STAR Early Literacy Assessment (STAR-ELA)?

On average, the students in this study improved 159.55 points on the STAR-ELA from pretest to posttest. This was a large and statistically significant increase. This pretest-to-posttest improvement was significant both in Tier 1 (149.29 points) and Tier 2 (215.30 points). These results are consistent with the conclusion that the Tier 1 and Tier 2 RTI evaluated in this study were effective in enhancing students’ reading achievement from the pretest to posttest. In terms of the null and alternative hypotheses associated with Research Question 1, the null hypothesis—that the interventions exerted no significant effect on reading achievement—was rejected; and the alternative hypothesis—that the interventions exerted a significant effect on reading achievement—was accepted.
Major finding 2. The second major finding is related to Research Question 2: Is there a relationship between improvement in reading score on the STAR ELA and the types of RTI provided to nonclassified struggling kindergarten readers?

Although both Tier 1 and Tier 2 students demonstrated significant improvements in reading achievement from pretest to posttest, the Tier 2 students showed significantly greater improvement than the Tier 1 students. At the fall pretest, the Tier 2 students scored 99.98 points behind the Tier 1 students, which is a statistically significant between-group difference. However, the extra small-group instruction received by the Tier 2 students eliminated most of this disparity by the time of the spring posttest. By spring, the gap between the Tier 1 and Tier 2 students had declined to only 33.97 points, which is not a statistically significant difference. Expressed in terms of the null and alternative hypotheses associated with Research Question 2, the null hypothesis—that types of RTI intervention were unrelated to improvements in reading achievement—was rejected; and the alternative hypothesis—that types of RTI intervention were related to improvements in reading achievement—was accepted.

Major finding 3. The third major finding is related to Research Question 3: To what extent, if any, does gender moderate the impact of a Tier 2 RTI framework have on the reading achievement of nonclassified struggling kindergarten readers as measured by scores on the STAR-ELA?

The data examined in this study provided no support for the conclusion that males and females benefitted differently from the Tier 2 reading intervention. The males and females in Tier 2 showed approximately equal and statistically significant gains in STAR-ELA scores from pretest to posttest. On average, the males gained 247.14 points
and the females gained 221.87 points. The small samples available to support the analysis (14 males and 23 females) call for a cautious interpretation of the findings of the analysis of Research Question 3. The statistical power available to support the test of the ANOVA interaction effect that was the key analysis for Research Question 1 was less than 56%. Consequently, the probability that the nonsignificant interaction effect was due to a Type II error was relatively high—over 44%.
Chapter 5: Discussion

This chapter reviews the implications of the findings, limitations of the study, and recommendations for future studies. The purpose of this research was to examine the impact a multitiered RTI framework has on nonclassified kindergarten students’ reading scores as measured by the STAR-ELA. The researcher sought to explore if there was a relationship between the type of RTI provided to kindergarten students and reading scores. In addition, the researcher examined if gender impacted the Tier 2 intervention and reading scores. Studies indicate there are gender differences in reading (Below et al., 2010; Klecker, 2006) and that females outperform males in areas of reading as early as kindergarten (Selafani & Dennis, 2017).

Introduction

This quasi-experimental, ex post facto, retrospective study examined RTI and the impact a three-tiered RTI model has on nonclassified kindergarten student reading scores. The literature review in this study indicates that over 10 million students struggle with reading before third grade (NCES, 2015). Students who do not learn to read proficiently by the third grade have difficulty reading during their entire academic career (Stahl, 2016). Reading serves as a basis for learning, and a deficit in reading can cause negative effects on employment and lead to an increase in high school dropout rates (Brynner, 2008; Wanzek et al., 2014). Research states that preventing reading failure is more beneficial than correcting reading failure (Hunter et al., 2015; Otaiba et al., 2014). Early identification of students who have difficulty reading is critical in solving this problem.
Research indicates that kindergarten is a vital time to begin screening and providing intervention to students who may be at risk of reading failure (Catts et al., 2016). Screening, also, in kindergarten can produce acceptable levels of accuracy (Schatschneider et al., 2004). Furthermore, screening children at the beginning of kindergarten can accurately identify children who may have reading difficulties (Catts et al., 2016). Kindergarten intervention measures are shown to be predictive of third-grade reading comprehension (Catts et al., 2016; Schatschneider et al., 2004). RTI is a framework that can prevent reading failure in children as early as kindergarten. However, many factors influence the effectiveness of RTI (Balu et al., 2015). According to Sharp et al. (2016), there are essential elements for the effective implementation of RTI, which include: (a) high-quality core curriculum, (b) universal screening, (c) increasingly intensive tiers, (d) progress monitoring, and (e) consistency of implementation. When RTI is implemented correctly, positive outcomes occur.

Second, utilizing pretests and posttests, this study was used to analyze the difference in reading scores between male and female students. For over 100 years, researchers have been examining the differences between males and females in terms of reading achievement (Ayers, 1909). One of the first large studies conducted on gender differences in reading was in Iowa in 1942, and it showed that girls outperformed boys in both elementary and high school reading comprehension (Stroud & Lindquist, 1942). Furthermore, the National Assessment of Educational Progress (NAEP) found that females outscored males on all reading assessments at every grade-level test (Hansen, Levesque, Valant, & Quintero, 2018). Gender differences are not limited to the United States. In 2011, a study conducted by the Program in International Reading Literacy
Study (PIRLS) indicates that females outscored males consistently in reading comprehension in over 49 nations (Hansen et al., 2018). Husain and Millimet (2007) found that “boys lag behind girls at the start of kindergarten and at the end of third grade” (p. 2). These studies indicate the need to further explore the gender-gap difference in the primary grades.

The purpose of this quantitative study was to examine the impact a three-tiered RTI model had on nonclassified kindergarten students’ reading scores. Furthermore, this study investigated if there was a difference in gender and reading scores. This study used an ex post facto design with a retrospective cohort. The researcher examined if the independent variables (RTI Tier and Time) had an impact or effect on the dependent variable (Reading Achievement Scores), by examining pretest and posttest scores. Therefore, this study examined the following questions:

1. To what extent does the implementation of Response to Intervention (RTI) impact nonclassified kindergarten students’ reading ability as measured by scores on the STAR Early Literacy Assessment (STAR-ELA)?

2. Is there a relationship between improvement in reading scores on the STAR-ELA and the types of RTI provided to nonclassified struggling kindergarten readers?

3. To what extent, if any, does gender moderate the impact of Tier 2 RTI on reading scores of nonclassified struggling kindergarten readers as measured by scores on the STAR-ELA?
Implications of Findings

First, this study examined the impact of RTI on nonclassified kindergarten students’ reading ability as measured by the STAR-ELA. The students in this study significantly improved. On average, both male and female students improve by 159.55 points from the pretest to posttest. Students in Tier 1 improved 149.29 points, and the Tier 2 students improved 215.30 points. These results are consistent with prior research that states RTI is an effective framework. Noltemeyer and Sansosti (2012) examined the relationship between student reading performance and the quality of the RTI. The results indicate that when all the components of RTI are implemented, student performance improves. Sharp et al. (2016) found similar results and further added that school districts should analyze the implementation of their RTI model.

Particularly, this study showed that RTI had a significant positive effect on nonclassified kindergarten students’ reading scores. Furthermore, this study incorporated all the elements for effective RTI implementation, which include: (a) STAR-ELA as the universal screener, (b) the Wonders Reading Program as the high-quality research-based core curriculum, (c) data-based decision making, (d) a three-tiered model with increasing levels of intensity, (e) progress monitoring using the Wonders Reading Program, and (f) consistency of implementation. All these elements can be easily duplicated by school districts looking for an effective RTI model.

In addition, this study indicates that students identified as at risk in the beginning of kindergarten and who are provided with strategic interventions can be prevented from reading failure. Specifically, the Wonders Reading Program was implemented district wide in Grades K-5 in the subject school. Wonders incorporated Tier1 and Tier 2
interventions including small-group instruction, progress monitoring, direct instruction, differentiated instruction, and guided reading. This notion is consistent with the literature that states that the early grades can provide a unique opportunity to implement RTI instruction in which reading failure can be prevented (O’Connor et al., 2005). The literature review, as well as the results of this study, indicates the positive effect of targeted interventions in kindergarten (Cunningham & Stanovich 1997; Juel, 1988; O’Connor et al., 2005; Sparks et al., 2014).

RTI can identify student-learning issues early so that school districts can intervene prior to a large academic gap occurs. Most states follow an RTI framework; however, within each state, schools have the flexibility to use different research-based methods to identify struggling readers (Gersten et al., 2009). Within school districts, there are many unanswered questions about the fidelity of RTI implementation (Bradshaw et al., 2008). Furthermore, the RTI framework presented in this study can be an exemplar model for other school districts to assess the effectiveness of their RTI model. As shown in this study, kindergarten students’ reading achievement increased with early intervention, and the gap between at-risk and typical students decreased significantly. States and school districts should contemplate creating a common research-based framework (McInerney & Elledge, 2013). First, school districts should incorporate a three-tiered RTI framework with increasing levels of intervention. The RTI framework should include the following core components: (a) standardized universal screener, (b) high-quality core curriculum, (c) data-based decision making, (d) increasingly intensive tiers, (e) progress monitoring, and (f) consistency of implementation (Otaiba et al., 2014).
A standardized universal screener can determine which students may need additional intervention with reading. It is recommended that school districts use the STAR-ELA as the universal screener. The STAR-ELA can be administered multiple times throughout the year, and it can help to identify students who may need additional intervention. This study analyzed the data for fall and spring; however, it is recommended that districts administer the STAR-ELA three times each year to further identify students who may need Tier 2 or Tier 3 interventions. This additional assessment can also serve as a progress-monitoring tool. The Wonders Reading Program should be implemented as a school district’s high-quality core curriculum (Tier 1). Wonders ensures that all students are exposed to an effective research-based curriculum and differentiated Tier 2 interventions. In addition, if Wonders is implemented district wide, there will be consistency within and across-grade levels. This RTI framework can be implemented state wide. The state of California adopted the Wonders Reading Program, which was created on the California Framework and the California State Standards. Implementing a common RTI framework is imperative because 10 million children in the United States are at risk of reading failure before third grade (NCES, 2015). Implementing this framework can reduce the number of students needing intervention as early as kindergarten.

Second, this study examined the relationship between improvement in reading scores on the STAR-ELA and the types of RTI provided to nonclassified struggling kindergarten readers. This study found that both Tier 1 and Tier 2 students demonstrated significant improvements in reading achievement from the pretest to posttest. However, the Tier 2 students showed significantly greater improvement than the Tier 1 students. At
the fall pretest, Tier 2 students scored 99.98 points behind the Tier 1 students, a statistically significant between-group difference. By spring, the gap between the Tier 1 and Tier 2 students had declined to only 33.97 points, which is not a statistically significant difference. There are several reasons why this occurred. First, the small-group additional intervention, based on direct instruction, aided the students with their reading. Second, the fidelity of the intervention was consistent. In addition, the explicit instruction provided by the Wonders Reading Program assisted the students with increasing their reading scores. These findings are consistent with the research that states that small group size, as well as frequency and length of the intervention, are important factors in effective Tier 2 instruction. In this study, small groups of students were pulled out of the classroom three to five times a week and given interventions.

Third, this study examined if gender had an impact on the reading scores of nonclassified struggling kindergarten readers. McGlinn (2003) found that a gender gap existed in reading. Research as far back as 1909 states that males enter school with a deficit in reading (Ayers, 1909). In addition, preschool girls have had higher levels of emergent literacy skills than preschool boys (Justice at al., 2005). According to Chatterji (2006), gender differences exist between male and female children when entering school. The research indicates that females have been outperforming males on reading assessments for decades (Hansen et al., 2018). However, in this study, the males and females in Tier 2 showed approximately equal and statistically significant gains in STAR-ELA scores from pretest to posttest. On average, the males gained 247.14 points and females gained 221.87 points. This could be due to the small sample of students
analyzed in Tier 2 (14 males and 23 females). Therefore, caution should be taken when interpreting the findings of this analysis.

This study can benefit suburban school districts in implementing an RTI framework for the improvement in reading achievement for at-risk kindergarten students. Closing the achievement gap prior to the third grade is imperative in terms of student reading achievement (Fuchs et al., 2012). The results of this study indicate that RTI in kindergarten can significantly close the achievement gap between at-risk and typical kindergarten students. The data from this study can be used to guide a school district in the effectiveness of an RTI model on student reading scores, as well as add to the body of research involving RTI models and kindergarten reading scores.

In addition, school districts can significantly reduce their budgets by applying the framework used in this study. Implementing interventions early in kindergarten, and closing the reading achievement gap, can result in less remedial and intervention staff having to be utilized in upper grades throughout the district. Allocating money and resources in the primary grades will build a solid reading foundation for students. Although research suggests that gender does impact reading achievement (Chatterji, 2006; Fryer & Levitt, 2009; Husain & Millimet, 2009; Robinson & Lubinski, 2011) in this study, there was no difference between males and females in terms of levels of remediation as measured by reading scores on the STAR-ELA.

Furthermore, the McGraw-Hill Educational Company, creator of the Wonders Reading Program, would be captivated by the findings of this research study. This was an independent research study that indicates the effectiveness of the Wonders Reading Program in kindergarten as a core curriculum in addition to an effective Tier 2
intervention. Kindergarten students were able to generalize strategies learned in small group instruction to a standardized assessment.

**Limitations**

The purpose of this study was to explore the impact a three-tiered RTI model had on nonclassified kindergarten students’ reading achievement as indicated on the STAR-ELA. There are several possible limitations to this research study. It was conducted at one elementary school within one suburban school district with a homogenous socioeconomic status and ethnicity. The results of this study may not be generalizable to an urban school district with a diverse population.

A second limitation of this study is the small sample size that was utilized specifically in the Tier 3 intervention. In addition, research has indicated that there is a gap in the literature within Tier 3 interventions (Austin, Vaughn, & McClelland, 2017). Increasing the number of students who received Tier 3 interventions in this study would have allowed the researcher to include this tier within the study as well as allow for a more robust analysis.

**Recommendations**

There are several recommendations that can be made regarding future studies. First, this study was conducted with one cohort, in one suburban school district with a homogenous population, during one school year. Although students improved in both the Tier 1 and Tier 2 interventions, it would be interesting to conduct a longitudinal study with this cohort and analyze the students’ reading ability in third grade to see if the students who received intervention in kindergarten needed any further intervention in reading. Conducting a longitudinal study that expands this study over a 4-year period
would provide even more data and information on the effectiveness of this RTI model. District leaders may wish to consider a larger study that includes more grade levels. In addition, further research with a larger sample is needed before any solid conclusions can be drawn about whether or not gender moderates the impact of a Tier 2 reading intervention on the reading achievement of struggling kindergarten readers. Also, further research should be conducted on intervention in Tier 3. In this study, the sample size within Tier 3 was too small to analyze.

In addition, this study did not include classified students. Further studies should include students who are classified with an individualized education plan. According to O’Connor et al. (2005), many classified students have not been exposed to core instruction and are placed in Tier 2 interventions due to poor instruction. In addition, O’Connor et al. (2005) revealed that the incidence of placement into special education prior to their study averaged 15%, and 4 years following participation in the study, placement into special education decreased by 7%. Including this population into a study would give researchers and schools more information about targeted interventions with special education students. Similarly, including students who are English language learners could reveal how interventions can be effective with this population.

An additional topic of interest that might be important to explore is the possible connection of incarceration rates and low literacy skills. Studies indicate that if students are not reading at grade level by third grade, they will continue to remain behind their typical peers (Fuchs & Fuchs, 2009; Juel, 1988; Vellutino et al., 2008; Wanzek et al., 2014). In addition, between 2006 and 2007, an estimated 1 in 10 male high school
dropouts was incarcerated (Khatiwada, McLaughlin, Palma, & Sum, 2009). Furthermore, according to Harlow (2003) 40% of prison inmates do not have a high school diploma.

A 19-year study conducted by Gamier, Stein, and Jacobs (1997) found that dropping out of high school is a “multiply determined process, with early influences beginning in childhood . . . such as child ability prior to school entry, and early school performance” (p. 395). This point is further exacerbated by the findings that “early school difficulties may be associated not only with continued academic difficulties but also emotional problems or difficulties with interpersonal relationships that continue and are related to more stressful experiences in adolescence” (Gamier et al., 1997, p. 414).

Moreover, students who arrive at kindergarten with high levels of readiness are less prone to involvement in the criminal justice system and to quit high school (Shonkoff & Phillips, 2001). Closing the achievement gap by third grade might alleviate the number of people who become incarcerated.

**Conclusion**

Reading failure is a significant problem for students across the United States. Many students in kindergarten through third grade are struggling considerably with reading, which, if not remediated, can lead to harmful effects on future school performance (Lam & McMaster, 2014). Furthermore, poor reading skills in adolescence and adulthood have been found to have negative effects on employment and to contribute to social segregation (Brynner, 2008). Many researchers believe that early intervention can be an effective tactic to reducing the number of children who are at risk of reading failure (Denton et al., 2006b; Fuchs et al., 2014; Sparks et al., 2014). RTI is an early-
intervention framework that can assist in the prevention of reading failure (Otaiba et al., 2014).

RTI is a multtiered framework that emphasizes research-based instruction with increasing tiers of intensity based on students’ instructional need (Clarke et al., 2011). In a multtiered framework, instruction is intensified for struggling readers who are not showing growth with less intensive instruction (Turse & Albrecht, 2015). Tier 1 is designed to deliver high quality, research-based instruction to all students (Burns & Gibbons, 2012). Tier 2 is designed to provide additional intensive supports for students who do not make progress within Tier 1 (Fuchs et al., 2014). Tier 3 is the most intensive tier of support, and it is reserved for students who fall significantly below their peers (Denton et al., 2006b). The purpose of RTI is to provide additional supplementary support to students who are at risk of reading failure (Stahl, 2016).

The purpose of this quantitative, quasi-experimental ex post facto study was to examine the impact RTI has on kindergarten reading scores as measured by the standardized test for the achievement of reading early literacy assessment (STAR-ELA). A baseline universal screener, STAR-ELA, was used as pretest and posttest scores. Nonclassified kindergarten students received Tier 1 intervention (Wonders Reading Program) from their classroom teacher. Support staff implemented small-group instruction to students in Tier 2 and Tier 3, using the Wonders Reading Program. In addition, this study investigated the relationship between the dependent variable (Reading Scores) and the independent variables (Tier and Time). Lastly, this study examined the potential role of gender in moderating the effect of Tier 2 interventions on reading achievement. This study investigated the following research questions:
1. To what extent does the implementation of Response to Intervention (RTI) impact nonclassified kindergarten students’ reading ability as measured by scores on the STAR Early Literacy Assessment (STAR-ELA)?

2. Is there a relationship between improvement in reading scores on the STAR-ELA and the types of RTI provided to nonclassified struggling kindergarten readers?

3. To what extent, if any, does gender moderate the impact of Tier 2 RTI on reading scores of nonclassified struggling kindergarten readers as measured by scores on the STAR-ELA?

In this study, kindergarten students improved in reading scores by 159.55 points from pretest to posttest, which is a large, statistically significant improvement. Furthermore, students who were at risk for reading failure and were exposed to the RTI framework showed significant improvements in reading scores. Both students in Tier 1 and Tier 2 showed significant improvement from the fall pretest to the spring posttest. Tier 1 students improved an average of 132.59 points, and Tier 2 students improved an average of 239.39 points. These findings were consistent with previous studies that indicate that targeted early intervention within an RTI framework is effective in increasing reading scores (Catts et al., 2016; Fuchs & Fuchs, 2006; Fuchs et al., 2014; Lam & McMaster, 2014; Otaiba et al., 2014; Sparks et al., 2014). Scanlon et al. (2005) found that children at risk for reading difficulties should be identified at the beginning of kindergarten, and they should be provided with strategic interventions in order to prevent reading failure. In addition, 84% of students who received intervention at the beginning of kindergarten met grade-level expectations at the end of third-grade (Scanlon et al.,
Furthermore, Scanlon et al. (2005) found that the use of a universal screening at the beginning of kindergarten significantly reduced the number of children who were at risk of reading failure.

Furthermore, this study found that both Tier 1 and Tier 2 students improved significantly from pretest to posttest, and the improvement shown by Tier 2 students was significantly greater than that seen among Tier 1 students. This finding is consistent with the importance of a research-based high-quality core curriculum such as the Wonders Reading Program in Tier 1. In addition, the findings of this study suggest that RTI is imperative to student achievement. The advantage for students, who are at risk for reading failure to participate in early intervention is critical in closing the achievement gap prior to third grade. Children who are at risk for reading failure should be identified at the beginning of kindergarten with the use of a universal screener, such as the STAR-ELA, and they should be provided with research-based Tier 1 and Tier 2 interventions to prevent reading failure and close the achievement gap.

According to previous studies, males score lower than females on reading assessments (Mullis, Martin, Gonzalez, & Kennedy, 2003; Grigg, Daane, Jin, & Campbell, 2002). One finding in this study that was not consistent with previous research is that females outperform males on reading assessments. In this study, males and females in Tier 2 showed approximately equal and statistically significant gains in STAR-ELA scores from their pretest to posttest. On average, males gained 247.14 points and females gained 221.87 points. The small samples available to support the analysis (14 males and 23 females) might have had an impact on the results. Subsequently, the probability that the nonsignificant interaction effect was due to a Type II error was
relatively high—over 44%. Further studies with larger samples are needed in order to make a conclusive decision regarding the role of gender in moderating the effectiveness of the Tier 2 reading intervention.

The results of this study indicate that early intervention can close the reading achievement gap of nonclassified kindergarten students’ reading scores as measured by the standardized test for the STAR-ELA. In addition, implementing an RTI framework that is based on increasing tiers, student data, progress monitoring, small group size, and a research-based curriculum can improve reading scores and close the reading achievement gap.
References


Bornstein, J. (2015). If they’re on Tier 1, there are really no concerns that we can see: PBIS mediatizes compliant behavior. *Journal of Ethnographic & Qualitative Research, 9*(4), 247-267.


Ellis, E., Larkin, M., & Worthington, L. (n.d.). *Executive summary of the research synthesis on effective teaching principles and the design of quality tools for educators.* University of Alabama, Birmingham, AL.


Individuals with Disabilities Education Act of 2004, P.L. 108-446


Reynolds, M., Wheldall, K., & Madelaine, A. (2011). What recent reviews tell is about the efficacy of reading interventions for struggling readers in the early years of


Appendix A

RTI Model Implemented by School District

![Flowchart of RTI Model](image-url)
Appendix B

Three-Tiered RTI Model

Tier III – Intensive Intervention
- Individualized plans for at-risk students
- Frequently monitor effects of intervention

Tier II – Small Group Instruction
- Targeted intervention based on screening results
- Increased level of service from support teachers

Tier I – Whole Group Instruction
- All students receive Tier I support in class from the classroom teacher
- Students are screened on a periodic basis
Appendix C

Data Request

September 28, 2017

Dr. Walter Moran
Superintendent
Eastchester Union Free School District
580 White Plains Road
Eastchester NY 10709

Dear Dr. Moran:

As a doctoral student in the Executive Leadership Program at St. John Fisher College I am conducting research to examine the impact Response to Intervention has on non-classified kindergarten student reading scores as achieved by the STAR (standardized test for the assessment of reading) Early Literacy Assessment. The purpose of my research study is to examine to what extent response to intervention impacts kindergarten reading scores and if there is a relationship between the type of intervention and improvement in reading scores. The title of my research study is: Preventing Reading Failure: The impact of a response to intervention framework on non-classified kindergarten student's reading scores.

I am writing to request your permission to utilize data from the kindergarten STAR Early Literacy Assessment during the 2016-2017 school year. For the purposes of the study, I am requesting fall (September) and spring (May) data. Student data from self-contained special education classes will not be included in the sample collection. To protect the identity of the students identifying information (the student’s name) will be removed prior to the researcher obtaining the data. Identification numbers will be assigned to each student and the numbers will be matched on the pretest and posttest data. Physical assessment data will be kept in a locked file box stored securely in the private home of the researcher. In addition, online data will be kept in a password-protected file until analyzed by the researcher.

Thank you for considering my request. As this is archival data, and the researcher is not interacting with students I will not need individual permission from each student. However, in order to obtain approval for the St. John Fisher Institutional Review Board I will need approval from the Eastchester School District to use the data in my research study.

I am requesting a signed statement on official school letterhead indicating your approval of the use of the STAR data in my research study. Please respond by email to victorabursella@gmail.com

Once I get approval from St. John Fisher College IRB I will send a copy of the approval to you, in order to obtain the data.

Sincerely,
Victoria Bursella
St. John Fisher College Doctoral Student
Appendix D

Consent for Data

EASTCHESTER UNION FREE SCHOOL DISTRICT

Walter R. Moran III, Ed.D.
Superintendent of Schools

September 28, 2017

To whom it may concern:

I hereby grant approval of the use of data from the 2016-2017 school year STAR Early Literacy Assessment of kindergarten students in the Eastchester Union Free School District for a doctoral research study by Ms. Victoria Borsella at St. John Fisher College. I have been fully informed of the care in which this data will be utilized and protected.

I extend my best wishes to Ms. Borsella as she moved forward with her doctoral dissertation.

Sincerely,

Walter R. Moran III

Dr. Walter R. Moran III
Appendix E

Default Benchmark Scores

<table>
<thead>
<tr>
<th>Grade</th>
<th>Percentile</th>
<th>Scaled Score</th>
<th>Est. ORF&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Scaled Score</th>
<th>Est. ORF&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Scaled Score</th>
<th>Est. ORF&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>10</td>
<td>399</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>437</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>452</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>496</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>522</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>582</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>647</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>b</sup> Est. ORF: Estimated Oral Reading Fluency
Appendix F

Items Assessed in the STAR ELA


2. Concept of Word (CW) – Understanding of print concepts regarding written word length and word borders and the difference between words and letters.

3. Visual Discrimination (VS) – Differentiating both upper- and lowercase letters, identifying words that are different and matching words that are the same.

4. Phonemic Awareness (PA) – Understanding of rhyming words, ability to blend and segment word parts and phonemes, isolating and manipulating initial, final, and medial phonemes and identifying the sounds in consonant blend.

5. Phonics (PH) – Understanding of short, long, variant vowels and other vowel sounds, initial and final consonants, consonant blends and digraphs, consonant and vowel substitution and identification of rhyming words and sounds in word families.


7. Vocabulary (VO) – Knowledge of high-frequency words, regular and irregular sight words, multi-meaning words and words used to describe categorical relationships, position words, synonyms and antonyms (Renaissance Learning, 2009, p.28).