Southern Methodist University
SMU Scholar

Department of Teaching and Learning Theses and Dissertations

Summer 2020

Exploring Response to Intervention: How Are We Narrowing the Gap for Tier 3 Minority Learners?

KRISTI BAKER
Southern Methodist University, klbaker@smu.edu

Follow this and additional works at: https://scholar.smu.edu/simmons_dtl_etds

Part of the Curriculum and Instruction Commons, and the Disability and Equity in Education Commons

Recommended Citation
BAKER, KRISTI, "Exploring Response to Intervention: How Are We Narrowing the Gap for Tier 3 Minority Learners?" (2020). Department of Teaching and Learning Theses and Dissertations. 8.
https://scholar.smu.edu/simmons_dtl_etds/8

This Dissertation is brought to you for free and open access by the Department of Teaching and Learning at SMU Scholar. It has been accepted for inclusion in Department of Teaching and Learning Theses and Dissertations by an authorized administrator of SMU Scholar. For more information, please visit http://digitalrepository.smu.edu.
EXPLORING RESPONSE TO INTERVENTION:
HOw ARE WE NARROWING THE GAP FOR TIER 3 MINORITY LEARNERS?

Approved by:

_______________________________________
Dr. Stephanie Al Otaiba
Professor and Centennial Chair, Teaching
And Learning

_______________________________________
Dr. Paul Yovanoff
Professor of Teaching and Learning

_______________________________________
Dr. Francesca Jones
Clinical Associate Professor; Director of
Graduate Studies and Engaged Learning

_______________________________________
Dr. Kristen McMaster
Professor of Educational Psychology
NARROWING THE GAP FOR MINORITY LEARNERS

EXPLORING RESPONSE TO INTERVENTION:
HOW ARE WE NARROWING THE GAP FOR TIER 3 MINORITY LEARNERS?

A Dissertation Presented in Partial Fulfillment of the Requirements
for the degree of Doctor of Philosophy

by

Kristi Baker
Southern Methodist University

June, 2020
ACKNOWLEDGMENTS

This work could not have been accomplished without the support of my amazingly wonderful advisor, Dr. Al Otaiba, the support of Dr. Yovanoff, and so many faculty on whose expertise I have come to rely in the Department of Teaching and Learning here at SMU. Another thank-you to Dr. Jones who encouraged me to consider the Ph.D. program. I further acknowledge the support of the National Center on Leadership in Intensive Intervention (NCLII) through which I have been afforded so many opportunities and have been introduced to be most brilliant minds in special education. Thank you to my PhD cohort who have become family. Finally, I am so thankful to my family and friends who have continued to encourage me throughout this journey. Not enough words can express my gratitude. A special thanks to my Ken who checked on me during the late nights, prepared my food when I forgot to eat, and reminded me I was great when I just wanted to cry. For every act of kindness, thank you.
Abstract

The purpose of this mixed-methods study was to 1) analyze the relationship between RTI tier of intervention, minority status, and reading growth and 2) explore school level factors that might explain differential growth of minority students in Tier 3 relative to their peers. I used the lens of Critical Race Theory in Education to explore the reading growth of minority students in Tier 3 relative to their peers which included non-minority students in Tier 3 as well as minority and non-minority students in Tier 1. I sampled a total of 1,002 minority and non-minority elementary students from three schools in one largely minority school district. I first conducted regression analyses to correlate tier of intervention and minority status with student reading growth. I then conducted a more in depth case study of the three schools to explore if there were plausible explanations of differential patterns of growth related to (1) articulation of data-based decision making for eligibility for Tier 3, (2) proportion of students in special education are also served in Tier 3, and (3) intensity of intervention in Tier 3. I found that two of the schools demonstrated patterns of more growth for minority students in Tier 3, whereas one school demonstrated a pattern in which minority students grew less in relation to their peers. These two schools may represent intentional efforts to narrow the reading gap for minority students with reading difficulties.

Key words: response to intervention, minority students, Critical Race Theory, reading growth
# TABLE OF CONTENTS

LIST OF FIGURES ................................................................................................................. viii

LIST OF TABLES ..................................................................................................................... ix

CHAPTER 1: Introduction ......................................................................................................... 1

1.1 Critical Race Theory ...................................................................................................... 2

1.2 Response to Intervention ............................................................................................ 6

1.3 Relevant Literature ..................................................................................................... 14

1.4 Research Questions ..................................................................................................... 20

1.5 Key Definitions ........................................................................................................... 21

CHAPTER 2: METHODS ........................................................................................................ 24

2.1 Context of Larger Study ............................................................................................. 24

2.2 School Demographics and Student Participants ......................................................... 25

2.3 Measures and Data Sources ........................................................................................ 28

2.4 Data Analytic Methods ............................................................................................... 33

CHAPTER 3: Results ............................................................................................................... 39

3.1 Research Question 1 ................................................................................................... 39

3.2 Research Question 2 ................................................................................................... 39

3.3 Research Question 3 ................................................................................................... 40

CHAPTER 4: Discussion .......................................................................................................... 53
4.1 Research Question 1 ................................................................................................... 53
4.2 Research Question 2 ................................................................................................... 54
4.3 Research Question 3 ................................................................................................... 54
4.4 Strengths and Limitations of Study ........................................................................... 57
4.5 Conclusion .................................................................................................................. 59
REFERENCES ......................................................................................................................... 62
APPENDICES .......................................................................................................................... 74
APPENDIX 1- LIST OF FIGURES

Figure 1  Prisma Diagram of Literature Review ................................................................. 74
Figure 2  Sample Selection Tree ...................................................................................... 75
Figure 3  MAP Growth Norms .................................................................................. 76
Figure 4  Excerpt of RTI Implementation Worksheet ..................................................... 77
Figure 5  Excerpt of ICE-R Template ............................................................................. 78
Figure 6  Observation Instruction Lower Elementary Results Tier 3 .............................. 79
Figure 7  Observation Grouping Lower Elementary Results Tier 3 ................................. 80
Figure 8  Observation Instruction Lower Elementary Results Tier 1 ............................... 81
Figure 9  Observation Grouping Lower Elementary Results Tier 1 ................................. 82
Figure 10 Observation Instruction Upper Elementary Results Tier 3 .............................. 83
Figure 11 Observation Grouping Upper Elementary Results Tier 3 ............................... 84
Figure 12 Observation Instruction Upper Elementary Results Tier 1 .............................. 85
Figure 13 Observation Grouping Upper Elementary Results Tier 1 ............................... 86
APPENDIX 2- LIST OF TABLES

Table 1  Summary of Studies Examining RTI ................................................................. 87
Table 2  Summary of Recommended Tier 3 Components Across Studies.......................... 95
Table 3  Summary of Recommended Tier 1 Components Across Studies.......................... 99
Table 4  Summary of Research Rigor, or Quality Indicators in Quantitative Studies.......... 101
Table 5  Number of Students in Each Tier by Minority Status ......................................... 103
Table 6  School Demographic Data ................................................................................. 104
Table 7  Results from Regression Models of Reading Growth ......................................... 105
Table 8  Mean Scores on RTI Implementation Rubric ...................................................... 106
Table 9  Number of Students with Disabilities................................................................... 107
This is dedicated to parents. I wish you were here to see.
CHAPTER 1
INTRODUCTION

Reading skills are critical for students throughout and beyond years of formal education. However, national data continues to show that students often read far below grade level, which limits their ability to understand grade level content. In its biannual report, the National Assessment of Educational Progress (NAEP) identifies four categories of reading: 1) advanced, 2) proficient, 3) at basic, and 4) below basic. According to the NAEP website, students who read at a basic level are able to “locate relevant information, make simple inferences…use their understanding of the text to identify details that support a given interpretation or conclusion [and] interpret the meaning of a word as it is used in the text” (NAEP, 2018). However, the NAEP data also show that by fourth grade, while about 65 percent of students on average read at basic or above, there is gap in reading performance between minority and non-minority students that continues throughout high school. Data indicate a much smaller percentage of students from minority backgrounds read at or above “Basic” compared to their non-minority counterparts (i.e. Black American = 48 percent; LatinX = 52 percent, and White = 78 percent; NAEP, 2018). For Black American and LatinX students, this represents a gap of 27 points and 24 points less growth, respectively; and this gap has remained largely unchanged for several decades. When considering vulnerable populations, those who qualify for free and reduced lunch programs (the majority of whom are also minority students), 50 percent of those students read at or above “Basic” compared to 79 percent for students who do not qualify. Finally, students with disabilities comprise 70 percent of the students who read below basic compared with only 29 percent of students without disabilities (National Center for Education Statistics; NCES, 2019).
The inability to read proficiently is a public crisis linked to systemic inequities, which lead to a host of undesired consequences including higher rates of dropout and limited future employment opportunities (Alliance for Excellent Education, 2011). The need to improve outcomes is deeply concerning given that policy reforms have been enacted at the federal level to support students’ reading growth and success. However, policy alone has failed to reveal substantial impacts leading to noticeable change in narrowing the gap in reading performance for vulnerable students, and for students with disabilities. Prior to reauthorization of the Individuals with Disabilities Act (IDEA), which was reauthorized in 2004, states generally used the IQ-Achievement Discrepancy Model to identify students with learning disabilities. This model required students to demonstrate a “severe discrepancy” between intelligence and overall achievement in a particular area. However, this model was problematic in that students from a variety of diverse and minority backgrounds had to wait too long and “fail enough” to receive special education, and a reason for this discrepancy could have been a lack of timely exposure to quality instruction (Fuchs, Fuchs, & Morgan, 2003). After IDEA-2004, states and local education agencies (LEAs) were no longer required to use the discrepancy model. With the reauthorizations, states and LEAs could choose instead to adopt a framework of tiered interventions to address quality of education and opportunity to learn to prevent difficulties through early intervention, and to provide timely intensive interventions and reduce unnecessary placements in special education, especially in learning disabilities. This would be particularly helpful for students who may appear to have a learning disability, but instead may demonstrate learning struggles resulting from lack of quality and access to needed learning experiences.

**Critical Race Theory in Education and Serving Students of Color**
As previously mentioned, it is particularly troubling that most students who read below proficiency are of minority backgrounds (NCES, 2019). It is of no surprise that students who are minority and who qualify for FRL have the largest gaps in achievement. Data indicates that Black American and LatinX students are more than twice as likely to be in poverty than White and Asian students (Asante-Muhammad, 2016; Kaiser Family Foundation, 2018). These gaps are well documented and have a long history in American education, however, there has not yet been a solution to remedy this phenomenon. Critical Race Theory (CRT) in Education (Ladson-Billings & Tate, 1995; Dixson & Rousseau, 2005) may help to explain the large discrepancies.

CRT was first used in the mid-1970s as a means to explore the role of race and racism on legal matters, with lawyers and activists desiring to gain traction with civil rights efforts which had started to wane (Delgado and Stefancic, 1993). Later, Ladson-Billings and Tate (1995) applied CRT to education (CRT-E) asserting that the characteristics of CRT are also mirrored within education. Since its introduction, there has been no singular definition of CRT and several scholars have proposed tenets to guide how CRT and CRT-E scholars define the work. However, Ladson-Billings (2005) cautioned the importance of ensuring that proposed definitions of CRT-E remain closely aligned with its legal roots. Dixon and Rousseau (2005, 2015) further explained that there is often overlap between how CRT looks in both legal and education sectors. Extending Ladson-Billings and Tate’s work, Dixon and Rousseau have identified broad concepts through which to analyze CRT-E. These include:

1) **Using voice.** Voice is the of narrative to describe and analyze the experiences of people of color and other marginalized groups.

2) **Moving from restrictive to expansive views of inequality.** An expansive view of systemic inequality and racism related to the outcomes experienced by students of
color and considers ways to ensure that all students are successful, regardless of race. A restrictive view of inequality only considers whether students of color are treated equally. As such, this view often leads to considering one group as inherently inferior compared to the majority group (Dixson & Rousseau, 2005).

3) **Rejecting color-blindness.** Color-blindness ignores that race is a social construct used to categorize people of different origins thereby making it impossible to combat the forms of racism presented in educational contexts. The construct of race has long been a factor in how people are treated and the opportunities afforded to them.

4) **Advocating for change.** CRT-E should not just identify the problems but also propose solutions that lead to better outcomes for students of color.

Ladson-Billings & Tate (1995) argued that: (a) embedded racism within society is demonstrated in that minority students often find the least success in public schools intended to provide educational for all; and (b) civil rights’ initiatives intended to end segregation, have unintentionally led to greater segregation in urban, public schools. The combination of these alarming outcomes have led to inequities in education within and across schools such that quality educational opportunities are often limited for students of color compared to their non-minority counterparts. Across schools, this is demonstrated through differences in educational access where large, urban public schools, with limited resources and funding, are most often populated by minority Black American and LatinX students. For example, recent studies have shown that gerrymandering- manipulating school boundaries- has led to increased segregation of minority students between districts, such that districts, rather than schools, include predominately minority or non-minority students (Richards, 2014, 2017). Within schools, this is demonstrated in differences in “tracking”, where racial and ethnic minority students are disproportionally
represented in academically less intensive educational tracks such as remedial classes and special education compared to their White counterparts, resulting in an imbalance of educational opportunity (Ladson-Billing & Tate, 1995; Oakes, 1995; Dixson & Rousseau, 2005). Additional research has shown that students from minority backgrounds are over-identified in receiving services in some areas of special education such as intellectual disabilities and behavioral disabilities (Gardner, Rizzi, & Council, 2014; Skiba, Poloni-Staudinger, Gallini, Simmons, & Feggins-Azziz, 2006), and are under-identified in other areas such as autism, dyslexia, and gifted and talented programs (Hosp & Reschly, 2004; Morgan et al., 2015; Robinson, 2016).

Considering these factors, educational successes for students, especially those from diverse backgrounds, may increase when schools and researchers plan for how to support students who have been traditionally underserved and ensure the quality of reading interventions that are provided to them. Potential supports may include: (a) high quality instruction and intervention provided to all students in Tier 1 that is consistent with the science of reading (NRP, 2000); (b) a clearly articulated plan for using data to immediately provide intensive intervention for students who may be “at risk” (Al Otaiba et al, 2014); (c) culturally relevant supports that account for the experiences minority students have and that may supplement mainstream instruction by strategically focusing on developing and promoting students’ academics, cultural competence, socio-political consciousness (Ladson-Billings, 1995); or (d) instructional strategies such as antecedent behavior strategies during instruction which have been shown to increase student engagement and decrease problem behaviors, and using data to monitor response and individualize supports, which ultimately leads to increased student achievement (Deno, 1998; Gunter & Denny, 1998; Gunter et al., 1993; Skinner & Belmont, 1993; Sutherland, Copeland, & Wehby, 2001). Through a more expansive view of inequality, these and other
supports may lead toward narrowing the gap in reading performance for students of color by accelerating learning and achievement. To that end, the combination of properly implemented structures of tiered support such as RTI, and intentionality to intensify interventions that are data-driven to meet the targeted needs for students in order to increase successful outcomes for students of color, may provide a promising and inclusive framework for all students.

Response to Intervention

Response to Intervention (RTI) for academics, particularly reading, has a long tradition. The broader use of “multi-tiered system of supports” includes not only academic, but social and emotional learning and behavioral interventions in which students who struggle receive increasingly intensified levels of intervention (Fuchs, Fuchs, and Compton, 2012). Gersten et al. (2009) conducted a review of the literature for the Institute for Education Sciences to describe the strength of evidence and provide recommendations for RTI implementation. RTI models vary, but generally follow a three-tier process that is typically portrayed as a pyramid divided into three parts; each representing a tier. On the bottom level of the pyramid (Tier 1), the vast majority of students (80 - 85 percent) are served and represent what is considered general education. In the middle level (Tier 2), a much smaller percentage of students (10 – 15 percent) who do not respond to well-implemented Tier 1 are identified to receive supplemental instruction. At the top of the pyramid (Tier 3), a small portion of students (five percent) would be expected to need support through intensive interventions (National Center on Intensive Intervention, n.d.).

Tier 1

Research has shown that most students respond to high quality, evidence-based, and explicit Tier 1 instruction. Encouragingly, this level of high-quality Tier 1 instruction is sufficient for students who may be initially considered “at-risk” for future reading difficulties;
however, it is also most effective in the early grades (Foorman et al., 1998; Snow, Burns, & Griffin, 1998; Jitendra et al., 2004; Denton, 2012).

Experts define high-quality Tier 1 as instruction that includes the five components of foundational reading (phonemic awareness, phonics, fluency, vocabulary, and comprehension) and allows students multiple opportunities to engage with written text (Ehri, 2004; National Reading Panel, 2000). These reading skills can be further subdivided into code-focused and meaning-focused skills (e.g., the Simple View of Reading (SVR); Gough & Tunmer, 1986). Proponents of SVR posit that proficient reading includes both code-focused skills, such as decoding and reading with fluency, and meaning-focused skills or the ability to comprehend while reading. Though the goal is often to ensure that these skills are developed together, instruction in primarily code-focused skills tends to focus on early elementary when children are learning to read, then tapers off by third grade. At this point reading with prosody and building comprehension skills takes greater precedence in reading instruction as children read to learn (Adams, 1990). In early elementary, we would expect interventions to be heavily focused on phonemic awareness, fluency, and decoding strategies; in fourth grade and beyond, we expect interventions to be heavily focused on vocabulary and comprehension (Adams, 1990).

Within the meaning-focused skills, evidence also suggests that language and vocabulary skills are heavily tied to reading achievement (Snow, Burns, & Griffin, 1998; Beck & McGowan, 2007). Thus, the not-so-simple views of reading theories incorporate other skills and competencies such as emerging bilingualism, language code-shifting, home language environments and culture, and behavioral and motivation (e.g., Li et al., 2020; Terry et al., 2012, Saenz et al., 2012). However, schools, teachers, and education policy may need to consider barriers of “typical” language acquisition that some diverse students may exhibit during
traditional language instruction. These would include instructional supports for students who speak dialects of the English language that differ from what is referred to as Mainstream American English (Oetting & McDonald, 2002; Edwards et al., 2014; Gatlin & Wanzek, 2015) or English Language Learners who need supports in both languages as they develop literacy (Ortiz et al., 2012). In sum, it is important to provide direct vocabulary instruction to support the unique needs of diverse learners that likely inhibit reading achievement. Therefore, we would expect reading instruction to include vocabulary development throughout all grades.

**Universal screening and progress monitoring.** Another component of RTI, particularly for Tier 1 core reading instruction, is the use of universal screening. Universal screening is assessment provided to all students to determine academic skills relative to a normed population of students who are in the same age/grade (Gersten et al., 2009). Gersten and colleagues recommend that the tools be standardized and implemented with fidelity, and at the time curriculum-based measures (CBM) were suggested. For decades, even before RTI was encouraged, schools have used curriculum-based measures (CBMs) which have been shown to be reliable screening assessments that are predictive of how students perform on annual state accountability assessments (Fuchs et al., 1994; Silberglitt, Burns, Madyun, & Lail, 2006). There is also a rich history of using CBMs for data-based individualization during progress monitoring (e.g., Espin et al., 2012; Jenkins & Fuchs, 2012) and data-based individualization is considered a high leverage practice (CEC HP6, 2017; Jung et al., 2018). However, those screeners often only reflect one subset of reading skills, typically fluency, rather than encompassing student’s collective reading abilities. Additionally, they can be time consuming as they often require individual administration.
Computer adaptive tests. In recent years, schools have started to rely on computer-adaptive measures (CATs) which can be group administered and are able to assess multiple areas of reading skills (Klingbeil et al., 2015). In these assessments, a series of questions are presented to the student based on how the student answered the previous question(s) (Shapiro, 2012). If the student answers incorrectly, an easier question is presented. If the student answers correctly, a more challenging question is presented. The questions continue to “adapt” based on student responses until the student completes the assessment. Different from pen and pencil assessments, computer adaptive tests allow for multiple students to be assessed simultaneously.

Computer adaptive tests are most relevant for my dissertation study, as it focuses on data from the Northwest Education Association Measures of Academic Proficiency or the MAP (Northwest Evaluation Association; NWEA, 2015). The MAP is one CAT that is widely used to screen students’ overall reading skills, including those identified by the NRP (2000). According to NWEA, the reading assessment was normed on a large sample of students in kindergarten through 11th grades and has demonstrated high reliability and validity. Upon completion of the assessment, students are provided a Rasch Unit (RIT) score which is scaled vertically and allows students to compare their individual growth across years and compare their growth to other students in their same age/grade range (NWEA, 2015).

Recent studies reveal the MAP assessment to be a reliable and accurate screener (Klingbeil et al, 2015; Vanderheyden, Burns, & Bonifay, 2018). In one study, using a sample of 500 second and third grade students, Klingbeil and colleagues assessed the accuracy of students’ scores on MAP and two additional screeners (i.e. oral reading fluency, ORF; and Fountas and Pinnell Benchmark Assessment System, BAS). They found the MAP to be the most accurate measure of the three at predicting reading performance on end of year assessment data which
was used to measure reading achievement. Later, Vanderheyden and colleagues conducted a study to assess the benefit of using multiple screeners given the cost and time associated continuously collecting screening data. Again, using three assessments (i.e. MAP, 2009; Developmental Reading Assessment–Second Edition, DRA-2, 2011; CBM), on a larger sample of 797 second and third grade students, the research team found the MAP to be the most accurate at predicting student reading outcomes on their end of the year state assessment. Given ease of administration and evidentiary support, many schools will likely continue to move toward using CATs such as the MAP to screen students for future reading difficulty.

To summarize, CBM and CAT may be used for universal screening purposes. Both can be adapted for progress monitoring.

**Tier 2**

When students demonstrate discrepancies between their peers and do not respond adequately to the general curriculum, or Tier 1, they are identified as “at-risk” for potential reading failure and should be provided Tier 2 interventions. Tier 2 is supplemental to Tier 1 instruction and provided three to five times per week in a small group for an additional 20 to 40 minutes per session (Gersten et al, 2008). Moreover, the intervention should be targeted in a student’s area(s) of need. Students are monitored in Tier 2 for a predetermined time frame to measure response to the intervention.

Gersten et al.’s (2008) review of the literature indicated a strong evidence base for the effect of Tier 2 intervention (with relatively more evidence at the time for Tier 1 or 3). A few systematic reviews have additionally provided insight into the effectiveness of Tier 2 interventions. In one review, Wanzek et al. (2016) conducted a meta-analysis of 72 studies providing Tier 2 reading interventions with students in kindergarten through third grade.
reporting standardized and unstandardized foundational reading and language-comprehension measures. Overall, they found higher effects in unstandardized language-comprehension measures ($ES = 1.03$) and lower effects on standardized language-comprehension measures ($ES = 0.38$). Further, they found moderate effects on standardized foundational reading ($ES = 0.49$) and unstandardized foundational reading ($ES = 0.62$). The researchers reported that the effects were not dependent on type of intervention, duration, group size, grade, or implementer.

In a later study, Gersten et al. (2017) systematically reviewed 27 studies examining the effect of 20 Tier 2 interventions from 2002-2014 provided to students in first through third grades. They found that of the 20 interventions, only one demonstrated no effects on student reading outcomes. The highest intervention effects were found in foundational reading (e.g., word reading and pseudo-word) reading ($ES = 0.456$) while interventions seemed to produce no impact on improving vocabulary. For comprehension, effects sizes ranged from 0.37 (in first grade) to 0.33 (in second and third grades). The researchers reported that all tier 2 intensified interventions were administered either small group (3-5 students) or individually.

Most of the reviewed research involved researcher-implementation. In one large study, Balu et al. (2015) conducted a regression discontinuity to evaluate the intensity of reading services in schools that had fully implemented RTI. Concerning findings indicated that rather than being effective, providing more intensive intervention produced negative effects at first grade and non-significant effects at second and third grades. However, in reviewing the study, Fuchs and Fuchs (2017) noted that while a majority of schools reported to have fully implemented RTI on their campuses, they did so in ways that were inconsistent with the recommendations for implementation by Gersten et al. (2009).

*Studies on Tier 3*
If, after a pre-determined time of progress monitoring, students continue to make inadequate growth with Tier 2 interventions, they are provided the most intensive and individualized Tier 3 interventions, and possibly referred to special education (Fuchs, Fuchs, & Compton, 2003). Gersten et al. (2009) reported that at the time, there was a limited evidence-base for Tier 3, when delivered to students who had not previously responded adequately to interventions in Tier 2. In their IES practice guide, Gersten and colleagues recommended that Tier 3 should include daily instruction that is more intensive than Tier 2 and individualized to student needs along with more frequent progress monitoring. These procedures may help reduce over and under-identification in special education through targeted and intensive support for students whose struggles stem from lack of access. However, research has identified a lack of specificity regarding what constitutes effective Tier 3 instruction and how it should look. In some states, this has led to a “blurring of special education” often leading to confusion and overlap between Tier 3 provided by general education staff and interventionists, and Tier 3 reserved for students who do not respond to Tier 2 and are then identified as needing special education (Fuchs, Fuchs, & Stecker, 2010).

More recently, researchers have begun evaluating Tier 3 interventions and their overall impacts for students through randomized control trials and literature syntheses. In one randomized control trial, Al Otaiba et al. (2014) analyzed the impacts of receiving RTI immediately. The authors screened first grade students to determine their level of risk and randomly assigned them to one of two treatment conditions. In the Dynamic condition, students at the greatest risk received Tier 3 immediately. In the Typical condition, students began Tier 1 and were not eligible for Tier 3 until they demonstrated inadequate Tier 2 response. Additionally, in Tier 1, all teachers received RTI training and screening data. Students received intense
instruction four days per week, 45 min per sessions focused on code-focused instruction in
phonics, phonemic awareness, fluency, alphabetsics, and morphology and on meaning-focused
instruction in listening and reading comprehension. They reported average overall effects on
reading ($ES = 0.36$) favoring the Dynamic group, such that students who were immediately
provided with the needed supports performed better on reading outcome measures than
comparison students who had to systematically progress through intervention supports.

In two reviews of Tier 3 interventions, one research team analyzed the effects for
students in kindergarten through third grade (Wanzek & Vaughn, 2007) and extended that study
to include students in fourth through 12th grades (Wanzek & Vaughn, 2013). Given the lack of
consistent or clear definitions across the studies in defining Tier 3, the researchers used the term
“intensive intervention” to mean Tier 3 and so studies were only reviewed that provided a
minimum specified number of sessions (100 sessions for early elementary, and 75 sessions for
fourth through 12th grades). Also, only standardized interventions (not individualized to student
need) were included.

The two studies demonstrated stronger effects for early intervention than upper
elementary. In the elementary study, Wanzek et al. (2007), the highest effects favored studies
focused on phonics (e.g. $ES = 0.91$) and text reading (e.g. $ES = 1.33$) and with participants in
kindergarten through first grade. The smallest effects were found for interventions provided in
small groups ($ES = 0.18$). No effects were found for standardized measures compared to
unstandardized measures. In the middle school study, the researchers found small but positive
mean effects on measures of comprehension ($ES = 0.10$), word reading ($ES = 0.15$), word
reading fluency ($ES = 0.16$), and reading fluency ($ES = 0.16$).
In a recent synthesis, Austin et al. (2017) conducted a systematic review of the effect of Tier 3 reading interventions in kindergarten through third grade for students who demonstrated insufficient or inadequate response to Tier 2 intervention. In this review, 12 studies met inclusion criteria. Overall, results demonstrated that students who did not respond to Tier 2 were able to make significant improvements with more intensive instruction through Tier 3 with high effects favoring inadequate responders who received intensive intervention compared to those who did not. However, effects were varied and ranged from -3.81 to 1.70 on measures of word identification, word attack, reading fluency, and, reading comprehension. Austin and colleagues also reported the need for further information in Tier 1 and finding varied definitions across the studies about what constituted inadequate response to Tier 2. They highlighted the need for further research that reported the observed characteristics of Tier 3 to extend knowledge about how Tier 3 is actually implemented for students with intensive reading needs.

In summary, many researchers have studied the effectiveness of reading interventions provided within RTI. Although we know what effective instruction should be included at Tier 1, there is far less research about whether or not students actually received quality Tier 1 or core reading instruction prior to needing more intensive levels of intervention. The strongest evidence of effectiveness has focused on supplemental and standardized Tier 2 interventions. Available research on Tier 3 intensive interventions often has not observed how interventions reflect recommendations for Tier 3. As such, there exists a need to contrast growth for children in Tier 1 only versus growth of students who receive Tier 3 intervention. The current lack of research warrants exploration to identify how minority students in Tier 3 are supported relative to their peers.

**Observations of Tier 1 and Tier 3**
Within the research base on RTI, few studies have explored how Tier 1 and Tier 3 differ and have rarely reported the quality of Tier 1 core instruction (Austin et al., 2017; Wanzek & Vaughn, 2007). Though some syntheses have examined Tier 3 intensive interventions for elementary grade students, none of them have systematically compared whether the observed components of Tier 3 instruction align with the recommended practices by Gersten et al. (2009). Additionally, they have not compared observations of components of Tier 3 to what students received in Tier 1 or core instruction (Austin et al., 2017; Wanzek & Vaughn, 2007).

Researchers (e.g. Lam & McMaster, 2014; Austin et al., 2017) have reported that quality of Tier 1 is often overlooked when examining the impacts of more intensive tiers of intervention. If response to Tier 1 determines who receives more intensive intervention, then it is problematic if research does not document that Tier 1 is of high quality.

In a prior systematic literature review, I reviewed studies in which research teams observed intensive intervention across the literature, and the extent to which intensive intervention was provided to students from minority backgrounds. I synthesized 17 studies that 1) were published after 2004, 2) were published in journals available through online databases, 3) included kindergarten through 5th grade participants; 4) provided information about Tier 3 implementation (which was expanded to include intensive intervention or special education) and Tier 1 instruction; and 5) included observations of Tier 3 (Baker, 2019). Figure 1 visually details my overall search process. Additionally, Table 1 (i.e., Summary of 17 Studies Examining RTI that Included Tier 1 Instruction and Tier 3 Intervention) provides overall Table 2 summaries of all 17 studies including: (a) research design, (b) setting, (c) student participants, (d) student diversity (i.e. minority status; SES status), (e) teacher participants, (f) Tier 3 setting, (g) Tier 3 description, and (h) impact or effect sizes.
Minority Representation Across Studies

Twelve of the seventeen studies \((k = 9\) quantitative; \(k = 3\) qualitative) reported majority-minority student samples (see Table 1, where the * before the author designates these 12 studies). Of these, 11 studies were conducted in early elementary (K-3rd grade) and only one study (Wanzek et al., in review) was conducted with in upper elementary (4th grade). Among these studies, the majority of authors \((k = 9)\) also conducted studies in schools with high populations of students from socioeconomically disadvantaged backgrounds; FRL ranged from 42 percent to 99 percent. For the purposes of this dissertation, I focused my synthesis on 12 studies that included greater minority representation (see also Baker et al, in preparation).

Effects of Tiered Instruction

Quantitative (intervention) studies. Consistent with the NRP (2000) and the Simple View of Reading (Gough & Tunmer, 1986) the intervention studies that I reviewed \((k = 9)\) often reported using interventions focused primarily on building both code-focused skills (phonological and phonemic awareness, letter sounds, and concepts of print) and meaning-focused skills (vocabulary and text reading). Yet, Tier 3 interventions typically included only three of the five components identified by the NRP: phonics, fluency, and comprehension. It was rare for studies to include phonemic awareness and vocabulary as part of the intervention, and assessing those skills was often only folded into measures of phonics or comprehension. It should be noted that although many of the studies were conducted in schools and districts with high minority populations, none disaggregated data that would inform comparisons of the impact of RTI for Black, LatinX, or White students. As a result, it was not possible to ascertain the degree to which intensive interventions were effective for minority students.
In Tier 3 intervention studies that included a majority of minority students, effects in these three skills—phonics, fluency, and comprehension—varied. The largest effects were found in measures of fluency and ranged from very large (e.g. $ES = 2.87$; Volpe et al, 2011) to very small (e.g. $ES = 0.08$; Wanzek et al., in review). Similarly, in phonics, effects also ranged from very large (e.g. $ES = 1.32$; Denton et al, 2006) to very small (e.g. $ES = 0.08$; Wanzek et al, in review). On measures of comprehension, effects were smaller than fluency and phonics, but also ranged from large (e.g. $ES = 1.00$; Denton, 2006) to small (e.g. $ES = 0.09$; Wanzek et al., in review).

Overall, the largest effect sizes favored interventions that used quasi-experimental designs, had smaller sample sizes, and were provided in early elementary. The smallest effects were found in the only study eligible for review that provided Tier 3 reading intervention to older elementary students with reading difficulties (Wanzek, et al, in review).

**Qualitative (case) studies.** In the three qualitative studies that directly observed instruction in schools with majority-minority student populations, two studies (Orosco & Klinger, 2010; Swanson et al., 2012) observed and reported typical school practice, and one study (Rinali et al., 2011) described the implementation of an RTI model and reported how the model worked. These studies often thematically described that intensive intervention showed promise, yet findings also varied (see Table 1). Orosco and Klinger (2010) specified the need to consider diverse learners when implementing RTI. In their case study they focused on English Language Learners (ELLs) in one large urban and majority minority school where 85 percent of students were LatinX and 99 percent of students received FRL. The researchers conducted a relatively large number of observations in Tier 3 and Tier 1 which included 48 observations of eight school professionals who provided RTI support. As a result of these observations, one important finding centered around the lack of consideration of the student demographics when
providing RTI supports. The authors reported that though Tier 3 was frequent and highly individualized, instruction for students appeared to be inappropriate for their linguistic skills. This finding was consistent in Tier 1 instruction as well; outside of bilingual classrooms, Tier 1 instruction often did not consider the linguistic needs of the high population of English language learners. Researchers also noted that teachers often attributed students’ reading difficulties to limited English proficiency without considering quality of instruction.

In the other study that described what typically occurred in schools, Swanson et al. (2012) described observing instruction of special education teachers, who provided intensive intervention. The researchers used a low-inference observation protocol, the Instructional Content Emphasis in Reading (ICE-R; Edmonds & Briggs 2003) to explore teacher perceptions of RTI. They found that while it was beneficial to students and they were highly engaged, teachers perceived implementing RTI as difficult due to time and scheduling constraints, paperwork load, and limited staff.

Finally, Rinaldi et al. (2011) implemented an RTI model at one school and described the overall consequences of that implementation. They described that implementing RTI led to more data usage and less referrals to special education

**Recommended Components of Tier 3 and Tier 1 Support**

Across the studies I reviewed, Tier 3 and Tier 1 supports varied; but authors often reported using or observing only some of Gersten et al.’s. (2009) recommended RTI components. For Tier 3, authors often reported that intervention was provided frequently (i.e. four to five sessions a week for 45 minutes per session), and in small groups up to four students (see Table 2 for a Summary of the Tier 3 Recommended Components across the 17 studies). However,
materials used for intervention were often standardized and authors provided little information about whether it was individualized based on student data.

Reporting for Tier 1 was similar in that many of the essential components were not described except for universal screening (see Table 3, *Summary of Tier 1 Components across 17 Studies*). All but one of the studies that had majority-minority participants reported that universal screening occurred with most ($k = 8$) reporting the frequency of the screening. A few authors ($k = 4$) also indicated the screening measures used which included *AIMSweb* (Pearson, 2012) and *Dynamic Indicators of Basic Early Literacy Skills (DIBELS)* (Kaminsky & Good, 2002). Broadly, none of the studies reported that computer adapted testing was used, including the specific MAP measure I will use for this dissertation study.

**Rigor of Observations of RTI**

Across the 12 studies of focus from my literature synthesis, only three research teams specified using standardized and validated instruments to conduct observations to document overall quality of reading instruction in both tiers of intervention; two of these were conducted in high minority schools (i.e. Swanson et al., 2012; Wanzek et al., in review). Both Swanson et al. (2012) and Wanzek et al. (in review) used the *Instructional Content Emphasis- Revised (ICE-R)* (Edmonds & Briggs, 2003), which I used in this dissertation study to observe the amount and quality of reading instruction.

**Observations in quantitative studies.** Gersten et al. (2005) suggested that rigorously-conducted observations within quantitative (especially experimental) studies should include: (a) regular observations conducted throughout the study, (b) inter-observer reliability, and (c) collection of field notes. Across the nine studies, highlighted in Table 4, observation frequency varied and ranged from two observations throughout the study, to two observations per week.
Inter-observer reliability was high and ranged from 85 to 100 percent. Moreover, all researchers reported using field notes or researcher created checklists to report the quality of Tier 3 observations.

**Observations in qualitative studies.** Brantlinger et al. (2005) additionally recommended that observations conducted within qualitative studies should include much descriptive information including: repeated observations, collection of field notes, and a description of the setting. As with quantitative studies, frequency of observations was often reported and ranged from one to seven. Likewise, in all studies, authors reported that they collected field notes. Lastly, each study provided rich descriptions of the school settings.

**Summary**

Researchers often conducted studies in schools with large populations of minority students, yet, data was not disaggregated by minority status to determine the effectiveness of Tier 3 for these populations. Also, given the high minority student involvement, we might expect that interventions would include all components of foundational reading, with special attention to the provision of vocabulary development, consistent with prior research demonstrating the importance of building this skill (Beck & McGowan, 2007; Li et al., 2020; Ortiz et al., 2012; Terry et al., 2012, Saenz et al., 2012). Though many important aspects of foundational reading were address, vocabulary was not observed. These findings also indicated the need for more studies that provide insight into the specific impacts of RTI for students overall, and especially minority students.

**Study Purpose and Research Questions**

The purpose of this study was to use the lens of CRT to add to the existing RTI research by exploring: (a) how receiving RTI relates to students’ reading growth, and more specifically
minority students’ growth, and (b) what occurs within schools that might explain how minority students grow relative to their peers. This study was guided by the following three research questions:

*Research Question 1:* Relative to students who were only classified as receiving Tier 1 instruction, what is the growth of students who received Tier 3?

*Research Question 2:* What is the growth of minority students classified as receiving Tier 3 compared to other students (including non-minority in Tier 3, non-minority in Tier 1, and minority in Tier 3)?

*Research Question 3:* Why might minority student growth differ across schools?

To answer these, I focused on the 1,002 students in three schools across one district for whom NWEA MAP (NWEA, 2015) reading scores were collected. Of the total number of students included in the study, observation data was also collected on 18 minority students. In the methods section, below, I describe the participant demographics, measures used, and data analysis.

**Key Definitions**

While a set of guidelines have been provided to help educators support students through RTI, there is still varied implementation and definitions of these services. For the purposes of this dissertation the following definitions will be used.

*Response to intervention (RTI).* RTI as an academic instructional framework that consists of tiers of increasing instructional intensity provided to students. As listed below, this typically consists of three tiers. This dissertation focused only on reading, not other content areas, or multi-tiered systems of support (MTSS).
**Tier 1.** Within Tier 1, all students receive general education or core instruction. The RTI Practice Guide (Gersten et al., 2009) recommends that Tier 1 reading instruction for elementary students should be 1) evidence-based, 2) include the five components of phonemic awareness, phonics, fluency, vocabulary, and comprehension as identified by the National Reading Panel Report (2000), and 3) differentiated to student need. Additionally, Tier 1 includes universal screening of all students at least twice per year to identify students at-risk for reading difficulties.

**Tier 2.** Tier 2 is supplemental and systematic instruction provided to students who, through universal screening and benchmark testing, do not respond adequately to Tier 1 core instruction (Gersten et al., 2009). For the purposes of this dissertation, Tier 2 students were not included in the sample due to the nature and purposes of the extant data available in the larger study.

**Tier 3.** Tier 3 is synonymous with the term *intensive intervention*, and is used to specify interventions that include a combination of frequent instruction, increased duration, small or individualized student grouping, and ongoing progress monitoring (Gersten et al., 2009) provided to students who do or are likely to respond inadequately to both Tier 1 and Tier 2 instruction. Using this set of characteristics, participating schools identified students who received intensive intervention. Across schools, intensive intervention most commonly consisted of a specialized, pull-out intervention (e.g. special education, dyslexia services) provided to a small percentage of students with or at risk of disabilities and who did not respond adequately to less intensive instruction.

**Special education.** Special education is individualized instruction provided to students who qualify to receive an individualized education plan (IDEIA, 2004). Across the sample, students who received special education services were also identified receiving either Tier 1, Tier
2, or Tier 3. Though I do not focus on special education specifically, identification and type of support provided through special education may be alternative reasons for differential outcomes for students across schools.

**Progress monitoring.** Gersten et al. (2009) outline that progress monitoring within the RTI framework typically starts when students are in Tier 2 to determine how they are responding to the intervention. If students are able to return to Tier 1, progress is monitored to ensure that students are maintaining learned skills and continuing to grow. If students need support through Tier 3, the frequency of progress monitoring increases to determine how students respond to instruction.

**Negative gain.** Negative gain is the case of students scoring lower at the end of the year than at the beginning of the year. For this dissertation, this reflects negative gain in RIT points on a CAT test, the MAP. Across the three schools, a small number of students demonstrated negative growth on the universal screener from fall to spring. Some scholars (Dahlin, 2016) agree that negative growth is mostly uninterpretable for students in regards to academic achievement throughout the year. Rather, it is an indication of measurement or testing error. Though testing procedures would ideally be consistent across classrooms, grade levels, and schools, it may be reasonable to assume that negative gain indicates inconsistency of testing procedures. I will describe procedures used to reduce outliers – both students who demonstrated too much positive or negative gain.

**Minority status.** For the purpose of this dissertation study, all racial and ethnic information was provided by the schools. In Texas, students’ families provide this data to schools and Texas follows (Texas Education Agency; TEA, 2020). For my dissertation, given the CRT, I defined minority students as those from either Black American, or LatinX descent, either
by race or ethnicity. All students who identified as LatinX by ethnicity were categorized as LatinX regardless of race. Non-minority students are those from White descent and who were not identified as LatinX by ethnicity.

**Other.** Other students are those who did not identify as Black American, LatinX, or White by race and did not identify as LatinX by ethnicity. Note that I did not include (other races), because they were not part of my intended sample.
CHAPTER 2
METHODS

Larger Dataset

Data for this dissertation were taken from a larger grant-funded project, Project FOCUS. In three years of data collection, this project has explored the relation of Response to Intervention (RTI) implementation approaches and student reading outcomes, with a focus on students with intensive needs – specifically students classified by schools as receiving Tier 3 services. Given that data were collected across the nation, different schools used different assessments to measure reading ability. The overall project analyzed data using mixed methods to describe: (a) interviews with school leaders’ about school practices for interventions; (b) teachers’ knowledge about implementing RTI; (c) observations of intensive interventions for students with or at-risk for disabilities; and (d) the relations between student reading growth and collected RTI implementation data. We also explored whether student disability status, RTI status, and demographics moderated these relations.

Restatement of Purpose for the Current Study

As previously stated, the purpose of this study was to add to the existing research on RTI by using the lens of Critical Race Theory in Education. I used a mixed methods design to explore the relation of RTI (Tier 3) eligibility on overall student reading growth and a specific focus on reading growth for minority students. In the quantitative portion of my analyses, I explored the relation between RTI tier (Tier 1-only vs Tier 3 eligibility), minority status, and student growth. I analyzed data within schools which allowed to me to learn if and how student growth differed by school. In the qualitative portion of my analyses, using comparative case study procedures I conducted a case study of each school to describe similarities and differences between them.
including school profiles, RTI implementation, and student observations at each school. The aim of the comparative case study was to explore plausible explanations for differences in student growth between the schools to build theories and to inform future research. To the greatest extent possible, with the data available to me, I followed quality indicators for rigorous research in both the quantitative and qualitative portions of my study.

**School Settings and Student Participants**

Within this section, I report at the school level, the demographics for my sample, then I describe the larger school setting, including the entire school demographics, and other contextual information about the school. Through this dissertation, I focused on minority student reading growth during the 2017-2018 school year across three schools in one district.

From the FOCUS data (Al Otaiba et al., 2018), I selected schools that administered the MAP assessment at the beginning of the year (Fall 2017 administration) and end of year (Spring 2018 administration). Then I took a series of steps to refine school (and subsequently, participant, which is described below in student participant section) eligibility for my study. As seen in Figure 2, I first selected schools from the larger dataset that administered the MAP assessment as a universal screener. This provided a starting sample of 3,274 students across three districts and seven schools. Second, the data appeared to have a high percentage of students who had negative scores at fifth grade, so in consultation with Dr. Yovanoff and Dr. Al Otaiba, I eliminated fifth graders from the dissertation dataset to maximize the credibility of my data. Third, I selected schools in which there were valid observations for at least four minority students (one student per grade in first – fourth grades). Fourth, given that some schools only provided data for students in Tier 3 and Tier 1 for the broader project, I removed students from the data who were designated as “Tier 2” or who were missing a tier designation. Fifth, there
were students \((n = 30)\) for whom there was extreme positive or negative growth. I trimmed the data for those outliers (Identifying Outliers: Upper Fence & Lower Fence, 2017) which is a process of removing students with extreme positive or negative scores. Finally, in order to focus on minority students compared to their non-minority peers and because there were too few students in the other racial/ethnic subgroups to make meaningful comparisons, I removed students \((n = 24)\) who did not self-identify as Black, LatinX, or White. This resulted in a final sample of 1,002 minority and non-minority students from three schools (School E, School B, and School J) who reflect 61% of the total student population across each school (see Figure 2).

**Student Participants**

As indicated in Table 5, which summarizes the number of students by tier and minority status and indicates those that received only Tier 1 (Tier 1) or who were Tier 3 eligible (Tier 3), 1,002 students were sampled for this dissertation. At School E, 254 students were sampled for data analysis \((n = 169\) students in Tier 1 only; \(n = 85\) students in Tier 3. At School B, 352 students were sampled for data analysis \((n = 177\) in Tier 1 only; \(n = 175\) in Tier 3). Finally, at School J, 396 students were sampled for data analysis \((n = 359\) in Tier 1 only; \(n = 37\) in Tier 3). Of this, 17 minority students \((n = 2,\) Black American; \(n = 15,\) LatinX) students represented participants with the most intensive needs and were also observed receiving tiered instruction in both Tier 1 and Tier 3. Across the sample, schools reported that students who qualified for free or reduced lunch programs, which is an indicator of economic disadvantage, ranged from 78.3 percent to 89.5 percent.

**School Demographics and Settings**

Each of the three schools is located within a mid-sized suburban district in the southeastern region of the United States that serves a majority-minority student population. In the
sections below I first describe the general school demographics, and then describe details about
the school setting. Overall descriptions of the school demographics for each school can be seen
in Table 6.

**School E.** Like all of the schools represented in the sample, School E had a majority-
minority student sample, with LatinX as the largest ethnic category. At School E, 75 percent of
students identified as LatinX, five percent identify as Black American, 14 percent identify as
White and the remaining students identified as a different race or ethnicity. Seventy-eight percent
of students qualified to receive free or reduced lunch.

School E is located in the southernmost part of the district and, at data collection, had a
population of 697 students, led by the principal who was a LatinX female. The school employed
43 teachers, including 26 White, 15 LatinX, and two Black American. About half of the teachers
\((n = 21)\) had five or more years of classroom experience. The district website reported that nine
percent of students at School E receive special education services (compared with the eight
percent district average). On a grading scale of A-F, the school has an overall accountability
rating of B and data (The Texas Tribune, 2020) suggests that 71 percent of the student
population is at risk for future dropout.

**School B.** At School B, 76 percent of students identified as LatinX, nine percent as Black
American, and nine percent identify as White. The remaining students identified as a different
race or ethnicity. Additionally, 85 percent of students received free or reduced lunch.

School B is located about a mile and a half from School E. At the time of data collection,
it had a population of 816 students, led by the principal who was a White male. The school
employed 49 teachers, including 24 White, 23 LatinX, and two Black American, and 19 had five
or more years of classroom experience. Lower than School E, the district website reported that
just over six percent of students at School B received special education services (compared with the eight percent district average). On a grading scale of A-F, the school has an overall accountability rating of C and data (The Texas Tribune, 2020) suggests that 75 percent of the student population is at risk for future dropout.

**School J.** At School J, 77 percent of students identified as LatinX, five percent identify as Black American, 10 percent identify as White and the remaining students identified as a different race or ethnicity. Ninety percent of students qualified to receive free or reduced lunch.

School J is located about three miles from School E and two miles from School B. At data collection it had a population of 745 students, led by the principal who was a White female. The school employed 47 teachers, including 22 White, 22 LatinX, two Black American, and one Asian with 20 teachers who had five or more years of classroom experience. The district website reported that seven and a half percent of students at School J receive special education services (compared with the eight percent district average). On a grading scale of A-F, the school has an overall accountability rating of C and data (The Texas Tribune, 2020) suggests that 58 percent of the student population is at risk for future dropout, much less than in Schools E or B.

**Measures and Data Sources**

**Student Reading Growth Measures: MAP**

As described in Chapter 1, the MAP is a computer-adapted, norm-referenced assessment that many schools use as a universal reading screener (NWEA, 2015). According to NWEA, the reading assessment was normed on a sample of a minimum of 72,000 students at each grade level in Kindergarten through 11th grades. The Internal consistency ranges from 0.71 - 0.85 for first thorough ninth grade. Concurrent Validity ranges from 0.58 – 0.83. Predictive Validity
ranges from 0.63 – 0.82, and Criterion Related Validity ranges from 0.37 – 0.67. Growth norms are provided for each grade.

Upon completion of the assessment, students receive a score that is a Rasch Unit Score (RIT Score) which is scaled vertically and with equal intervals. This allows for students to be compared in terms of their individual growth within and across years, and to compare their growth to other students in their same age/grade range. As seen in Figure 3, NWEA provides means and standard deviations for growth from beginning to end of year for each grade and testing cycle, with standard deviations between six to eight points of the mean (i.e. first grade, \( M = 16.8 \) points, \( SD = 8.09 \); second grade, \( M = 14.00 \) points, \( SD = 8.20 \); third grade, \( M = 10.30 \) points, \( SD = 7.59 \); fourth grade, \( M = 7.7 \) points, \( SD = 7.05 \)) (NWEA, 2015). Normative data provided by NWEA (2015) indicates that students demonstrate the greatest fall to spring gains on RIT scores in first grade, trending toward less growth in each subsequent grade level.

Although MAP data may be collected at beginning, middle, and end of year, for the present study I focused on growth from the first assessment at the beginning of the year and final end of the year. Middle of year was excluded to maximize the sample of students for whom growth data was available. In summary, this study focused on these three schools in one suburban district that collected MAP data for universal screening monitoring during the 2017-2018 school year (i.e., within the second year of the larger project).

**School Level RTI Implementation: RTI Essential Components Structure Interviews**

As part of the larger study, principals at each of the schools were interviewed using a standard interview protocol, the *RTI Essential Components Worksheet* (Center on Response to Intervention; AIR, 2014; see Figure 4), to understand essential RTI practices within their schools. This Worksheet was developed by the American Institutes for Research (AIR);
however, to date, no studies have reported the reliability and validity of the worksheet. The essential categories included six components, which our team further sub-divided (note the categories we subdivided are noted by an *) into eight categories covering the following:

1) Assessment: Screening*
2) Assessment: Progress monitoring*
3) Data Based Decision Making
4) Multi-level Instruction: Tier 1
5) Multi-level Instruction: Tier 2*
6) Multi-level Instruction: Tier 3*
7) Infrastructure and support Mechanisms
8) Fidelity and Evaluation

After the interviews, responses were rated by the research team conducting the interviews on a rubric that incorporated a 1-5 Likert scale (RTI Fidelity of Implementation Rubric, AIR, 2014). A question response rating of 5 indicated that during the structured interview, principals detailed at least three components in response to the questions asked. A rating of 3 indicated at least two components was described, and in contrast, a rating of 1 indicated that principals provided very limited or no detail for the component. Ratings were then averaged to provide an RTI implementation score in each category.

Principal interviews from the 2017-2018 school year were conducted by myself and another trained graduate research assistant. Interviews were rated and transcribed for analysis. Notes were made to support ratings. To establish interrater reliability, the two interviewers double coded 25 percent of the interview transcripts, and established reliability on the rubric ratings of 81 percent.
In addition to the rated questions, interviewers asked a series of nine open-ended questions as an addendum related to supporting English Language Learners (ELLs) and how schools included social-emotional learning (SEL) into the curriculum. Across the three schools, a total of 81 pages of interview data was available for analysis. A discourse review of these data sources regarding themes from interviews provided further insight into the strength of RTI practices at the school level given that principals outline the direction of academic practices within the school and may moderated the relation between students’ minority status and reading outcomes.

**Tier 1 and Tier 3 Reading Instruction Observations: Instructional Content Emphasis in Reading**

Observation data for the larger study was collected by a team of trained research assistants. Observations conducted in the three schools included in this study were completed by myself and five additional observers. Of the six observers, one was a Black American female (me), two were LatinX females, two were White females, and one was an Asian male.

Prior to conducting observations, graduate research assistants were trained to use a low-inference observation tool, the *Instructional Content Emphasis in Reading* (ICE-R; Edmonds & Briggs, 2003; see Figure 5). They were required to meet 90 percent reliability gold standard using videotaped reading instruction as part of ICE-R standard procedure. Researchers of the tool established content validity through consultation with field experts and review of literature (Edmonds & Briggs, 2003).

As part of the larger study, schools assisted recruiting and consent procedures and were asked to nominate students for observation (eligibility criteria included either participation in
Tier 3 or special education). The graduate research assistants observed literacy instruction for students who received general education and Tier 3 reading support as defined by the school.

Observers completed one observation during participating students’ Tier 1 and Tier 3 instruction. These observations were intended to represent a snapshot of literacy instruction at each Tier, within that school. Following ICE-R standard procedures, observers coded for length of instruction, type of instructional activity, and instructional grouping for each minute of instruction. Along with type of instruction, observers coded overall student engagement on a 3-point Likert scale (3 = high engagement and 1 = low engagement). Observers also coded for overall quality of the lesson on a 4-point Likert scale (4 = excellent and 1 = weak) (Edmonds & Briggs, 2003). Observers determined the quality rating by highlighting when teachers met observable behaviors specified by the ICE-R Quality Indicators and Descriptions.

These school-level observations may contribute important data about how schools operationalized RTI implementation in Tier 1 and Tier 3. This represents school-level data and we did not link specific observations to student reading growth data. Recall that students who qualified for intensive intervention (operationalized as either Tier 3 or special education) were identified by their schools and they were observed receiving reading instruction both in Tier 1 core instruction (General Education) or Tier 3 intensive intervention (defined by the school, but typically consisted of special education or dyslexia services). Furthermore, findings from these observations may provide insight into the strength of RTI practices that students received and whether or not observed differences inform plausible explanations for differences in reading growth patterns for minority students in Tier 3 relative to other students (i.e. non-minority students in Tier 3, non-minority students in Tier 1, and minority students in Tier 1).

Data Analytic Methods
Given my conceptual framework for the study (Critical Race Theory in Education; Ladson-Billings & Tate, 1995), and my empirical framework related to RTI (i.e. hypothesized differences in intensity between Tier 1 and Tier 3, and differentiated instruction for Tier 3), I used a combination of quantitative and qualitative data to answer my three research questions. My data analytic methods were conducted as within school analyses, which considered each school as an individual case study and answered the research questions within the context of each school. My initial research questions required quantitative analytic methods; therefore, described below are the methods I used to address these, including the specific series of regression models. This series of models informed how I selected cases for in-depth qualitative analyses. Thus, the final research question addressed plausible explanations for differences in observed reading growth for my population of interest (minority students in intensive reading interventions). Using the data sources, a-priori plausible explanations included: (1) articulation of data-based decision making for eligibility for Tier 3, (2) proportion of students in special education are also served in Tier 3, and (3) intensity of intervention in Tier 3, and might provide plausible explanations for differences in student reading growth for minority students relative to peers at the three schools.

The first two research questions were answered using a correlational design to explore the relation between RTI tier, minority status, and student growth within schools. I conducted a series of linear regressions (See Table 7, Models 1-4). Regression models are flexible models and allowed me to test variables without relying on analysis of variance assumptions (Heck, Thomas, & Tabata, 2013). Analyzing the growth patterns in each school separately served to set up the analysis of the individual cases within my case study.
To address research question 1, “Relative to students who were only classified as receiving Tier 1 instruction, what is the growth of students who received Tier 3?” I estimated a simple linear regression which allowed me to test the correlation between my scale dependent variable (Reading growth) and my scale independent variable (RTI Tier). In Model 1, I estimated the following regression model:

\[ R = b_0 + b_1(tier3) + e \]

Where,

\[ R = \text{reading fall-to-spring growth} \]
\[ b_0 = \text{mean score for tier 1 students} \]
\[ b_1 = \text{the effect of RTI ‘Tier 3 status’ on ‘reading growth’. ‘Tier3’ is a dummy variable coded 0,1 indicating Tier 3 status (1) relative to Tier 1 status (0)} \]
\[ e = \text{estimation error} \]

In addition, I estimated two additional models to assess the main effects of minority status across the schools. To do this, I first replaced the tier dummy coded variable with a minority status dummy coded variable using the following model (Model 2):

\[ R = b_0 + b_1(minority) + e \]

Where,

\[ R = \text{reading fall-to-spring growth} \]
\[ b_0 = \text{mean score for Tier 1 students} \]
\[ b_1 = \text{the effect of ‘Minority status’ on ‘Reading growth’. ‘Minority’ is a dummy variable coded 0,1 indicating Minority status (1) relative to Non-minority status (0)} \]
e is the estimation error

Next, I estimated the main effects both Tier 3 status and Minority status on reading growth outcomes using the following model (Model 3):

\[ R = b_0 + b_1(Tier3) + b_2(Minority) + e \]

Where,

- \( R \) = reading fall-to-spring growth
- \( b_0 \) = mean score for tier 1 students
- \( b_1(Tier3) \) = the effect of RTI ‘Tier 3 status’ on ‘Reading growth’. ‘Tier3’ is a dummy variable coded 0,1 indicating Tier 3 status (1) relative to Tier 1 status (0)
- \( b_2(Minority) \) = the effect of ‘Minority status’ on ‘Reading growth’. ‘Minority’ is a dummy variable coded 0,1 indicating Minority status (1) relative to Non-minority status (0)
- \( e \) = estimation error

To address RQ2, “What is the reading growth of minority students in Tier 3 relative to other students, including non-minority students in Tier 3, minority students in Tier 1, and non-minority students in Tier 1?” I estimated a multiple linear regression to test the significance of the interaction between Minority status and RTI tier. Using the variables described above, I ran the following model (Model 4):

\[ R = b_0 + b_1(Tier3) + b_2(Minority) + b_3(Tier3 \times Minority) + e \]

Where,

- \( R \) = Reading fall-to-spring growth
- \( b_0 \) = the mean score for Tier 1 Non-minority students
b₁ (Tier3) = the effect of RTI ‘Tier 3 status’ on ‘Reading growth’, controlling for ‘Minority status’, where ‘Tier3’ is a dummy variable coded 0,1 indicating Tier 3 status (1) relative to Tier 1 status (0)

b₂ (Minority) = the effect of “Minority status” on ‘Reading growth’, controlling for ‘Tier 3 status’, where ‘Minority status’ is dummy variable coded 0,1 indicating Minority status (1) relative to Non-minority status (0)

b₃ (Tier3 x Minority) = the Minority by Tier 3 interaction effect on ‘Reading growth’ (which is a test of moderation)

e = estimation error

To address RQ 3 “Why might student growth differ across schools? What is the experience of Tier 1 vs Tier 3 or Special education for minority students in schools implementing RTI?” I conducted a case study (Yin, 2014) given that schools are different and that students’ experiences often vary depending on the school they attend. A case study allowed me to explore the unique characteristics at each school and explore and describe the differences in RTI implementation and student reading growth that might be plausible explanations for trends in different student growth for minority students in Tier 3 vs peers. These explanations may inform theory and future research directions.

Yin (2014) highlighted that case studies should be guided by theory and are helpful when investigating questions within a set of data relating to that theory. Yin further specified that while case studies can use both quantitative and qualitative data, they must include multiple sources of data to allow for triangulation between sources. In addition to the quantitative data attained
though Research Questions 1 and 2, two additional data sources were 1) the principal interviews for overall school level RTI implementation and 2) the observation data. Through this case study approach, I was able to enact the power of voice through storytelling, a necessary component of CRT-E, to highlight the experiences of minority students receiving RTI supports. Furthermore, through descriptive analyses, I was able to explore differences in schools that, due to power, could not be numerically demonstrated.

Through 81 total pages of transcripts of the principal interviews as well as ratings assigned to each interview from the *RTI Implementation Rubric* (AIR, 2014) I analyzed how the school level data, as reported by the principals, might account for variability in student growth outcomes across the schools. This included a combination of analyzing the ratings assigned to each category of questions on which principal interviews (see Table 7), and the actual responses within the transcripts.

Finally, through the observation data (Figures 6-8), I was able to triangulate student data and principal interviews with what students were observed receiving. As previously mentioned, each student was observed a total of two times- once in Tier 3 and once in Tier 1. Therefore, they were snapshots representative of typical instruction that minority students received at each tier of RTI at the school level. I analyzed 152 pages of ICE-R protocols to unpack similarities and differences in overall literacy instruction which included duration, instructional grouping, global student engagement, and instructional quality across Tier 3 and Tier 1 observations for each student.
CHAPTER 3

RESULTS

I used a mixed methods design to study how minority students are supported within the RTI framework. I first quantitatively analyzed the data within each school to see the relation between RTI tier, minority status, and student growth. Analyzing within schools allowed me to learn if and how student growth differed by school. This also allowed me to then analyze each school through qualitative case studies to understand the differences between schools that might account for variability in student growth.

Research Question 1. Relative to students who were only classified as receiving Tier 1 instruction, what is the growth of students who received Tier 3?

Research Question 2. What is the growth of minority students classified as receiving Tier 3 compared to other students (including non-minority in Tier 3, non-minority in Tier 1, and minority in Tier 3)?

Table 7 (Student Growth in Tier 3) summarizes results for the series of models I conducted within four regression analyses to address these first two research questions, respectively. In Model 1, I used linear regression to analyze the overall growth of students in Tier 3 relative to those in Tier 1 at each of the three schools. Then, to analyze whether the relation between growth in tiers depended on minority status, I first conducted two additional linear regressions to test the main effect of minority status (Model 2) and the main effects of tier and minority status with each controlling for the other (Model 3). Across the schools, there was not a main effect of minority status, indicating that student growth was not solely related to race. Additionally, the inclusion of race did not meaningfully change the significance of the main
effects of tier when modeled together therefore, I will not be discussing the Models 2 and Model 3 estimates.

In my final model (Model 4) I conducted a multiple linear regression to model the interaction between minority status and tier of intervention to analyze the reading growth of minority students in Tier 3 relative to other students. Given my desire to understand how minority students have been supported in intensive intervention (Tier 3), to develop my case studies, I focused on the Model 4 estimates. In the following sections, I first describe the findings from Model 1, then the findings from Model 4 for each school. I describe the findings starting with Schools E and B; both schools had effects favoring growth of minority students in Tier 3. I then describe the findings for School J which indicated less growth for minority students in Tier 3.

School E

At School E, and as seen in Table 7 Model 1, results were significant and demonstrated that on average, students in Tier 3 demonstrated two points less growth on RIT scores relative to students in Tier 1, $b = -2.34$, $t (253) = -2.18$, $p = .03$ with a small effect of $\beta = 0.13$. When analyzing specifically how minority students grew in relation to their peers (Model 4), the data indicated that there was still a significant main effect of RTI tier such that students in Tier 3 scored nearly seven points lower than students in Tier 3, $b = -6.94$, $t (253) = -2.12$, $p = 0.04$ and a moderate effect of $\beta = -0.39$. Although the interaction between tier and minority status was not significant statistically at alpha = 0.05, it indicated a positive trend for minority students such that they averaged five points higher growth relative to all other students, $b = 5.16$, $t (253) = 0.29$, $p = .14$ with a moderate effect of $\beta = 0.29$.

School B
At School B, the regression analysis in Model 1 demonstrated that from Fall administration to Spring Administration, students in Tier 3 grew, on average, nearly two RIT points more than students in Tier 1. These results were statistically significant, $b = 2.068$, $t (351) = 2.461$, $p = 0.014$ and represented a small effect of $\beta = 0.13$. When adding minority status and the interaction between minority status and tier to the model (see Model 4), the main effect of RTI tier dropped, indicating no difference between growth of students in Tier 3 in relation to Tier 1. There was also no interaction effect which indicated that growth in either tier of intervention did not depend on minority status. However, the interaction term suggested a pattern that minority students in Tier 3 at B Elementary grew nearly three RIT points more relative to all other students in the school regardless of tier or minority status, $b = 2.94$, $t (351) = .92$, $p = 0.36$, with a small effect of $\beta = 0.18$.

**School J**

Finally, at School J, results from the overall regression (see Model 1) indicated a trend that students in Tier 3 score higher by nearly three points than students in Tier 1. However, this finding was not significant, $b = 2.65$, $t (395) = 1.87$, $p = .06$, and had a small effect of $\beta = 0.09$. However, this changed when analyzing growth of minority students in Tier 3 in Model 4. These results indicated a significant main effect of tier in favor of Tier 3 students such that on average, students in Tier 3 made about 10 points more growth than students in Tier 1, $b = 10.54$, $t (395) = 2.22$, $p = 0.03$, with a moderate effect of $\beta = 0.38$. Interestingly, this pattern of overall growth did not hold true for minority students in Tier 3. Though non-significant, the interaction between tier and minority status indicated that minority students in Tier 3 demonstrated approximately 8 points less growth in relation to all other peers, $b = -8.20$, $t (395) = -1.65$, $p = .1$, with a moderate effect of $\beta = -0.28$. 
Research Question 3. Why does minority student growth differ across schools?

For research question 3, I used a qualitative comparative case study approach (Yin, 2018) to explore plausible explanations for the trend in differences (from Model 4) among the three schools for growth of minority students in Tier 3 relative to other students (including non-minority students in Tier 3 and all students who participated in Tier 1). This comparative case approach is framed by the initial conceptual frameworks (e.g., effective, well-implemented Tier 3 should help close the reading gap and, as suggested by CRT, closing the gap signifies progress toward equitable outcomes for students of color) and by exploring plausible explanations for differences among the schools.

Given that all three schools are located within the same district and had similar demographic profiles (see Table 4), I considered real-world alternative plausible explanations that might provide for the quantitative findings (Yin, 2018). As described in the school settings section of the Methods (p. 26), these schools were located within a three-mile radius. The demographics in schools were similar in several ways (e.g., school size, percentages of students from different ethnic groups, and teacher demographics). All schools had relatively high participation in FARL. However, there were some differences among principals racial/gender. School E was led by a LatinX female principal, School B by a White male, and School J by a White female. I explored what qualitative factors may have contributed to the differences in RTI growth profiles demonstrated between each school.

In this comparative case study, the schools represented three unique cases of RTI implementation. School E and School B represented two cases in which results demonstrated positive reading growth trends for minority students who received Tier 3. In School E, minority
students in Tier 3 grew, on average, nearly five RIT points higher growth on the MAP from fall to spring compared to all other students. This represented nearly a third of a standard deviation more growth for minority students in Tier 3 relative to other students. A similar trend was observed in School B, but growth was relatively less; minority students in Tier 3 grew approximately three RIT points more than either reference group, which was nearly one-fifth of a standard deviation more growth. Thus, Schools E and B represent schools trending toward narrowing the gap in reading performance for minority students in Tier 3 relative to other students (non-minority students in Tier 3 as well as the reference group of students who received only Tier 1) and I expected some commonalities between the two that would lead to similar patterns of growth. By contrast, School J represented a case of a school demonstrating a negative trend for minority students in Tier 3 who averaged eight points less growth (or just over a fourth standard deviation) than all other students. I hypothesized that something different may have occurred in School J which would lead to different outcomes.

The case study explored whether there were plausible explanations that emerged through the principal RTI Implementation interviews and the observations of reading instruction and intervention.

**Schools E and B vs. School J: Two Cases of Greater Growth and One Case of Less Growth for Minority Students in Tier 3**

**Plausible Explanation 1: Articulation of Decision Making about Eligibility**

When reviewing transcripts from the principal interviews and the subsequent ratings on the *RTI Implementation Rubric* (AIR, 2014), visual analysis of the data (see Table 7 for the Mean Scores on the RTI Implementation Rubric) appeared to demonstrate little overall variability in the eight broad RTI components (*i.e.* 1) Assessment: Screening, 2) Assessment: Progress
monitoring, 3) Data Based Decision Making, 4) Multi-level Instruction: Tier 1, 5) Multi-level Instruction: Tier 2, 6) Multi-level Instruction: Tier 3, 7) Infrastructure and support Mechanisms, and 8) Fidelity and Evaluation; see Table 7 for the descriptive statistics of ratings on these categories across the schools). However, each category was subdivided into several additional sections of questions for a total of 31 rated sections. Per the rubric, which was scaled from 1-5, a rating of “5” indicated that which interviewed, the principal described that the school met all of the essential elements described within the question. Conversely, a rating of “1” indicated that in the interview with the principal, the school met one or less of the described elements. Though there was little variability in the ratings averaged across the eight broad components of RTI implementation, a closer look into the categories by which the components were subdivided revealed vastly different ratings on questions related to decisions made to move students between tiers and place students in higher tiers of intervention.

One aspect/example of the theme of articulation of decision making emerged about how principals used data to guide decisions. I examined two sets of questions that queried principals on their schools’ use of data to make decisions. The first set of questions, under the heading “Data Points to Verify Risk”, were asked within the component of “Assessment: Screening” and asked “Do you review other information to help verify that the results of the initial screening are accurate before placing a student in secondary-level or intensive intervention? If so, what other types of assessment data do you use?” (RTI Essential Components Worksheet; AIR, 2014). The second set of questions, under the heading “Decision Making Process” were under the broad component of “Data-Based Decision-Making” and asked “Can you describe how decisions are made to move students between tiers?”. Though the questions were listed under two different
components, there was often overlap in principals’ responses for the two questions. Therefore, it seemed fitting to examine responses to both as an indication of the decision making process.

**School E and School B.** At Schools E and B, the principals indicated that when making decisions about moving students between tiers, they took into account multiple sources of student data, conducted meetings among relevant team members, and included teacher judgement as part of the decision making process. Specifically, in the case of School E, where minority students made stronger reading growth by eight RIT points than other students in the school, when answering the question about how decisions were made to move students between tiers, the principal at School E articulated the decision making process in this way:

All of our students are screened. And then according to the results, we move down and we rank our students beginning with the lowest reading levels and math scores…And then this is where the review of previous data…we look at whether or not there's previous parent meetings or teacher interventions [on progress monitoring assessments]. And then we look and see whether they didn't qualify for sped or dyslexia or any other services. And then we group students by need and really finalize that schedule.

Similarly, in the case of School B, where minority students in Tier 3 grew 3 points more than other students, the principal articulated three ways in which they made decisions. As seen in the following quote, they used screening and other data to identify students who did not respond to general instruction or supplemental interventions, they moved responders back to less intensive intervention, and they intensified instruction for non-responders.

So our grade levels meet as grade level teams on a regular basis…And we look at that data, we analyze the data and determine which students were not successful in that Tier 1 curriculum. And then we form groups and discuss how we're going to best remediate
those students…and what would really be the process for looking between Tier 1 and
Tier 2 kids fluctuating between, you know, the tiers there…Once they are identified as
Tier 3, they receive a separate intervention. And…then we will we do that intervention
for an entire semester and then, at the middle of the year when we do the screener again,
we look to see that they have made progress and that they are now on or above grade
level. And if they are, then they’re removed from that intervention and just put back into
Tier 1. If they are still below grade-level then that intervention continues. (Principal,
School B)

School J. In contrast, in the case of School J, where minority students in Tier 3
demonstrated less reading growth than other students, the principal provided relatively less
overall detail than principals in Schools E and B about how they utilized multiple data sources
including screeners, how decisions were made, and who was involved in making those decisions.
Instead, the following excerpt from the principal demonstrated a sole reliance on test scores
when making the decision.

Yes. [We use] common assessments, reading levels, DRA levels, previous state testing
scores, common summative and common formative assessments.

Though not readily apparent in the schools’ mean scores on the RTI Essential
Components Rubric, these different responses about data based decision making that reflect
teacher input and a team approach to making decisions. School E and School B both specified
using multiple data sources and a team level approach to understand the unique needs of each
student in intensive intervention. Given that CRT specifies the need for equitable outcomes for
students of color as a demonstration of movement toward fairness and equality, the deliberate
consideration of data is likely of particular benefit to minority students who are served in Tier 3.
Plausible Explanation 2: Identifying Students in Tier 3

Research typically presents the supporting structure of RTI as a pyramid with three levels (Fuchs, Fuchs, & Compton, 2004). On the bottom level of the pyramid, the majority of students (80 -85 percent) are served in Tier 1. In the middle level, a much smaller percentage of students (10 – 15 percent) who do not respond to well-implemented Tier 1 are identified to receive supplemental support in Tier 2. At the top of the pyramid, only a handful of students (five percent) would be expected to need Tier 3 support, or intensive intervention (National Center on Intensive Intervention, n.d.). However, converging research has shown that for students from underserved communities, including minority students and those from low-SES backgrounds, the pyramid may be inverted with a much larger percentage of students needing Tier 3 intervention relative to those in Tier 1.

To examine how this presented in the three cases, I first explored in which tier students, both minority and non-minority, were served. As a reminder, the percentages represented only correspond to the students who were in the sample, rather than the total number of students in the school. Of the students for whom tier data was provided, at each school roughly similar percentages of minority students relative to non-minority students were served in each tier of intervention (see Table 5). However, an interesting pattern emerged: there were vast differences in percentages of students who received Tier 3 supports across the schools. School E and School B had the highest percentages of students in Tier 3 (33.50 percent and 49.70 percent respectively). Comparatively, School J had a much lower percentage of students across the school who received Tier 3 support, with only 9.90 percent of students across the school in Tier 3 and 90.10 percent of students identified only as Tier 1.
I then explored the percentages of students at each school who received special education services. Nationally, during the 2017–18 school year, the prevalence of public school students who received special education services under the Individuals with Disabilities Education Improvement Act (IDEIA, 2004) totaled seven million (14 percent), including 34 percent who qualified under the category of specific learning disabilities (NCES, 2019). As shown in Table 9, the sample reflected that percentages of students with disabilities was similar at each school: School E \((n = 5.4\text{ percent})\), School B \((n = 6.3\text{ percent})\), School J \((n = 6.6\text{ percent})\). As shown in Table 5, at Schools E and B, students who were in special education, regardless of minority status, were often identified as receiving Tier 3 services as well, rather than Tier 1 only. However, at School J, nearly all students in special education received support solely in Tier 1.

Schools E and B more closely resembled the inverted pyramid of tiered student support whereas, though demographically similar, School J more resembled the typical RTI triangle. These vast differences in proportions of students receiving Tier 3 between the schools could be an indication that students were not appropriately identified for receiving necessary services either in Tier 3 or in special education at School J. This could also signify programmatic differences in types of reading instruction across the schools as a plausible explanation for differential reading growth for minority students in Tier 3. Given the fact that at School J, students in Tier 3 overall demonstrated more growth than those in Tier 1 (as seen above in Table 7, Model 1), for minority students, the lack of targeted Tier 3 instruction might be especially harmful.

**Plausible Explanation 3: Instruction and Intervention in Early Elementary**

Another source for exploring plausible explanations for the trend of differences favoring reading growth for minority students in Tier 3 was observations of reading instruction and
intervention. Converging empirical evidence from the National Reading Panel (NRP, 2000) and theoretical support from The Simple View of Reading (Gough & Tunmer, 1986) emphasized the importance of building foundational reading skills in early elementary through early intervention. This guidance suggests that instruction should include a balance of both code-focused and meaning focused skills in early elementary (learning to read) and progress to more meaning focused skills in later elementary (reading to learn) (Adams, 1990). Additionally, research in RTI suggests that early intervention in early elementary may be more powerful than remediation in the upper grades (Foorman et al., 1998; Snow, Burns, & Griffin, 1998; Jitendra et al., 2004; Denton, 2012). Therefore, I sought to learn if there was a pattern of differences in the observations of Tier 1 vs Tier 3 in lower (first and second) or upper (third and fourth) elementary grades that might explain differences in performance growth across these three cases. These observations were conducted during both Tier 1, class-wide instruction and during the more intensive services in Tier 3. The summaries (as shown in Figures 6-13) represent a school level description of the supports provided to students identified as having the most intensive reading need.

Strikingly, there were many similarities across the schools in the observations. However, with only one observation per student in each tier of instruction (Tier 3 and Tier 1), it should be reiterated that observations represented a snapshot of instruction, rather than multiple observations for each student. Within the Case of School E, where minority students made stronger reading growth than other students, as seen in Figure 4, seven students ($n = 3$ early elementary; $n = 4$ upper elementary) were observed receiving Tier 3 instruction. Tier 3 observations averaged 21 minutes in length and Tier 1 averaged 44 minutes. Similarly, at School B, five students ($n = 3$ early elementary; $n = 2$ upper elementary) were observed receiving Tier 3
which averaged 24 minutes in length and Tier 1 which averaged 20 minutes. Finally, at School J, where minority students made less reading gains than other students, five students ($n = 3$ early elementary; $n = 2$ upper elementary) were observed receiving Tier 3 for 23 minutes in length on average and Tier 1 for 41 minutes on average.

**Overall descriptions of observations.** Across the schools, observed instruction included both code and meaning focused skills (Gough & Tumner, 1986) and many of the elements of described as necessary to foundational reading (NRP, 2000). As seen in Figure X, across the schools, comprehension was the largest category of instruction in both observed tiers as well as in early and upper elementary. Phonics was the next largest instructional category and occurred more often in Tier 3 relative to Tier 1. Fluency occurred sporadically across the schools, with only one school (School J) providing fluency instruction in Tier 3/upper elementary, and another school (School E) providing fluency instruction in Tier 1/upper elementary. Vocabulary instruction was also relatively small across all of the schools. Phonemic awareness was the least observed instructional category and across schools and only occurred in Tier 3.

**Early elementary**

**Tier 3 intervention.** Tier 3 intervention at School E, which demonstrated patterns of the greatest growth for minority students in Tier 3, composed many of the elements of code-focused and meaning focused skills including phonics (30 percent), comprehension (21 percent), and phonemic awareness (five percent; see Figure 6). There were similar patterns at School B, which also demonstrated patterns of growth for minority students and early elementary Tier 3 instruction included phonics (26 percent), vocabulary (14 percent), and phonemic awareness (four percent). At School J much of the early elementary intervention in Tier 3 was comprised of comprehension (42 percent) and fluency (32 percent). Interestingly, fluency intervention in Tier
3 was not observed either at Schools E or B. As seen in Figure 7, instructional grouping was similar across the schools and students received the majority of Tier 3 early elementary instruction in small groups (School E = 82 percent; School B = 100 percent; School J = 82 percent).

**Tier 1 instruction.** Tier 1 instruction for early elementary students, shown in Figure 8, varied across the schools. At School E, which demonstrated relatively larger growth for minority students in Tier 3 among the three schools, Tier 1 was mostly comprised of comprehension instruction (72 percent). Schools B and J spent relatively less time on comprehension (32 percent and 24 percent respectively) and more time on writing (22 percent and 25 percent respectively) and phonics (12 percent and seven percent respectively). No school was observed to provide fluency or phonemic awareness instruction in Tier 1 to minority students in early elementary. Instructional grouping was quite varied between the schools (see Figure 9). At School E, nearly all Tier 1 instruction in early elementary was provided in whole class format (86 percent). At School J, which was the only school to demonstrate negative growth for minority students, most instruction (55 percent) was provided whole class followed by small group (22 percent). At School B, students mostly did independent work (52 percent) followed by small group (39 percent).

**Upper elementary**

**Tier 3 intervention.** The most surprising comparisons were found in upper elementary Tier 3 intervention, as shown in Figure 10. At School E, a balanced amount of time was spent between code-focused and meaning focused skills. Code focused skills included phonics (29 percent) and text reading (19 percent). Meaning focused skills included comprehension (32 percent) and grammar (10 percent). Surprisingly, observed intervention at School J nearly
mirrored that of School E in phonics (22 percent), text reading (17 percent), and comprehension (35 percent). In contrast, tier 3 upper elementary instruction at School B was largely focused on meaning focused skills which included comprehension (86 percent) followed by some vocabulary (10 percent). Figure 11 highlights the similarities and differences across the schools were also noticeable in grouping. Tier 3 grouping was mostly observed to occur in small groups in School E (83 percent) and School J (74 percent) relative to School B (41 percent).

**Tier 1 instruction.** A relatively large proportion of instructional time was dedicated to meaning focused skills, which can be seen in Figure 12. At School E, instruction as dedicated mostly to comprehension (48 percent). Unexpectedly, School E was also observed having the largest amount of non-instructional time, which accounted for 27 percent of their Tier 1 upper elementary instruction. At School J, observed instruction included comprehension (70 percent) followed by writing (14 percent) and grammar (eight percent). At School B, comprehension comprised 100 percent of observed instruction at Tier 1 for minority students in upper elementary. Across the schools, Tier 1 instruction in upper elementary was typically provided as whole class instruction (School E = 56 percent; School B = 70 percent; School J = 61 percent; see Figure 13).

As seen in the description of observations, there were many unexpected patterns unique to each of the schools. Instructional type and grouping varied across the schools and offered few plausible explanations or patterns of why minority student growth was stronger at Schools E and B relative to School J. Also, observations appeared to show more similarities than differences at the two schools that showed the strongest growth (School E) and least growth (School J) for minority students in Tier 3.
CHAPTER 4
DISCUSSION

Using the RTI framework which was introduced to help address disproportionate academic failure, and Critical Race Theory in Education, the goal of this study was to explore how minority students who were identified as most “at risk” in three schools made reading growth in relation to their peers. To add to the literature about supporting students through Tier 3, with a specific emphasis on support for minority learners, I conducted a study to address three research questions. First, relative to students who were only classified as receiving Tier 1 instruction, what was the growth of students who received Tier 3? Second, what was the growth of minority students classified as receiving Tier 3 compared to other students (including non-minority in Tier 3, non-minority in Tier 1, and minority in Tier 3)? Third, why might minority student growth differ across schools? In the following sections, I summarize the major findings from the study, discuss the limitations, and provide directions for future research.

Research Question 1- Relative to students who were only classified as receiving Tier 1 instruction, what is the growth of students who received Tier 3?

As summarized in Table 7, averaged across first through fourth, students who received Tier 1 only demonstrated mean growth at School B of 10.87 RIT points (SD = 0.59), at School E of 13.22 RIT points (SD = 0.61), and at School J of 12.77 RTI points (SD = 0.42). For students in Tier 3, regardless of minority status, two of the schools demonstrated encouraging findings for students in Tier 3. I found that in School E and School J, students who received Tier 3 intervention demonstrated patterns of greater growth than students who received Tier 1 instruction only (School B: $M = 2.07, SD = 0.84$; School J: $M = 2.65, SD = 1.42$), though it was
only significant at School B. Though, at School E, Tier 3 students overall demonstrated significantly less growth ($M = -2.34; SD = 1.08$; see Table 7).

Prior research, such as Al Otaiba et al. (2014), has shown a positive overall effects of Tier 3, relative to a control group ($ES = 0.36$). As a reminder, through a randomized control trial, she and colleagues found that students assigned to Tier 3 immediately showed improved reading performance relative to first grade classroom peers with the same initial skills but waited for Tier 3 until demonstrating inadequate growth in Tier 1 and Tier 2. Given the correlational design of my study, I cannot directly compare my findings to prior studies; nor can I compare my effect sizes to effect sizes of experimental studies. However, as shown in Table 7, though not as large as Al Otaiba et al., I do find overall small positive mean effect at two of the schools and uniquely extend the research base by exploring the potential benefit that Tier 3 may have. Also, though not directly comparable due to designs, my results indicate, unlike Balu et al’s (2015) findings about the effects of Tier 2 versus Tier 1 using RDD, that Tier 3 is helpful for students, though the benefit likely depends on factors unique to the school.

**Research Question 2: What is the growth of minority students classified as receiving Tier 3 compared to other students?**

**Research Question 3: Why might minority student growth differ across schools?**

Since the direction of the findings in research question 2 informed the case study in research question 3, I discuss their combined implications. Overall, there was variability in how minority students who received Tier 3 support demonstrated growth relative to their peers which included minority students in Tier 1, and non-minority students both in Tier 3, and Tier 1. Clear patterns emerged across the schools even though the sample size at each school limited the ability to detect significance of the results. A trend emerged at School E that, while overall
students in Tier 3 made less growth relative to Tier 3, minority students in Tier 3 actually made more growth than other students irrespective of tier status or minority status \((M = 5.16, SD = 3.48)\). Similarly, minority students in Tier 3 at School B also demonstrated more growth than other students \((M = 2.94; SD = 3.18)\). However, at School J, minority students in Tier 3 demonstrated less growth than all other students \((M = -8.20, SD = 4.97)\). It is possible that at School E and School B, policies in place and school staff may have done something differently from School J which helped to narrow the gap for minority students who are most at risk for reading failure.

Guided by my conceptual framework of CRT-E (Ladson-Billings & Tate, 1995), enacted the power of voice to explore how minority students in Tier 3 were supported. Using transcripts from principal interviews about RTI implementation and observations of minority students receiving Tier 3 and Tier 1 support, I then conducted a case study of each of the schools to explore plausible a-priori explanations for the patterns I detected. Each school was demographically similar in nearly all aspects, yet three themes emerged from the more in-depth analysis of each school. First, at Schools E and B where minority students demonstrated more growth, principals were able to clearly articulate how they used data to support students in each tier of intervention. Related to that, the specificity by which they outlined their support for students demonstrated a reliance on multiple formative sources of data and a team-based approach through teacher input to determine how to best support students, which is consistent with literature on DBI (CEC HP6, 2017; Jung et al., 2018). At School J, it was less clear how the school made decisions to support students. This careful attention to detail may be of particular benefit to minority students in Tier 3 as it may represent a focus on ensuring equitable outcomes.
for minority students (as specified in CRT-E; Ladson-Billings & Tate, 1995) rather than merely equal decision making.

An additional theme was related to how students were identified for tiered support. Researchers have highlighted the importance of early literacy interventions (e.g. Foorman et al., 1998; Snow, Burns, & Griffin, 1998; Jitendra et al., 2004; Denton, 2012) and advocates for the pyramid model have suggested that Tier 1 and 2 should be robust enough for all but 10% of children (Fuchs, Fuchs, and Compton, 2012). Yet, scholars have reported that in schools with high populations of minority students or those from low-SES backgrounds, the typical pyramid of RTI in which most students receive Tier 1 and far fewer students receive supplemental instruction in Tier 2 or intensive intervention in Tier 3, is actually inverted. In this study, Schools E and B most closely resembled this inverted pyramid, whereas School J more resembled the traditional structure of RTI. Given that each school served high populations of minority students, the consideration of having a higher percentage of students in Tier 3 than is traditionally deemed necessary may be especially important in helping to narrow the achievement gap for students of color. For minority students specifically, attempting to maintain a traditional RTI model with only a few students served in more intensive tiers of intervention, may do a disservice to minority students. However, in my study, though percentages of students in Tier 3 and special education were similar regardless of minority status, it also potentially suggests that too many students needed these services, thus representing the blurring of special education (Fuchs, Fuchs, & Stecker, 2010) in which Tier 3 services are provided as special education and vice versa.

The final theme was related to instruction and intervention in early elementary. As supported in prior research, the observations revealed clear differences between types and amounts of instruction provided in early elementary compared to late elementary. Similar to
Swanson et al. (2012) and Wanzek et al. (in review), students were observed using the ICE-R (Edmonds & Briggs, 2003) to document the amounts and types of reading instruction they received. However, unique to this study, I conducted an in-depth analysis of what minority students received in Tier 3 and Tier 1 to make comparisons between instruction in each tier, and to serve as a snapshot of how minority students were supported at the school level. Broadly, and similar to Swanson et al., instruction in Tier 3 included a balance of comprehension and phonics instruction. More narrowly, Tier 3, intervention incorporated a balanced combination of code-focused and meaning-focused skills both in early and late elementary at School E and School J. While this was true for early elementary at School B, Tier 3 in late elementary was heavily weighted toward comprehension. Across all schools and grades, Tier 3 instruction was mostly provided in small groups. Unlike Tier 3, Tier 1 was typically provided as whole group regardless of grade, and instruction leaned more toward comprehension. Early elementary Tier 1 instruction largely focused on comprehension at School E, whereas at School B and School J, much of the instructional time was split between comprehension and writing. Lastly, in late elementary, comprehension was the largest instructional category at all of the schools, and accounted for all observed instruction at School B. Interestingly, across the schools, there were far more similarities in observed instruction, especially at School E and School J which demonstrated opposing patterns for minority student reading growth. It was also perplexing that no clear patterns emerged that differentiated schools from each other. However, it was clear that in each school, observations rarely included instruction in vocabulary and phonemic awareness, as was similar in prior research (Baker, 2019).

Strengths, Limitations, and Directions for Future Research
This study included three main strengths. First, in reporting of effect sizes, I was able to unearth encouraging patterns indicating that depending on supports provided in schools, minority status may be related to student reading growth. This may provide insight in how the needs of minority students are supported. Second, my measures demonstrated good reliability data (MAP and ICE-R) maximizing ability to accurately assess the correlation.

Findings from this study should also be viewed in light of some limitations. First, by nature of correlational designs, I was unable to make causal comparisons between minority growth to non-minority growth. Due to having small samples at each tier, the study was underpowered to analyze the interactions between minority status and tier of intervention at each school. Third, all students were in one district. As such, I did not have the data to make comparisons of minority students in a district serving large populations of minority students versus a district serving fewer minority students. Future research that includes more districts using the same outcome assessments is needed to address these limitations.

Fourth, due to the data from the larger study, students in Tier 2 were not included in the analyses. Future research should consider including Tier 2 students to understand the progressions between and student outcomes in each tier. Fifth, observations were of varying lengths and occurred only once in each tier. Relatedly, the ICE-R (Edmonds & Briggs, 2003) protocol required observations of particular instruction and grouping to last at least a minute, therefore there might have been brief opportunities of instruction that were not accounted for (e.g. 48 seconds of “phonemic awareness” during a 5-minute instructional period coded as fluency) because they did not last for a full 60 seconds of instructional time. Therefore, they only reflected a snapshot of the supports students received. Multiple observations during entire reading blocks would add a more comprehensive view the instruction students receive at each
tier and would be more revelatory of how different schools include foundational components of
reading instruction.

My final limitations pertain to the principal interview RTI protocol and rating rubric. I was not able to assess cultural responsivity at the schools. The interview protocol was worded in such a way that principals mostly discussed linguistic supports for students whose first language was not English, rather than accounting for how other part of culture might affect instruction. Additionally, the ratings may have had limited sensitivity. Through the rubric clearly specified when to rate the interview as a 1, 3, or 5, it was not always clear when to rate a response as a gray rating of 2 or 4. Finally, there were no questions on the rubric related to observations of services provided in each tier. Therefore, future research should include the following three elements: 1) clear differentiation in the rubric between “gray” ratings, 2) questions related to observations, and 3) an addendum to the interview protocol specifying how schools included cultural responsivity and awareness within curriculum and teacher training - separate from linguistic supports- would help to highlight how the unique characteristics of students’ cultures are addressed.

**Conclusion**

The outcomes of American education have often demonstrated a crack in the foundation, or system, in which some succeed while others are left behind. As highlighted by countless statistics, those who benefit the least are most often students from minority backgrounds. If systemic inequities ingrained within the fabric of society are at the root of negative results that disproportionately affect minority students, as posited in CRT in education (Ladson-Billings & Tate, 1995; Dixson & Rousseau, 2005), researchers and educators must realize that the problem is not within students, but within the foundation of the system, and must push for change. Much
of the desired reforms that need to occur to rectify the frustrating outcomes are beyond the
capabilities of the schools alone yet, education is one of the foundational layers and it is
therefore important to explore what occurs at the school and district levels to meet the needs of
all. Given that Black and LatinX students are at a higher risk of failing to meet imposed
standards of reading proficiency, then researchers must continue to work to clearly articulate
ways to support this community of learners. As such, it is also necessary to continue to advocate
for change by advancing research and practices aimed toward narrowing the gap in reading
performance, and ultimately maximize the American ideals of equity and success for all.
REFERENCES

Note: An * indicates the study was one of the 17 studies in my review of the literature.


growth/.


The influence of curriculum-based measurement on education. U of Minnesota Press.


based individualization for students with intensive learning needs: A meta-

category/demographics-and-the-economy/people-in-poverty/.

validity and diagnostic accuracy of screening measures of reading skills. *Psychology in
the Schools*, 52(5), 500-514.


Kozleski, E. B., & Smith, A. (2009). The role of policy and systems change in creating equity for


Li, M; Koh, P.; Geva, E.; Joshi, R.; Chen, X. (2020). The componential model of reading in
bilingual learners, *Journal of Educational Psychology*.

https://doi.org/10.1037/edu0000459

*Liu, K. H. (2009). Development of an assessment rubric for the implementation of Response to
Intervention (RTI) at elementary schools* (Doctoral dissertation, University of Oregon).

dissertation.

Morgan, P. L., Farkas, G., Hillemeier, M. M., Mattison, R., Maczuga, S., Li, H., & Cook, M.
(2015). Minorities are disproportionately underrepresented in special education:


National Center on Intensive Intervention. (n.d.) https://intensiveintervention.org/


Shapiro, E. S. (2012). New thinking in response to intervention.


Voyager Sopris Learning (2008). *Passport*. Dallas, TX.


Appendix 1: Figures
Figure 1

Prisma Diagram of Literature Review

Records identified through data-base searching  
\( k = 974 \)

Duplicates removed  
\( k = 289 \)

Records screened (titles, abstracts)  
\( k = 699 \)

Full-text articles reviewed for eligibility  
\( k = 104 \)

Studies included in synthesis  
\( k = 17 \)

Records identified through other sources  
\( k = 14 \)

Records excluded  
\( k = 595 \)

Reasons for exclusion:
- Study either did not observe RTI practices
- Did not specify providing an intervention in Tier 1 (core) or Tier 3 (intensive)
  - Not elementary (K-5)
  - Not a study
  - Not reading
- Survey only of teachers’ practices (no intervention/observation)
  - Preservice teachers
  - Tier 2 only
- Described Tier 1 or Tier 3 only
- No observation component

Note. This figure demonstrates the flow of the search process taken to identify and screen articles for eligibility into the systematic literature review, summarized in Chapter 1.
Figure 2

Flowchart of Sample Selection

Starting Sample
\( n = 3274 \)

\( n = 2663 \)
- Delete 5th grade negative scores (\( n = 611 \))

\( n = 1348 \)
- Remove schools with less than four observations of minority students receiving Tier 3 (\( n = 1315 \))

\( n = 1059 \)
- Recode Tier into Tier 3 and Tier 1
- Remove missing tier data (\( n = 96 \))
- Remove Tier 2 (\( n = 193 \))

\( n = 1029 \)
- Compute upper/lower quartiles (\( n = 27 \))

Final Sample
\( n = 1002 \)

Note. This figure demonstrates the process taken to identify students and schools for eligibility into the study.
**Figure 3**

*Excerpt of NWEA MAP Norms*

<table>
<thead>
<tr>
<th>Grade</th>
<th>2015 READING Student Status Norms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Begin-Year</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>K</td>
<td>137.5</td>
</tr>
<tr>
<td>1</td>
<td>160.7</td>
</tr>
<tr>
<td>2</td>
<td>174.7</td>
</tr>
<tr>
<td>3</td>
<td>188.3</td>
</tr>
<tr>
<td>4</td>
<td>198.2</td>
</tr>
<tr>
<td>5</td>
<td>205.7</td>
</tr>
<tr>
<td>6</td>
<td>211.0</td>
</tr>
<tr>
<td>7</td>
<td>214.4</td>
</tr>
<tr>
<td>8</td>
<td>217.2</td>
</tr>
<tr>
<td>9</td>
<td>220.2</td>
</tr>
<tr>
<td>10</td>
<td>220.4</td>
</tr>
<tr>
<td>11</td>
<td>222.6</td>
</tr>
</tbody>
</table>

*Note.* This figure demonstrates the anticipated student growth on RIT scores in reading, published by the NWEA MAP (2015).
Figure 4

Excerpt of RTI Implementation Worksheet

RTI Essential Components Worksheet

School: ____________________________ District: ____________________________ Date: __________

Person(s) Interviewed: ____________________________

Interviewer(s): ____________________________

Purpose:

The purpose of this worksheet is to provide a tool for collecting relevant information and for recording a school’s rating on various items related to response to intervention (RTI) implementation. Descriptions of ratings for each item are provided on the RTI Essential Components Integrity Rubric (the Rubric).

Information about school-level implementation (Grades K–8) may be collected through interviews with school personnel and through observations and document review. After all of the information has been collected, use your notes and the Rubric to rate the school on each item. The Rubric provides a five-point rating scale and descriptions of practices that would score a 1, 3, or 5. Data collectors may assign the school a rating of 2 or 4 if the information collected suggests the school falls between the rubric descriptions. For example, if the reviewer judges a school to be performing at a level higher than the Rubric describes for a 3 rating but not quite at the level described for a 5, then the reviewer should rate the school as performing at a 4.
Figure 5

Excerpt from ICE-R Template
Figure 6

*Tier 3 Instruction at Lower Elementary*

![Graph showing Tier 3 Instruction at Lower Elementary](image)
NARROWING THE GAP FOR MINORITY LEARNERS

Figure 7
Tier 3 Grouping at Lower Elementary

Tier 3 Grouping- Lower Elem.

School B
School E
School J

0 0.2 0.4 0.6 0.8 1 1.2

T3 Tutoring  T3 Individual  T3 Independent  T3 Paired  T3 Small Group  T3 Whole Class
Figure 8

Tier 1 Instruction at Lower Elementary

Tier 1 Instruction Lower Elementary

- T1 Behav Man
- T1 Non-Instruction
- T1 OA
- T1 Writ
- T1 Gram
- T1 Text Read
- T1 Spelling
- T1 Comp
- T1 Vocab
- T1 Fluency
- T1 Phonics
- T1 PA

School B  School E  School J
Figure 9

*Tier 1 Grouping at Lower Elementary*

![Tier 1 Grouping Lower Elem.](image-url)
Figure 10

Tier 3 Instruction at Upper Elementary

Tier 3 Instruction- Upper Elem.

T3 Behav Man
T3 Non-Instruction
T3 OAI
T3 Writ
T3 Gram
T3 Text Read
T3 Spelling
T3 Comp
T3 Vocab
T3 Fluency
T3 Phonics
T3 PA

School B | School E | School J

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Figure 11

Tier 3 Grouping at Upper Elementary
Figure 12

Tier 1 Instruction at Upper Elementary

Tier 1 Instruction- Upper Elem.

- T1 Behav Man
- T1 Non-Instruction
- T1 OAI
- T1 Writ
- T1 Gram
- T1 Text Read
- T1 Spelling
- T1 Comp
- T1 Vocab
- T1 Fluency
- T1 Phonics
- T1 PA

[Bar chart showing Tier 1 Instruction at Upper Elementary for School B, School E, and School J]
Figure 13

*Tier 1 Grouping at Upper Elementary*

![Bar chart showing Tier 1 Grouping at Upper Elementary for School B, School E, and School J.](image-url)
Appendix B: Tables
### Table 1

#### Summary of Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Method</th>
<th>Setting</th>
<th>Student Participants</th>
<th>Student Diversity</th>
<th>Teacher Participants</th>
<th>Tier 3 Setting/Duration</th>
<th>Tier 3 curriculum/skills</th>
<th>Effect Size/Theme</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quantitative Experimental Designs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>O'Connor et al., 2005</em></td>
<td>NR</td>
<td>n = 2 Schools; SES; 45% FRL</td>
<td>n = 11 KG-3rd grade students</td>
<td>68% White (Total population across schools)</td>
<td>n = 3 Research staff (Tier 3)</td>
<td>20 weeks; 30 min/day; 5 days/week</td>
<td>Decoding, high frequency words, and fluency</td>
<td>ES ranged from 0.40-1.80</td>
</tr>
<tr>
<td><em>Denton et al., 2006</em></td>
<td>Quasi-Experimental Pre-post experimental design</td>
<td>n = 4 Schools; SES: NR; “highly diverse”</td>
<td>n = 27 1st-3rd grade students</td>
<td>52% Black, 22% Hispanic; 4% Asian; 22% White (Total population across schools)</td>
<td>n = 6 Research staff (Tier 3); n = 36 1st-3rd grade GenEd teachers (Tier 1)</td>
<td>16 weeks; Decoding = 2 hours/day for 8 weeks and Fluency= 1 hour/day for 8 weeks</td>
<td>Decoding and fluency</td>
<td>ES ranged from .84-1.53</td>
</tr>
<tr>
<td>Study</td>
<td>Method</td>
<td>Setting</td>
<td>Student Participants</td>
<td>Student Diversity</td>
<td>Teacher Participants</td>
<td>Tier 3 Setting/Duration</td>
<td>Tier 3 curriculum/skills</td>
<td>Effect Size/Theme</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------</td>
<td>--------------------------------</td>
<td>----------------------</td>
<td>------------------------------------------------------------------------------------</td>
<td>----------------------</td>
<td>-------------------------</td>
<td>---------------------------------------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td><em>Denton et al., 2013</em></td>
<td>RCT</td>
<td>Schools; SES: 89% FRL</td>
<td>n = 10</td>
<td>n = 72 2nd grade students</td>
<td>n = 6</td>
<td>24-26 weeks; 45 min/day</td>
<td>Word study, fluency, reading comprehension, applying reading strategies, and written expression; Fluency as needed</td>
<td>ES ranged from .34-.40 statistically significant for all measures but ORF</td>
</tr>
<tr>
<td><em>Al Otaiba, Connor, et al., 2014</em></td>
<td>RCT</td>
<td>Schools in first year of RTI; SES: 90% FRL</td>
<td>n = 7</td>
<td>n = 522 1st grade students</td>
<td>Research staff (Tier 3); n = 34 first-grade GenEd teachers</td>
<td>24 weeks; 45 min/day, 4 days/week</td>
<td>Dynamic RTI- students immediately placed in most appropriate tier</td>
<td>ES = .36</td>
</tr>
<tr>
<td>Study</td>
<td>Method</td>
<td>Setting</td>
<td>Student Participants</td>
<td>Student Diversity</td>
<td>Teacher Participants</td>
<td>Tier 3 Setting/Duration</td>
<td>Tier 3 curriculum/skills</td>
<td>Effect Size/Theme</td>
</tr>
<tr>
<td>--------------------</td>
<td>-------------------------------</td>
<td>------------------------------</td>
<td>----------------------</td>
<td>-------------------</td>
<td>----------------------</td>
<td>------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td><em>Al Otaiba, Kim, et al., 2014</em></td>
<td>Longitudinal follow-up study to RCT</td>
<td>*n = 6 Schools (originally 7 but 1 school opted out); SES:90% FRL</td>
<td><em>n = 278 in 3rd grade (after attrition)</em></td>
<td>58% Black (Total sample) 30% Black or other (in Tier 3)</td>
<td>Research staff (Tier 3); <em>n = 34 1st grade GenEd teachers</em></td>
<td>24 weeks; 45 min/day, 4 days/week</td>
<td>Phonics, phonemic awareness, fluency, alphabetic, and morphology.</td>
<td>ES ranged from -0.81 to -0.44 for SR students 3rd grade</td>
</tr>
<tr>
<td><em>Greulich et al., 2014</em></td>
<td>All-subset regression</td>
<td>*n = 7 Schools in first year of RTI; SES:90% FRL</td>
<td>*n = 20 1st grade who were inadequate responders during the Al Otaiba et al (2014) study</td>
<td>58% Black (Total sample) 30% Black or other (in Tier 3)</td>
<td>Research staff (Tier 3); <em>n = 34 1st grade GenEd teachers</em></td>
<td>24 weeks; 45 min/day, 4 days/week</td>
<td>Phonics, phonemic awareness, fluency, alphabetic, and morphology.</td>
<td>Non-responders were more off task and less engaged than adequate responders</td>
</tr>
<tr>
<td>Study</td>
<td>Method</td>
<td>Setting</td>
<td>Student Participants</td>
<td>Student Diversity</td>
<td>Teacher Participants</td>
<td>Tier 3 Setting/Duration</td>
<td>Tier 3 curriculum/skills</td>
<td>Effect Size/Theme</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>---------</td>
<td>----------------------</td>
<td>-------------------</td>
<td>----------------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td><em>Wanzek et al., (In review)</em></td>
<td>RCT</td>
<td>$n = 15$ Public and charter schools; SES:86% FRL</td>
<td>$n = 306$ 4th grade students from 2 cohorts; ($n = 154$ treatment)</td>
<td>36.9% Black; 29.1% Hispanic; 2.3% American Indian; 2.6% Asian or Pacific Islander; 24.2% White (Total sample)</td>
<td>$n = 26$ Research staff (Tier 3)</td>
<td>100 sessions; 45 min/day; 5 days/week</td>
<td>Phonics, word recognition, fluency, vocabulary, and comprehension</td>
<td>ES ranged from .08-.25</td>
</tr>
</tbody>
</table>

**Other Quantitative Designs**

<table>
<thead>
<tr>
<th>Study</th>
<th>Method</th>
<th>Setting</th>
<th>Student Participants</th>
<th>Student Diversity</th>
<th>Teacher Participants</th>
<th>Tier 3 Setting/Duration</th>
<th>Tier 3 curriculum/skills</th>
<th>Effect Size/Theme</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Vaughn et al., 2009</em></td>
<td>Quasi-Experimental; RD</td>
<td>$n = 7$</td>
<td>$n = 48$, 2nd grade students ($n = 14$ low responders)</td>
<td>$n = 14$; $n = 3$ Black; $n = 9$ Hispanic; $n = 2$ White (In Tier 3)</td>
<td>$n = 16$ Research staff (Tier 3)</td>
<td>26 weeks; 50 min/day; 5 days/week</td>
<td>Sound review, phonics, vocabulary, fluency, text reading, and comprehension</td>
<td>ES ranged from .23-1.22 on measures of reading</td>
</tr>
<tr>
<td>Study</td>
<td>Method</td>
<td>Setting</td>
<td>Student Participants</td>
<td>Student Diversity</td>
<td>Teacher Participants</td>
<td>Tier 3 Setting/Duration</td>
<td>Tier 3 curriculum/skills</td>
<td>Effect Size/Theme</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------</td>
<td>----------------------------------</td>
<td>----------------------</td>
<td>---------------------------</td>
<td>----------------------</td>
<td>-------------------------</td>
<td>-------------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>*Volpe et al., 2011</td>
<td>Single Case:</td>
<td>Urban, public school; SES:NR</td>
<td>$n = 1$</td>
<td>$n = 4$ Black; $n = 2$ Hispanic (In Tier 3)</td>
<td>$n = 4$ Research staff</td>
<td>Intervention included tutoring</td>
<td>Letter sounds using, <em>Buddy</em>, the researcher created computer program</td>
<td>ES = 2.87</td>
</tr>
<tr>
<td></td>
<td>Multiple baseline design</td>
<td></td>
<td>KG students at risk for reading difficulty; KG at risk for reading difficulty</td>
<td>$n = 4$ Black; $n = 2$ Hispanic (In Tier 3)</td>
<td>$n = 4$ Research staff</td>
<td>Intervention included tutoring</td>
<td>Letter sounds using, <em>Buddy</em>, the researcher created computer program</td>
<td>ES = 2.87</td>
</tr>
<tr>
<td>*Yurick et al., 2012</td>
<td>Correlation</td>
<td>Urban schools; SES:95% FRL</td>
<td>$n = 3$</td>
<td>$n = 70$ Black; 7.9% Latino; 5.2% multiracial; 39.5% White (In Tier 3)</td>
<td>$n = 6$ School-based KG instructional assistants.</td>
<td>30 min/day, 3-5 days/week (depending on school)</td>
<td>Phonics, phonological awareness, word reading, and alphabetic principle</td>
<td>ES = .81 and .99</td>
</tr>
<tr>
<td>Liu, 2009</td>
<td>Mixed methods dissertation</td>
<td>Schools; SES:NR</td>
<td>$n = 2$</td>
<td>$n = 5$ 2nd grade students who received Tier 3</td>
<td>$n = 5$ School professional $s (n = 3$ GenEd teachers; $n = 1$ SpEd teacher; $n = 1$ instructional assistant)</td>
<td>30 min/day, 3 days/week (School 1); 30 min/day, 5 days/week (School 2)</td>
<td>Phonics, fluency, comprehension</td>
<td>Rubric was able to capture that School 1 fully implemented RTI, and School 2 only partially implemented RTI</td>
</tr>
</tbody>
</table>

**Qualitative Designs**

Liu, 2009: Mixed methods dissertation — $n = 2$ Schools; SES:NR

- 2nd grade students who received Tier 3
- 100% White (In Tier 3)
- $n = 5$ School professionals ($n = 3$ GenEd teachers; $n = 1$ SpEd teacher; $n = 1$ instructional assistant)
- 30 min/day, 3 days/week (School 1); 30 min/day, 5 days/week (School 2)
- Phonics, fluency, comprehension

Rubric was able to capture that School 1 fully implemented RTI, and School 2 only partially implemented RTI
<table>
<thead>
<tr>
<th>Study</th>
<th>Method</th>
<th>Setting</th>
<th>Student Participants</th>
<th>Student Diversity</th>
<th>Teacher Participants</th>
<th>Tier 3 Setting/Duration</th>
<th>Tier 3 curriculum/skills</th>
<th>Effect Size/Theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Orosco &amp; Klingner, 2010</td>
<td>Case study</td>
<td>n = 1 Large urban school; SES:99% FRL</td>
<td>KG-2nd grade majority Latino students (amount NS)</td>
<td>85% Latino (Total population)</td>
<td>n = 6 teachers, school-based (SpEd provided Tier 3)</td>
<td>15 – 20 min/session Tier 3 was SpEd; Intensive, individualized</td>
<td>NR; evidence based interventions</td>
<td>Instruction in RTI was not tailored to meet students’ linguistic needs</td>
</tr>
<tr>
<td>*Rinali et al., 2011</td>
<td>Case study</td>
<td>n = 1 Urban school; SES:NR</td>
<td>NR</td>
<td>59% Hispanic, 16% Black, 13% -Asian, 1% White, and 1% - other. 39% ELL (Total population)</td>
<td>n = 8 School based professions (4 GenEd teachers, 1 reading specialist, 3 SpEd teachers)</td>
<td>20 min of 1-1 Literacy support in addition to Tier 1 and Tier 2</td>
<td>NR</td>
<td>Implemented RTI model led to more data usage and less referrals to special education</td>
</tr>
<tr>
<td>Shepherd &amp; Salembier, 2011</td>
<td>Case study</td>
<td>n = 3 Rural schools in their first year of RTI; SES:47% FRL</td>
<td>NR</td>
<td>NR</td>
<td>n = 43 School based teacher including GenEd/SpEd/support teachers</td>
<td>Tier 3 was SpEd</td>
<td>NR</td>
<td>RTI led to greater data usage and greater collaboration between teachers</td>
</tr>
<tr>
<td>Study</td>
<td>Method</td>
<td>Setting</td>
<td>Student Participants</td>
<td>Student Diversity</td>
<td>Teacher Participants</td>
<td>Tier 3 Setting/Duration</td>
<td>Tier 3 curriculum/skills</td>
<td>Effect Size/Theme</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>---------</td>
<td>----------------------</td>
<td>-------------------</td>
<td>---------------------</td>
<td>-------------------------</td>
<td>------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>*Swanson et al., 2012</td>
<td>Not named- (likely; Case Study)</td>
<td>School district</td>
<td>NR</td>
<td>23% Black; 36% Hispanic; 9% Asian or Pacific Islanders; 32% White (Total population)</td>
<td>Tier 3= Year 1: n = 17 SpEd teachers; Year 2: n = 12 SpEd teachers</td>
<td>Provided outside of the GenEd classroom</td>
<td>Comprehension, instruction, word study, vocabulary, text reading, and fluency;</td>
<td>Teachers described RTI as difficult due to additional paperwork and time needed to implement</td>
</tr>
<tr>
<td>AlSuli-man, 2012</td>
<td>Mixed methods dissertation</td>
<td>n = 2 Schools; SES:20% FRL</td>
<td>n = 26 4th grade students receiving RTI services (4 students received Tier 3)</td>
<td>15% Black or other (Total population across schools)</td>
<td>n = 6 School staff (n = 2 SpEd teachers; n = 2 reading specialists; n = 2 GenEd teachers); Tier 3 by SpEd teachers</td>
<td>Students received intervention in small group settings.</td>
<td>Read Naturally (Ihnot, Matsoff, Gavin, &amp; Hendrickson, 2001) target individual need Triumph (McGraw-Hill, 2005) - support key skills found in the core curriculum</td>
<td>Tier 3 intervention was minimally beneficial for 3 of the 4 students in Tier 3</td>
</tr>
<tr>
<td>Study</td>
<td>Method</td>
<td>Setting</td>
<td>Student Participants</td>
<td>Student Diversity</td>
<td>Teacher Participants</td>
<td>Tier 3 Setting/Duration</td>
<td>Tier 3 Curriculum/ Skills</td>
<td>Effect Size/ Theme</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------------------</td>
<td>-----------------------</td>
<td>----------------------</td>
<td>-------------------</td>
<td>----------------------</td>
<td>--------------------------</td>
<td>--------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>McGlohorn, 2018</td>
<td>Mixed methods dissertation</td>
<td>n = 1 School; SES: diverse, NR</td>
<td>n = 80 KG-2nd grade who received Tier 3 instruction</td>
<td>23% Black; 12% Hispanic; 3% multiracial; 62% White (Total population)</td>
<td>n = 21 Teachers (n = 18 GenEd; n = 3 support staff); GenEd provided Tier 3</td>
<td>GenEd classroom Fountas and Pinnel (2010)-direct instruction and close reading; Project Read (2018) - phonological and phonemic awareness, letter sounds, concepts of print, vocabulary, and text reading</td>
<td>Tier 3 focused on students’ areas of need increased reading skills</td>
<td></td>
</tr>
</tbody>
</table>

Note. Studies denoted with an asterisk* were described in the current study. FRL = Free or reduced lunch; NR= Not reported; ES = Effect Size; RD = Regression Discontinuity; RCT= Randomized Control Trial; SpEd = Special Education; GenEd= General Education.
Table 2

**Tier 3 across studies**

<table>
<thead>
<tr>
<th>Study</th>
<th>Intensive Intervention</th>
<th>Individualized</th>
<th>Progress Monitoring</th>
<th>Small Group</th>
<th>Small Group Size</th>
<th>Intervention Frequency</th>
<th>Practice/Feedback Described</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>O’Connor et al., 2005</em></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>$n = 1-2$ Students per group</td>
<td>$n = 30$ min/day, 5 days/week</td>
<td>Yes</td>
</tr>
<tr>
<td><em>Denton et al., 2006</em></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>$n = 2$ Students per group</td>
<td>$n = 60$-120 min/day, 5 days/week</td>
<td>Yes</td>
</tr>
<tr>
<td><em>Denton et al., 2013</em></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>$n = 2-3$ Students per group</td>
<td>$n = 45$ min/day, 5 days/week</td>
<td>Yes</td>
</tr>
<tr>
<td><em>Al Otaiba, Connor, et al., 2014</em></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>$n = 1-3$ Students per group</td>
<td>$n = 45$ min/day, 4 days/week</td>
<td>NA</td>
</tr>
<tr>
<td><em>Al Otaiba, Kim, et al., 2014</em></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>$n = 1-3$ Students per group</td>
<td>$n = 45$ min/day, 4 days/week</td>
<td>NA</td>
</tr>
<tr>
<td><em>Greulich et al., 2014</em></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>$n = 1-3$ Students per group</td>
<td>$n = 45$ min/day, 4 days/week</td>
<td>NA</td>
</tr>
<tr>
<td><em>Wanzek et al., (In review)</em></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>$n = 2-3$ Students per group</td>
<td>$n = 45$ min/day, 5 days/week</td>
<td>Yes</td>
</tr>
<tr>
<td>Study</td>
<td>Intensive Intervention</td>
<td>Individualized</td>
<td>Progress Monitoring</td>
<td>Small Group</td>
<td>Small Group Size</td>
<td>Intervention Frequency</td>
<td>Practice/Feedback Described</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------------</td>
<td>----------------</td>
<td>--------------------</td>
<td>--------------</td>
<td>------------------</td>
<td>------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>*Vaughn et al., 2009</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>n = 2-4 students per group</td>
<td>50 min/day, 5 days/week</td>
<td>Yes</td>
</tr>
<tr>
<td>*Volpe et al., 2011</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>n = 1 student per group</td>
<td>4-6 min/day, 3 x per week</td>
<td>NA</td>
</tr>
<tr>
<td>*Yurick et al., 2012</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Other Quantitative Designs**

**Qualitative Designs**

<table>
<thead>
<tr>
<th>Liu, 2009</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>n = 1-2 students per group</th>
<th>School 1: n = 25 min, 2 days/week for 25 min plus Tier 2 for 30 min/day, 5 days/week; School 2: n = 60 min, 5 days/week; replaced Tier 1</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Orosco &amp; Klingner, 2010</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Study</td>
<td>Intensive Intervention</td>
<td>Individualized</td>
<td>Progress Monitoring</td>
<td>Small Group</td>
<td>Small Group Size</td>
<td>Intervention Frequency</td>
<td>Practice/Feedback Described</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------------</td>
<td>----------------</td>
<td>--------------------</td>
<td>--------------</td>
<td>------------------</td>
<td>------------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>*Rinaldi et al., 2011</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>( n = 1 ) student per group</td>
<td>( n = 20 ) min/day, 4-5 days/week</td>
<td>NA</td>
</tr>
<tr>
<td>Shepherd &amp; Salembier, 2011</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>ALSuliman, 2012</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>( n = 1-6 )</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>*Swanson et al., 2012</td>
<td>Yes</td>
<td>Yes (based on description of RTI)</td>
<td>Yes (based on description of RTI)</td>
<td>Yes</td>
<td>( n = 2-4 ) students per group</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>McGlohorn, 2018</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>( n = 2-4 ) students per group</td>
<td>( n = 30 ) min/day, 5 days/week</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Note. Studies denoted with an asterisk* were described in the current study.
### Table 3

**Tier 1 across studies**

<table>
<thead>
<tr>
<th>Study</th>
<th>Tier 1- Universal Screening</th>
<th>Screening Frequency</th>
<th>Curriculum Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>O'Connor et al., 2005</em></td>
<td>Yes</td>
<td>3x per year</td>
<td>Ongoing PD for GenEd teachers addressing skills in the NRP report plus PD for phonological blending and segmenting, letter names and sounds, the alphabetic principle, strategies for phonics instruction and integrating phonics with phonemic awareness, and vocabulary</td>
</tr>
<tr>
<td><em>Denton et al., 2006</em></td>
<td>Yes (see Mathes 2005)</td>
<td>ND</td>
<td>Unnamed basal curriculum that &quot;provided guidance for delivering a comprehensive reading curriculum&quot;; PD for GenEd teachers</td>
</tr>
<tr>
<td><em>Denton et al., 2013</em></td>
<td>Yes</td>
<td>ND- indicates at least 2x per year</td>
<td>Unnamed basal curriculum described as providing explicit instruction in phonics, word study, and comprehension</td>
</tr>
<tr>
<td><em>Al Otaiba, Connor, et al., 2014</em></td>
<td>Yes</td>
<td>ND- indicates at least 4x per year</td>
<td><em>Open Court</em> (McGraw Hill, 2002) 90 min/day. Focused on phonics and reading comprehension; 2 days of PD for GenEd teachers</td>
</tr>
<tr>
<td><em>Al Otaiba, Kim, et al., 2014</em></td>
<td>Yes</td>
<td>ND- indicates at least 4x per year</td>
<td><em>Open Court</em> (McGraw Hill, 2002) 90 min/day. Focused on phonics and reading comprehension; 2 days of PD for GenEd teachers</td>
</tr>
</tbody>
</table>
**NARROWING THE GAP FOR MINORITY LEARNERS**

<table>
<thead>
<tr>
<th>Study</th>
<th>Tier 1- Universal Screening</th>
<th>Screening Frequency</th>
<th>Curriculum Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Greulich et al., 2014</td>
<td>Yes</td>
<td>ND- indicates at least 2x per year</td>
<td><em>Open Court</em> (McGraw Hill, 2002) core reading curriculum for 90 min/day. Focused on phonics and reading comprehension; 2 days of PD for GenEd teachers</td>
</tr>
<tr>
<td>*Wanzek et al., In review</td>
<td>Yes</td>
<td>ND</td>
<td><em>Journeys Common Core</em> (Templeton et al., 2014)</td>
</tr>
</tbody>
</table>

**Other Quantitative Designs**

<table>
<thead>
<tr>
<th>Study</th>
<th>Tier 1- Universal Screening</th>
<th>Screening Frequency</th>
<th>Curriculum Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Vaughn et al., 2009</td>
<td>Yes</td>
<td>3x per year</td>
<td>Not described</td>
</tr>
<tr>
<td>*Volpe et al., 2011</td>
<td>NA</td>
<td>ND</td>
<td>Not described- but included <em>K-PALS</em> (Fuchs, et al., 2001)</td>
</tr>
<tr>
<td>*Yurick et al., 2012</td>
<td>Yes</td>
<td>2x per year (fall and spring)</td>
<td>School 1: <em>LACES</em> (school designed); School 2: <em>Trophies</em> (Beck, Farr, &amp; Strickland, 2003); School 3: Not described</td>
</tr>
</tbody>
</table>

**Qualitative Designs**

<table>
<thead>
<tr>
<th>Study</th>
<th>Tier 1- Universal Screening</th>
<th>Screening Frequency</th>
<th>Curriculum Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Liu, 2009</td>
<td>Yes</td>
<td>3x per year</td>
<td><em>Treasures</em> (McGraw-Hill, 2005)</td>
</tr>
<tr>
<td>*Orosco &amp; Klingner, 2010</td>
<td>Yes</td>
<td>ND- indicates at least 1x/year</td>
<td>Unnamed curriculum. 90-min per day covering phonemic awareness, vocabulary, and fluency</td>
</tr>
<tr>
<td>*Rinaldi et al., 2011</td>
<td>Yes</td>
<td>3x per year</td>
<td>Unnamed curriculum; provided instruction in PA, phonics, alphabetic principle, fluency, comprehension, text reading, and writing</td>
</tr>
<tr>
<td>Study</td>
<td>Tier 1- Universal Screening</td>
<td>Screening Frequency</td>
<td>Curriculum Description</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-----------------------------</td>
<td>---------------------</td>
<td>----------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>*Shepherd &amp; Salembier, 2011</td>
<td>Yes</td>
<td>3x per year</td>
<td>Unnamed curricula but specifically stated that chosen curricula had to be scientifically-based programs recognized by the WWC</td>
</tr>
<tr>
<td>ALSuliman, 2012</td>
<td>Yes</td>
<td>3x per year</td>
<td>Treasures (McGraw-Hill, 2005)</td>
</tr>
<tr>
<td>*Swanson et al., 2012</td>
<td>Yes (based on description of RTI)</td>
<td>ND</td>
<td>Unnamed. Included 90 min per day reading instruction. Focus of instruction not detailed.</td>
</tr>
<tr>
<td>McGlohorn, 2018</td>
<td>Yes</td>
<td>3x per year</td>
<td>ND</td>
</tr>
</tbody>
</table>

*Note. Studies denoted with an asterisk* were described in the current study. PA = phonemic awareness.
Table 4

**Quality Indicators in quantitative studies**

<table>
<thead>
<tr>
<th>Study</th>
<th>Regular observations throughout the study</th>
<th>Inter-observer reliability reported</th>
<th>Field notes collected</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quantitative Experimental Designs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O’Connor et al., 2005</td>
<td>$n = 3$ observations in Tier 1; $n = 2$ observations in Tier 3</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>*Denton et al., 2006</td>
<td>$n = 5x$ in each intervention; $n = 90$ in Tier 1</td>
<td>Yes- 100% on all but one with was 94%</td>
<td>NS</td>
</tr>
<tr>
<td>*Denton et al., 2013</td>
<td>$n = 5x$ in each intervention</td>
<td>Yes- 85%</td>
<td>NS</td>
</tr>
<tr>
<td>*Al Otaiba, Connor, et al., 2014</td>
<td>$n = 3$ observations of each interventionist</td>
<td>Yes- 98.1%</td>
<td>NS</td>
</tr>
<tr>
<td>*Al Otaiba, Kim, et al., 2014</td>
<td>$n = 3$ observations in Tier 3; $n = 2$ observations in Tier 1</td>
<td>Yes- 98.1%</td>
<td>NS</td>
</tr>
<tr>
<td>*Greulich et al., 2014</td>
<td>$n = 3$ observations in Tier 3; $n = 2$ observations in Tier 1</td>
<td>NR</td>
<td>Yes</td>
</tr>
<tr>
<td>*Wanzek et al., (In review)</td>
<td>$n = 4$ for each interventionist in Tier 3; $n = 2$ in Tier 1</td>
<td>Yes - 90%</td>
<td>NS</td>
</tr>
<tr>
<td><strong>Other Quantitative Designs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*Vaughn et al., 2009</td>
<td>$n = 1x$ a month in Tier 3</td>
<td>Yes - 92%</td>
<td>NS</td>
</tr>
<tr>
<td>*Volpe et al., 2011</td>
<td>NS</td>
<td>NR</td>
<td>NS</td>
</tr>
<tr>
<td>Study</td>
<td>Regular observations throughout the study</td>
<td>Inter-observer reliability reported</td>
<td>Field notes collected</td>
</tr>
<tr>
<td>------------------------</td>
<td>-------------------------------------------</td>
<td>-----------------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>*Yurick et al., 2012</td>
<td>$n = 1 - 2$ per week during intervention (unclear how many weeks)</td>
<td>NR</td>
<td>NS</td>
</tr>
</tbody>
</table>
### Table 5

**Number of Students in Each Tier by Minority Status and Disability Status**

<table>
<thead>
<tr>
<th>Minority Status</th>
<th>Disability Status</th>
<th>School B</th>
<th>School E</th>
<th>School J</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Tier 1</td>
<td>Tier 1</td>
<td>Tier 1</td>
</tr>
<tr>
<td>White</td>
<td>Disability = No</td>
<td>20</td>
<td>29</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>% 71.4%</td>
<td>% 85.3%</td>
<td>% 89.7%</td>
</tr>
<tr>
<td></td>
<td>Disability = Yes</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>% 25.0%</td>
<td>% 0.0%</td>
<td>% 100.0%</td>
</tr>
<tr>
<td>Black, Hispanic</td>
<td>Disability = No</td>
<td>152</td>
<td>140</td>
<td>303</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>% 49.8%</td>
<td>% 68.6%</td>
<td>% 91.5%</td>
</tr>
<tr>
<td></td>
<td>Disability = Yes</td>
<td>4</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>% 26.7%</td>
<td>% 0.0%</td>
<td>% 78.3%</td>
</tr>
<tr>
<td>Total n</td>
<td></td>
<td>177</td>
<td>169</td>
<td>359</td>
</tr>
<tr>
<td>Total %</td>
<td></td>
<td>50.3%</td>
<td>66.5%</td>
<td>90.1%</td>
</tr>
<tr>
<td>Tier 3</td>
<td>Disability = No</td>
<td>8</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>% 28.6%</td>
<td>% 14.7%</td>
<td>% 10.3%</td>
</tr>
<tr>
<td></td>
<td>Disability = Yes</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>% 75.0%</td>
<td>% 100.0%</td>
<td>% 0.0%</td>
</tr>
<tr>
<td>Total n</td>
<td></td>
<td>153</td>
<td>64</td>
<td>28</td>
</tr>
<tr>
<td>Total %</td>
<td></td>
<td>49.7%</td>
<td>31.4%</td>
<td>8.5%</td>
</tr>
<tr>
<td>Total n</td>
<td></td>
<td>175</td>
<td>85</td>
<td>37</td>
</tr>
<tr>
<td>Total %</td>
<td></td>
<td>49.7%</td>
<td>33.5%</td>
<td>9.9%</td>
</tr>
<tr>
<td>Total n</td>
<td></td>
<td>352</td>
<td>254</td>
<td>396</td>
</tr>
<tr>
<td>Total %</td>
<td></td>
<td>100.0%</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>
Table 6

School Demographic Data

<table>
<thead>
<tr>
<th>School</th>
<th>Student Enrollment</th>
<th>FARL</th>
<th>Caucasian</th>
<th>African American</th>
<th>LatinX</th>
<th>Asian</th>
<th>American Indian</th>
<th>Pacific Islander</th>
<th>&gt;1 Race</th>
</tr>
</thead>
<tbody>
<tr>
<td>School B</td>
<td>816</td>
<td>85.00%</td>
<td>9.40%</td>
<td>8.70%</td>
<td>75.70%</td>
<td>1.30%</td>
<td>3.20%</td>
<td>0.00%</td>
<td>1.60%</td>
</tr>
<tr>
<td>School E</td>
<td>697</td>
<td>78.30%</td>
<td>13.50%</td>
<td>5.30%</td>
<td>74.90%</td>
<td>&lt; 1%</td>
<td>3.20%</td>
<td>&lt; 1%</td>
<td>2.20%</td>
</tr>
<tr>
<td>School J</td>
<td>745</td>
<td>89.50%</td>
<td>10.30%</td>
<td>5.00%</td>
<td>77.20%</td>
<td>1.30%</td>
<td>6.00%</td>
<td>0.00%</td>
<td>&lt; 1%</td>
</tr>
</tbody>
</table>
Table 7

Results of Research Questions 1 and 2

<table>
<thead>
<tr>
<th>School</th>
<th>Model</th>
<th>Variable</th>
<th>B</th>
<th>S.E. B</th>
<th>β</th>
<th>t</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>School B</td>
<td>Model 1</td>
<td>(Constant)</td>
<td>10.761</td>
<td>0.587</td>
<td>18.341</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tier 3</td>
<td>(Constant)</td>
<td>2.068</td>
<td>0.84</td>
<td>0.129</td>
<td>2.461</td>
<td>0.014*</td>
<td>0</td>
</tr>
<tr>
<td>School E</td>
<td>Model 2</td>
<td>(Constant)</td>
<td>13.215</td>
<td>0.61</td>
<td>21.65</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tier 3</td>
<td>(Constant)</td>
<td>-2.343</td>
<td>1.076</td>
<td>-0.133</td>
<td>-2.179</td>
<td>0.030*</td>
<td>0</td>
</tr>
<tr>
<td>School J</td>
<td>Model 3</td>
<td>(Constant)</td>
<td>12.774</td>
<td>0.419</td>
<td>30.485</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tier 3</td>
<td>(Constant)</td>
<td>2.654</td>
<td>1.422</td>
<td>0.093</td>
<td>1.867</td>
<td>0.063*</td>
<td>0</td>
</tr>
<tr>
<td>School B</td>
<td>Model 4</td>
<td>(Constant)</td>
<td>13.179</td>
<td>1.29</td>
<td>10.214</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Minority</td>
<td>(Constant)</td>
<td>-1.339</td>
<td>1.351</td>
<td>-0.047</td>
<td>-0.991</td>
<td>0.322</td>
<td>0</td>
</tr>
<tr>
<td>School E</td>
<td>Model 5</td>
<td>(Constant)</td>
<td>11.836</td>
<td>1.13</td>
<td>10.473</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Minority</td>
<td>(Constant)</td>
<td>0.689</td>
<td>1.221</td>
<td>0.029</td>
<td>0.565</td>
<td>0.573</td>
<td>0</td>
</tr>
<tr>
<td>School J</td>
<td>Model 6</td>
<td>(Constant)</td>
<td>14.048</td>
<td>1.228</td>
<td>11.436</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Minority</td>
<td>(Constant)</td>
<td>-0.991</td>
<td>1.296</td>
<td>-0.038</td>
<td>-0.765</td>
<td>0.445</td>
<td>0</td>
</tr>
<tr>
<td>School B</td>
<td>Model 7</td>
<td>(Constant)</td>
<td>11.129</td>
<td>1.453</td>
<td>7.659</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tier 3</td>
<td>(Constant)</td>
<td>2.301</td>
<td>0.866</td>
<td>0.144</td>
<td>2.657</td>
<td>0.008*</td>
<td>0</td>
</tr>
<tr>
<td>Minority</td>
<td>(Constant)</td>
<td>-0.663</td>
<td>1.504</td>
<td>-0.024</td>
<td>-0.441</td>
<td>0.66</td>
<td>0</td>
</tr>
<tr>
<td>School E</td>
<td>Model 8</td>
<td>(Constant)</td>
<td>11.323</td>
<td>1.372</td>
<td>8.252</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tier 3</td>
<td>(Constant)</td>
<td>-2.369</td>
<td>1.108</td>
<td>-0.135</td>
<td>-2.137</td>
<td>0.034*</td>
<td>0</td>
</tr>
<tr>
<td>Minority</td>
<td>(Constant)</td>
<td>1.984</td>
<td>1.471</td>
<td>0.085</td>
<td>1.349</td>
<td>0.179</td>
<td>0</td>
</tr>
<tr>
<td>School J</td>
<td>Model 9</td>
<td>(Constant)</td>
<td>14.019</td>
<td>1.256</td>
<td>11.159</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tier 3</td>
<td>(Constant)</td>
<td>3.076</td>
<td>1.423</td>
<td>0.11</td>
<td>2.162</td>
<td>0.031*</td>
<td>0</td>
</tr>
<tr>
<td>Minority</td>
<td>(Constant)</td>
<td>-1.234</td>
<td>1.323</td>
<td>-0.048</td>
<td>-0.932</td>
<td>0.352</td>
<td>0</td>
</tr>
<tr>
<td>School B</td>
<td>Model 10</td>
<td>(Constant)</td>
<td>12.31</td>
<td>1.733</td>
<td>6.925</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tier 3</td>
<td>(Constant)</td>
<td>-0.4</td>
<td>3.051</td>
<td>-0.025</td>
<td>-0.131</td>
<td>0.896</td>
<td>0</td>
</tr>
<tr>
<td>Minority</td>
<td>(Constant)</td>
<td>-1.654</td>
<td>1.848</td>
<td>-0.06</td>
<td>-0.895</td>
<td>0.371</td>
<td>0</td>
</tr>
<tr>
<td>School E</td>
<td>Model 11</td>
<td>(Constant)</td>
<td>2.938</td>
<td>3.182</td>
<td>0.183</td>
<td>0.923</td>
<td>0.356</td>
</tr>
<tr>
<td>Interaction</td>
<td>(Constant)</td>
<td>12.31</td>
<td>1.522</td>
<td>8.086</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>School J</td>
<td>Model 12</td>
<td>(Constant)</td>
<td>13.459</td>
<td>1.299</td>
<td>10.364</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tier 3</td>
<td>(Constant)</td>
<td>-6.935</td>
<td>3.274</td>
<td>-0.395</td>
<td>-2.118</td>
<td>0.035*</td>
<td>0</td>
</tr>
<tr>
<td>Minority</td>
<td>(Constant)</td>
<td>0.79</td>
<td>1.674</td>
<td>0.034</td>
<td>0.472</td>
<td>0.637</td>
<td>0</td>
</tr>
<tr>
<td>School J</td>
<td>Model 13</td>
<td>(Constant)</td>
<td>5.155</td>
<td>3.478</td>
<td>0.285</td>
<td>1.482</td>
<td>0.14</td>
</tr>
<tr>
<td>Interaction</td>
<td>(Constant)</td>
<td>13.459</td>
<td>1.299</td>
<td>10.364</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>School J</td>
<td>Model 14</td>
<td>(Constant)</td>
<td>10.541</td>
<td>4.742</td>
<td>0.378</td>
<td>2.223</td>
<td>0.027*</td>
</tr>
<tr>
<td>Minority</td>
<td>(Constant)</td>
<td>-0.607</td>
<td>1.374</td>
<td>-0.023</td>
<td>-0.442</td>
<td>0.659</td>
<td>0</td>
</tr>
<tr>
<td>Interaction</td>
<td>(Constant)</td>
<td>-8.199</td>
<td>4.97</td>
<td>-0.282</td>
<td>-1.65</td>
<td>0.100*</td>
<td>0</td>
</tr>
</tbody>
</table>

Note. Dependent variable = reading growth. * = p < .05.
## Table 8

*RTI Implementation Rubric Scores*

<table>
<thead>
<tr>
<th>Assessment</th>
<th>DBDM</th>
<th>Multi-Level Instruction</th>
<th>Infrastructure and Support</th>
<th>Fidelity and Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Screening</td>
<td>PM</td>
<td>Tier 1</td>
<td>Tier 2</td>
</tr>
<tr>
<td>School B</td>
<td>5</td>
<td>3.5</td>
<td>3.67</td>
<td>4.4</td>
</tr>
<tr>
<td>School E</td>
<td>5</td>
<td>3</td>
<td>4.67</td>
<td>4.4</td>
</tr>
<tr>
<td>School J</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4.2</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>4.67</td>
<td>3.17</td>
<td>3.78</td>
<td>4.33</td>
</tr>
<tr>
<td><strong>Variability</strong></td>
<td>0.33</td>
<td>0.08</td>
<td>0.71</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>Confidence Band</strong></td>
<td>0.14</td>
<td>0.04</td>
<td>0.31</td>
<td>0.01</td>
</tr>
</tbody>
</table>

*Note.* Assessment contained two levels: screening and PM. PM = Progress monitoring. Multi-Level Instruction contained 3 levels: Tier 1, Tier 2, and Tier 3.
Table 9

<table>
<thead>
<tr>
<th>School</th>
<th>Disability = no</th>
<th></th>
<th>Disability = yes</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n )</td>
<td>%</td>
<td>( n )</td>
<td>%</td>
</tr>
<tr>
<td>School B</td>
<td>333</td>
<td>94.6%</td>
<td>19</td>
<td>5.4%</td>
</tr>
<tr>
<td>((n = 352))</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School E</td>
<td>238</td>
<td>91.5%</td>
<td>16</td>
<td>6.3%</td>
</tr>
<tr>
<td>((n = 254))</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J Haley</td>
<td>370</td>
<td>93.4%</td>
<td>26</td>
<td>6.6%</td>
</tr>
<tr>
<td>((n = 396))</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>