AN EXAMINATION OF THIRD GRADE READING ACHIEVEMENT FOR STUDENTS IDENTIFIED AS SLD, TIER 2, OR TIER 3

by

KIRA DANIELLE FEE

A DISSERTATION

Presented to the Department of Educational Methodology, Policy, and Leadership
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Education

March 2020

DISSERTATION APPROVAL PAGE

Student: Kira Danielle Fee

Title: An Examination of Third Grade Reading Achievement for Students Identified as SLD, Tier 2, or Tier 3

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Education degree in the Department of Educational Methodology, Policy, and Leadership by:

Gerald Tindal, Ph.D.	Chairperson and Advisor
Keith Hollenbeck, Ph.D.	Core Member
Julie Alonzo, Ph.D.	Core Member
Laura Lee McIntyre, Ph.D.	Institutional Representative

and

Kate Mondloch	Interim Vice Provost and Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded March 2020

DISSERTATION ABSTRACT

Kira Danielle Fee

Doctor of Education

Department of Educational Methodology, Policy, and Leadership

March 2020

Title: An Examination of Third Grade Reading Achievement for Students Identified as SLD, Tier 2, or Tier 3

The purpose of this study was to examine formative reading and vocabulary achievement results for third grade students recently found eligible for special education (SpEd) within specific learning disability (SLD) reading eligibility versus other third grade students not identified with SLD. The participant sample crossed four consecutive school years, and students were assigned to four different groups according to SLD status and Fall overall easyCBM© risk including: (a) students not in SpEd with a Fall designation of Low Reading Risk Group (non-SpEd low risk), (b) students not in SpEd with a Fall designation of Some Reading Risk Group (non-sped some risk), (c) students not in special education with a Fall designation of High Reading Risk Group (non-SpEd high risk), and (d) students in special education with a Fall designation of High Reading Risk Group (SLD high risk). The SLD high risk group was comprised of 90 students, so 90 students were randomly selected and assigned to each of the other three groups. Fall, Winter, and Spring easyCBM© passage reading fluency and vocabulary progress monitoring data were collected. Data were analyzed using a mixed effects model, including omnibus tests and pairwise comparisons. Specifically, the pairwise comparisons found SpEd High Group's mean PRF scores were significantly lower from

all other groups and, importantly, significantly lower than the non-SpEd High group in Fall ($p = .01$), Winter ($p < .00$), and Spring ($p < .00$). Vocabulary pairwise comparisons for the SpEd High Group versus the non-SpEd High Group found no significant difference in Fall ($p = .45$), but significant differences in Winter ($p < .00$) and Spring ($p < .00$). Implications of my findings focus on district policy regarding: (a) the use of CBMs for risk labels, (b) Response to Intervention (RtI) as an effective framework, and (c) reading achievement growth for students with SLD are discussed.

CURRICULUM VITAE

NAME OF AUTHOR: Kira Danielle Fee

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
Pacific University, Forest Grove

DEGREES AWARDED:

Doctor of Education, 2020, University of Oregon
Preliminary Administer License, 2017, University of Oregon
Master of Arts, 2007, Pacific University
Bachelor of Arts, 2006, University of Oregon

AREAS OF SPECIAL INTEREST:

Inclusive Practices in Special Education
Districtwide Policy and Implementation of Multi-Tiered Systems of Support
New Teacher Mentorship and Retention
Districtwide Development and Monitoring of Systems

PROFESSIONAL EXPERIENCE:

Director of Special Education, Redmond School District: 2020-present

Assistant Director of Student Services, Redmond School District: 2018-2020

Learning Specialist, Springfield School District: 2017-2018

Graduate Assistant, University of Oregon: 2016-2017

Life Skills Teacher, Springfield Public Schools: 2013-2016

Life Skills Teacher, Lane ESD: 2007-2011

Special Education Teacher, Lebanon School District: 2006-2007

ACKNOWLEDGEMENTS

I wish to thank my committee members for their support and guidance throughout this project: Dr. Keith Hollenbeck, Dr. Gerald Tindal, Dr. Julie Alonzo, Dr. Laura Lee McIntyre, and Dr. Brian Megert. I would also like to thank exceptional faculty member Dr. Jo Smith for her thoughtful instruction, feedback, and guidance.

I have been fortunate to have the guidance of several strong women administrators who have provided me with mentorship and friendship. I thank Dr. Brooke Wagner for being a relentless cheerleader through the D.Ed. degree, and providing an example of relational leadership with humility. I thank Calli Dean for supporting me with humor and several closed-door venting sessions, and for taking the leap into administration first. I can always count on Calli to say what needs to be heard. I thank Martha Hinman and Karen Jordan for encouraging me through the dissertation process, for checking on my progress, and expecting me to finish strong.

Finally, I wish to thank my family. My sons, Cooper and Logan, who never see my limitations and flaws, and believe I can do anything. My husband Tyler Fee is my greatest advocate and is always willing to work beside me in pursuit of my dreams. Without his love and encouragement, completing this dissertation and degree would not have been possible. I love you three more than I can say.

I dedicate this dissertation to my sons, Logan and Cooper. May you always be kind humans and know your dreams are attainable. I support you with my whole heart.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

Research has shown that identification as having a disability and eligibility for special education services can result in differences in a child's public school experience compared to children without disabilities (Cortiella & Horowitz, 2014). Within the special education service array, the most common designation is under *specific learning disabilities* (SLD) (Oregon Department of Education, 2016; Zirkel, 2013). In the 2011-2012 school year, more than 40% of all students identified for special education were identified as having SLD, a number representing 4.75% of all students enrolled K-12 nationwide (Cortiella & Horowitz, 2014; Zirkel, 2013). Students qualify for special education services under SLD after undergoing an assessment process to identify specific areas of academic struggle. The intention is the student will then have access to specially designed instruction and services and thus improved achievement. The premise is that these targeted intensive interventions should put the student on an exponential growth line, helping them to catch up to their peers without learning disabilities.

However, the 2016 Oregon State Report Card showed the opposite for students with SLD. Students with SLD achieved lower state test scores, had lower rates of graduation and fewer went on to post-secondary education than their general education peers (Oregon Department of Education, 2016). Importantly, graduation and post-secondary rates must be considered distal indicators of K-12 special education success. For these distal indicators to be positive markers, it is important to identify more proximal early variables closer to when the students first were identified with their SLD eligibility. In Oregon, few students are identified as having SLD prior to second grade (Oregon Department of Education, 2018); therefore, a more proximal and better early

indicator of early special education program success might be the students with SLD's formative reading assessments at third grade across the school year. Therefore, my study sought to evaluate third grade students with a Specific Learning Disability label in reading, and a high Fall reading risk designation on a district's formative reading achievement assessment progress versus their non-disabled peers across a school year.

Special Education Policy

Before progressing with my examination of SLD and reading achievement progress across a year, it is imperative to understand the identification processes as specified in policy, placement practices, and outcomes of special education. Special education law provides for specially designed instruction and accommodations and/or modifications to the regular education curriculum or environment for students who qualify for an individual education program (IEP), as well as establishing supplementary aids and services, related services, or consultation, depending on students' academic and behavioral needs (Individuals with Disabilities in Education Act [IDEA], 2004). To qualify for special education, a student undergoes an evaluation process specific to the suspected disability.

Special education policy provides provision for legal protections against discrimination for students and requires schools to provide access to additional supports as established by the student's IEP team. With the passing of Public Law 94-142 (PL 94-142, 1975) children with disabilities were guaranteed legally protected equal access to public schools, including the right to a free and appropriate public education (FAPE) in the least restrictive environment (LRE). An appropriate education for students with an IEP would be individualized to the students' needs, and may be comprised of education

in regular classes, education in regular classes with the use of related aids and services, or special education and related services in separate classrooms for all or portions of the school day. An IEP team must consider and document the student's need for specially designed instruction in classrooms, at home, or in private or public institutions, and this need may be accompanied by the provision of related services such as speech therapy, occupational and physical therapy, psychological counseling, and medical diagnostic services necessary to the child's education (U.S. Department of Education, 2010).

For example, the expectations related to FAPE opened the school doors for children with disabilities, guaranteeing their rights to a free public school education that met their individual academic and behavioral needs. The stipulation that students be educated in the LRE provided that to the maximum extent appropriate, children with disabilities were educated with children who were not disabled, and alternative placement occurred only when the nature or severity of the disability of a child was such that supplementary aids and services were not sufficient support in the regular education classroom (Catabay, 2017). Public Law 94-142 was reauthorized and renamed in 1994 as the Individuals with Disabilities in Education Act (IDEA, 1994). IDEA was then reauthorized twice more in 1997 and 2004, with FAPE and LRE firmly entrenched in each of the reauthorized versions of that legislation. There is no question that FAPE and LRE have been key elements of federal legislation and policy and a mainstay of special education practice in the public schools. Local service provisions for students with disability are considered through the state and national policy lenses, and resulted in a charge to local education agencies to provide students with disability, not only access to

public education, but also the support needed for participation with typically developing peers in the general education classroom and curriculum.

In Oregon, if a student meets state and federal eligibility requirements for one or more of 11 disability categories and experiences an adverse educational impact due to the disability, a team can find that student eligible for special education services (Oregon Rules for Special Education, 2013). In 2015-16, Oregon had a special education enrollment rate of 13.3% (Oregon Department of Education, 2016).

Specific Learning Disability

Specific learning disability (SLD) is the most common eligibility label students receive to access special education services in Oregon. In 2015-16, Oregon had 25,932 students with the SLD label, which constituted 33.8% of all students with disabilities in the state (Oregon Department of Education, 2016). Moreover, Oregon's special education laws mirror federal guidelines. Special education policy at the national level guide state and local school district policy for providing services for students with the SLD label. IDEA (2004) and Oregon Administrative Rule (OAR) 581-015-2170 defined specific learning disability as:

> A disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations, including conditions such as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. Specific learning disability does not include learning problems that are primarily

the result of visual, hearing, or motor disabilities, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage.

Within the SLD eligibility classification, students may qualify for special education services for reading, writing, or math, or a combination of the three. Importantly, Oregon's OAR SLD definition made clear that a learning disability was a condition separate from an intellectual or mental health disability and was not the result of another physical impairment. The student's lack of achievement in the identified subject could not be due to the child's culture or socioeconomic status, limited English proficiency, or lack of appropriate instruction in reading or math.

As previously discussed, IEP teams are required to consider LRE when deciding placement for students with IEPs, including those with eligibility under the SLD category. Nationally, McLeskey, Landers, Willimson and Hoppey (2012) found that in 1990, about 23% of students with SLD received their instruction in regular education classroom settings. By 2007, this proportion had increased to 59%, while the proportion of students with SLD placed in self-contained or special school settings declined from 24% in 1990 to 11% in 2007. Thus, students with SLD have experienced increasingly more inclusive environments, and since 1990 they have seen the greatest growth in time spent in the regular education classroom of any disability category (McLeskey, et al., 2012). While McLeskey and colleagues. did not make definitive conclusions on why placement rates changed so drastically, they pointed to federal and state policies as contributing factors. For example, rates of regular education classroom participation changed fairly slowly between 1993 and 2002, and then more rapid growth occurred

between 2002 and 2007, which correlated with major federal policy adoption and implementation.

Interestingly, in Oregon, students with SLD eligibility have traditionally received instruction in the general education classroom at rates much higher than the national rate. In the 1992-1993 school year, 65% of Oregon students with SLD spent 80% or more of their school time in regular education classroom (Oregon Department of Education, 2018). Comparatively, in the 2011-2012 school year, 84% of Oregon students with SLD spent 80% or more of their school time in the regular education classroom, 14% were in the regular classroom 40-79% of their school time, 1% were in the regular classroom less than 40%, and less than 1% of students were in another placement, such as homeschool, private school, corrections, or special school (Oregon Department of Education, 2016). Even though Oregon rates of inclusive environments were relatively high in the early 90s compared to national rates, the state paralleled national trends in increasing rates of regular education classroom participation across time.

Title 1 Federal Education Policy and SLD

The Federal Title 1 legislation was reauthorized and renamed in 2001 as No Child Left Behind (NCLB). With the implementation of NCLB (2001) came a standards-based approach to public education, including students with special needs, with corresponding accountability assessments and an emphasis on inclusive education and differentiated instruction (McLeskey & Waldron, 2011). The result of the NCLB emphasis on accountability was what Fuchs, Fuchs, and Stecker (2010) described as a *blurring* of special education, where roles and responsibility between special and general education overlap. In addition, there was more emphasis on the achievement outcomes for students

accessing special education supports, and an expectation that students with disabilities would experience school outcomes similar to students without disabilities as measured by test scores and school completion rates (Oregon Department of Education, 2016).

Long-term school success indicators showed that students with disabilities in Oregon were not finding a level of success commensurate with their peers without disabilities. During the 2015-2016 school year, 10.2% of students in special education in Oregon had at least one discipline incident compared to 4.5% of students not in special education. Chronic absenteeism was another area where students with disabilities were overrepresented, with 25% of students with disabilities absent greater than 10% of school days, compared to 19% for all students. Students with disabilities experienced similar results relative to high school graduation rates. While the overall graduation rate in Oregon was 74.8%, only 55.5% of high school seniors with disabilities graduated in 2015-2016 (Oregon Department of Education, 2016).

The achievement gap between students with and without disabilities was evident much sooner than high school and could clearly be seen in third grade state assessment data. For the 2014-2015 school year, Oregon set an annual measurable objective that 54.5% of students in grades 3 through 5, including students with IEPs, would *meet* or *exceed* the state English Language Arts (ELA) standards as measured by the Smarter Balanced Assessment. Although 46% of all third-grade students met or exceeded Oregon's elementary ELA benchmark in the 2014-2015 school year, only 19% of students with IEPs reached the same standard. In 2015-2016, 20% of students on IEPs compared to 47% of all students met or exceeded, and in 2016-2017, 19% of students with IEPs met or exceeded the state reading benchmark compared to 45% of all third-

grade students. (Oregon Department of Education, 2017). Despite state and federal policy protecting students with disabilities and high expectations through accountability, all evidence supports the assertation that students with disabilities are experiencing limited academic achievement and school success.

Study Purpose

Despite findings that students with SLD are not experiencing comparable achievement results to their peers without disabilities, scant research has evaluated reading trajectories among students with the lowest reading scores. Because not all students with the lowest reading scores are identified as having SLD, an investigation of early readers identified as having SLD and those not identified is important. There is a need to investigate the relationship between the designation of a student as eligible for special education services and the achievement of these students.

The purpose of my study was to examine the reading achievement growth for students who received special education eligibility and were labeled as having SLD in reading at the beginning of third grade, compared to other students with similar reading achievement profiles grouped by reading risk, in a local school district that utilized Response to Intervention (RtI) for both SLD identification and intervention. The review of literature guided my research questions and methods, with the aim of informing a local school district's practices and policy around intervention success of students with SLD.

Review of Literature

The purpose of my review of literature was to identify and analyze studies that focused on: SLD identification using the RtI method, how SLD enrollment trends may or may not have changed after implementation of NCLB and IDEA, and how students with

SLD may receive intervention services within an RtI framework. The results were used to identify current research gaps and guide research questions and corresponding methods, with the ultimate goal of informing policy and practices for identifying students as having SLD and providing intervention for those students within an RtI framework.

In the following sections I summarize the main themes that surfaced in my literature search including: (a) predictive validity of early reading screeners and later SLD identification, (b) trends in SLD enrollment, and (c) effectiveness of the RtI model for improving early reading achievement for at-risk readers and students with SLD.

RtI Method of SLD Identification

Under IDEA 2004 and Oregon Administrative Rules, Oregon school districts can choose between three methods of identifying students for SLD including: (a) Discrepancy Model, (b) Patterns of Strengths and Weaknesses, and (c) Response to Intervention (RtI). Because my study focused on RtI, I briefly define the Discrepancy Model and the Patterns of Strengths and Weaknesses Model, and provide an in-depth analysis of RtI.

Discrepancy Model. Historically, the most common model of determining student eligibility for SLD required assessing students with achievement and IQ tests (Ihori & Olvera, 2015). If a student demonstrated a *severe discrepancy* and was achieving below his or her potential by at least two standard deviations, as established by the IQ test, they could be found eligible for Special Education services as a student with SLD.

Patterns of Strengths and Weaknesses Model. The second method of identification allowed by IDEA (2004) was the pattern of strengths and weaknesses (PSW) model, which explored whether a student "exhibits a pattern of strengths and weaknesses in performance, achievement, or both, relative to age, state-approved grade-

level standards, or intellectual development, that is determined by the group to be relevant to the identification of a specific learning disability" (Ihori & Olvera, 2015, p. 1).

RtI Model. The RtI model allowed states and districts to adopt an SLD identification method utilizing processes to assess a student's response to scientific, researched-based interventions. RtI is a decision-making framework for providing comprehensive support to students. It is a prevention-oriented approach, linking assessment and instruction that can provide for early identification of students who are at-risk. (National Center on Response to Intervention, 2010). Through a tiered system of interventions, RtI provides students with or without disabilities access to instruction and supports matched to their identified needs. In an RtI model, universal screenings are conducted to identify students who are at-risk. These students are provided instruction in the regular education classroom, and their progress is monitored over time. If a student does not respond to Tier I instruction, he or she is provided Tier 2 intensive, research-based instruction, and progress monitoring continues. If the student continues to be non-responsive to instruction, he or she would move to Tier 3 for even more intensive instruction and possibly special education assessment (Ihori & Olvera, 2015).

One primary advantage of RtI as a means of identifying students with SLD is that it does not require an assessment team to wait for the gap between achievement and potential to be considered wide enough, as is the case for the discrepancy model. Students can be identified at lower grades as they fail to demonstrate responsiveness to increasingly intense interventions, and instead of measuring an achievement gap, the team is measuring responsiveness. This approach is not without its problems, however. Measures of responsiveness vary and can result in inconsistencies in which students are

identified as SLD eligible (Fuchs, Fuchs, & Compton, 2004; Spencer et al., 2014). For example, Fuchs et al. (2004) found that alternate methods of assessing responsiveness produce different prevalence rates of reading disability and different subsets of unresponsive children. Spencer et al. (2014) considered the different operational definitions of reading disability including (a) nonword fluency, (b) oral reading fluency, and (c) listening reading comprehension, and the different rates of SLD identification each definition would produce. They found limited stability over time or agreement across operational definitions, raising the concern that increased variability in the identification of reading disability can influence which students are selected to take part in intervention programs, as well as distort the potential effectiveness of interventions.

Regardless of SLD identification method, a commonality exists in that accurate identification of students with SLD relies heavily on both informal and formal academic achievement assessment. Thus, it is critical that assessments are reliable and have a high level of predictive validity for SLD, and that teams can confidently use data from those assessments to accurately predict which students truly present with an SLD long-term. As previously discussed, SLD enrollment trends indicate evaluation teams are identifying fewer students as eligible for SLD, and teams are positively identifying a higher percentage of students referred as eligible for special education services, which could indicate more accurate identification measures are being used, but questions still remain.

SLD Enrollment Trends

After the implementation of IDEA (2004), states had two methods of identification available, including RtI. There were three articles in my review of literature that analyzed enrollment rates of students with SLD after implementation of RtI

(VanDerHeyden, Witt, & Gilbertson, 2007; Wanzek & Vaughn, 2011; Zirkel, 2013). Zirkel completed the largest and most comprehensive analysis of SLD enrollment by conducting an extant data review of U.S. Department of Education data across 17 school years beginning with school year 1995-1996 and ending with school year 2011-2012. Using data that included all special education enrollment rates nationwide, Zirkel verified that SLD was by far the most common disability category for all years studied. In addition, Zirkel reported:

> SLD enrollment numbers show a continuation of growth until a leveling off during 2000-2001 and 2001-2002 and then a more gradual but unreversed decline until the latest year of available data—2011-2012. During the same overall period, SLD enrollments similarly reflected an ascending then descending pattern, but with the turning point 1 year earlier, that is, during 1999-2000 and 2000-2001. As a third variation, SLD enrollments as a proportion of special education enrollments did not follow the same up-then-down pattern, instead showing a slow but steady decline for the entire period. (p. 474)

Through this analysis of data, Zirkel (2013) found the decline in SLD enrollment occurred prior to implementation of RtI, and the change of enrollment is likely due to other factors such as pressure to reduce number of students identified, increased emphasis on early intervention and school readiness, and increased enrollment in other disability categories including autism and other health impairment.

Two other studies examined overall SLD enrollment after RtI implementation. Wanzek and Vaughn (2011) reviewed extant data from a study on the efficacy of

implementing a three-tiered intervention system in Texas. Their study included seven elementary schools in one school district and included three cohorts of students over five school years. Their data showed "a trend in the direction of a decreased percentage of students identified for special education through each cohort of students; however, there were no statistically significant differences in the overall percentages of students identified for special education across cohorts" (p. 167). Wanzek and Vaughn concluded that while the results were not statistically significant, they were practically significant to the school staff who saw an overall decrease in SLD enrollment of 5 percentage points.

VanDerHeyden et al. (2007) conducted a similar study in a suburban district in Arizona. They evaluated the effects of an RtI approach to screening and eligibility determination on outcomes including evaluation and placement in special education. An RtI model was implemented in five elementary schools across three school years beginning in 2003-2004, and SLD enrollment trends were analyzed. VanDerHeyden and colleagues found that "fewer evaluations were conducted and evaluated students were more likely to qualify for services when STEEP (RtI) data were included in the team decision-making process" (p. 249). Across all schools included in the study, it was discovered that after implementation of RtI, the number of initial referrals for SLD evaluation dropped, while the rate at which those students qualified rose. They concluded that through the RtI process, the data used for referring and evaluating students for special education services resulted in more accurate identification of students for SLD, and the teams spent less time evaluating students who would not qualify.

RtI Framework

The most common approach to implementing RtI is the use of the 3-tier model. Within a 3-tier RtI model, reading instruction can be provided through three avenues including high quality general education (Tier 1), evidence-based intervention instruction (Tier 2), and individualized intervention or special education instruction (Tier 3; Fuchs et al., 2010; Harn, Chard, Biancarosa, & Kame'enui, 2011). As described by Fuchs et al. (2010), these types of instruction may not be place or teacher specific, but they may be related to the type of supports students need. Students achieving at expected levels typically receive instruction through the regular curriculum, and students who are identified as at-risk or labeled with SLD receive additional or replacement instruction through intervention and/or special education. This decision-making is highly reliant on high-quality screening and progress monitoring data, considered essential components to the RtI framework (Jenkins & Terjeson, 2011). RtI is built on intervention research that poor reading achievement can be improved, or even prevented, and is an alternative to traditional remedial and special education service models (Simmons, et al., 2008).

Coyne, Kame'enui, Simmons, and Harn (2004) analyzed the reading progress of students subsequent to beginning reading interventions, referring to two possible effects as *inoculation* and *insulin*. The *inoculation* hypothesis theorized that:

> Early intervention, if carefully designed and delivered, is sufficient to remediate, within a specified window of time, the phonological and alphabetic deficits of a significant percentage of children who are initially identified as at risk for reading disability, making further intensive

> intervention at a subsequent time during reading development unnecessary (Coyne et al., 2004, p. 91).

This view asserted that with effective early intervention, future reading struggles could be avoided; thus, intervention acted like a vaccination against reading disability. In contrast, the *insulin* hypothesis theorized that:

> Short term effects (i.e., the elimination of at-risk status) gained through early intervention can be maintained only with continued intensive support… children at risk of experiencing reading difficulties not only require highly explicit and systematic instruction to gain initial access to the complex alphabetic writing system, but they also require ongoing and intensive intervention to acquire later, more advanced reading skills. (Coyne et al., 2004, p. 91)

The insulin theory maintained that, like insulin-dependent diabetes patients, students identified with phonological deficits would be vulnerable again if interventions were discontinued.

Coyne et al. (2004) considered these two theories by examining the first-grade reading progress of children who participated in an intensive beginning reading intervention in kindergarten, to see if that intervention prevented first-grade difficulties. Participants included 59 children at risk of developing reading difficulties who received a seven-month beginning reading intervention in kindergarten. They were assigned to one of two groups in first grade: (a) code-based classroom instruction and a supplemental maintenance intervention, or (b) only code-based classroom instruction. Posttest measures assessed oral reading fluency, word reading, nonword reading, and

comprehension, and between groups analyses indicated that instruction groups did not differ on any posttest measure. Between 75% and 100% of students in both conditions attained posttest achievement comparable to their average-achieving peers. The results of this study support the inoculation theory, suggesting that intensive intervention can provide long-term positive results without sustained intervention. Interestingly, some students who responded positively to the kindergarten reading intervention who were randomly assigned to first grade maintenance intervention, did not benefit from that intervention, perhaps because they did not need it. In addition, students who did not respond as strongly to the kindergarten intervention were not included in the study because they needed different intervention than the study applied. Coyne et al. (2004) concluded these results highlight the need to differentiate intervention support targeted to individual student need.

In 2015, Balu and colleagues published an Evaluation of Response to Intervention Practices for Elementary Reading (hereafter, *The National Evaluation*; Balu et al., 2015). The National Evaluation was contracted by the Department of Education to analyze RtI implementation and the impact of intervention on reading achievement. It informed the field's knowledge about RtI in three ways: (a) it described implementation of RtI practices in multiple states at the school level, (b) it described practices in schools that had adopted RtI on their own and had implemented it for three or more years, and (c) it answered a question about effective targeting, by comparing the outcomes for students just below and just above the cut point of eligibility for intervention. The Balu and colleagues identified 146 RtI schools in 13 states implementing RtI for at least 3 years and because of the completeness of their implementation of RtI components, they

classified them as *impact schools*. The researchers collected survey data from teachers and staff, and individual-level fall screening test scores and tier placements for fall and winter of the 2011-12 school year for all students in grades 1-3 in the 146 impact schools.

The National Evaluation sought to answer three sets of research questions using the first part of their evaluation was a comparison of practices between school samples (Balu et al., 2015). They analyzed how the prevalence of RtI practices differ between a representative reference sample of schools and schools selected for the impact evaluation. They also investigated to what extent impact schools were implementing more RtI practices than the reference schools in their sample. Finally, they evaluated how special education identification rates in the impact sample compared with rates for the states as a whole. Balu et al.'s (2015) evaluation found that a majority of schools in both groups reported full implementation of an RtI framework for reading, with 86% of impact schools and 56% of reference sample schools reporting full implementation. Impact schools were more likely to report providing time for Tier 2 intervention at least three times a week (97% compared to 80% of reference sample schools) and were also more likely to report providing time for Tier 3 intervention at least five times a week (68% compared to 47%). Among impact sample schools, 83% conducted universal screening assessments of students at least twice a year (compared with 59% of reference sample schools), and they were more likely to follow a prescribed sequence of steps to respond to students who read below grade-level benchmarks (95% compared to 88%, respectively). For students suspected of having SLD, impact sample and reference school samples were not significantly different in their use of data to monitor student progress following

implementation of reading interventions. They also found that special education identification rates were comparable between the impact sample in the 13 states.

Balu et al.'s (2015) second set of research questions were a comparison of reading services between reading groups at different skill levels. In impact schools (those with three or more years of implementing RtI), they wanted to investigate to what extend schools (a) placed students in tiers as suggested by earlier RtI models, (b) adjusted tier placement during the school year, (c) varied in how they organized reading services for specific reading levels, and (d) varied intensity of services for students reading below grade level versus for students reading at or above grade level. The evaluation found that impact schools were using data to initially assign students to Tiered instruction, and about 25% of students moved tiers within the school year. The evaluation found that although all impact sample schools complied with RtI implementation criteria, some schools showed variation on three aspects of RtI implementation described in literature.

1. Prior studies generally served only students reading below grade. However, 45% of impact sample schools offered reading intervention services to at least some students reading at or above grade level, as well as to those reading below grade level.

2. Previous studies often designated intervention as supplemental services provided in addition to the core general education reading curriculum. In this study, 69% of impact schools offered at least some intervention services during the core program.

3. Previous, more controlled studies of RtI relied on non-classroom teaching staff to provide intervention services. This study found that 37% of first grade intervention services were provided by classroom teachers rather than intervention support staff.

Balu et al.'s (2015) third set of research questions examined the impact of RtI on reading outcomes for students. For students who fell just below school-determined standards for each grade on screening tests, they evaluated (a) the effects on reading achievement of actual assignment to receive reading intervention services (in addition to core instruction), (b) the extent of variation in estimated impacts across RtI schools, and (c) the estimated impact associated with certain school features or student characteristics. The national evaluation found that assignment to Tier 2 or Tier 3 intervention services in impact sample schools had a negative effect on performance on a comprehensive reading measure for first-graders just below the Tier 1 cut point on a screening test. The estimated effects on reading outcomes in second and third grade were not statistically significant. They also found at the student level, for some outcomes and grades, students who had an IEP appeared to have been affected by the treatment more negatively.

In response to Balu et al.'s (2015) National Evaluation, Fuchs and Fuchs (2017) examined the study's methods, findings, and conclusions to explain what could and could not be learned from it about RtI effectiveness. After their analysis, they cautioned against accepting the national evaluation as "a proper, bottom-line, summative analysis of RtI effectiveness" (p. 259). They identified several concerns with the national evaluation including its narrow focus and several RtI implementation problems in the participating schools. They wrote,

> If its results are wrongly interpreted to mean "RtI doesn't work," we fear that the hard work of practitioners, researchers, and policy makers to make it effective will stop. We fear too that 15 years of research on RtI-related

> assessments and interventions will also be wrongly judged as ineffective and unimportant. (p. 266)

Fuchs and Fuchs (2017) doubted whether the lack of intervention achievement outcomes for impact schools participating in RtI were a result of RtI itself not being an effective framework for intervention, but rather the weak or inconsistent implementation of RtI at the impact schools studied. In their view, RtI is a complex reform, and data indicates that many schools across the country are struggling to implement RtI with fidelity. It requires a dynamic, well-orchestrated use of measures and intervention to optimize student performance (Simmons et al., 2008).

When it came to special education, specifically students with SLD, and intervention services within the RtI framework, Fuchs and Fuchs (2017) claimed the practice was less clear.

> Despite what we believe is strong agreement on this point, many students in America's schools who are unresponsive to Tiers 1 and 2 suffer one of two unfortunate fates. In one scenario, they linger indefinitely in Tier 2, receiving the same unhelpful intervention again and again. In a second scenario… they move from Tier 2 to "special education," which removes them from RtI frameworks and returns them to the general classroom that proved inadequate at the start. We believe that the most intensive level of an RtI system should be special education. (p. 263)

Notwithstanding what Fuchs and Fuchs (2017) stated to be a strong belief in research, some practitioners, advocates, and policy makers defined RtI as a general education reform, separate from the policies and practices of special education. While an important

means of SLD identification, students with SLD are sometimes excluded from the RtI instructional framework.

The National Center on Response to Intervention (2010) identified four essential components of RtI: (a) a school-wide, multi-level instructional and behavioral system for preventing school wide failure, (b) screening, (c) progress monitoring, and (d) data-based decision making for instruction, movement within the multi-level system, and disability identification. If students with SLD are to be included in these components, then special educators should compare their academic performance to peers without disabilities.

Study Context and Research Questions

As noted earlier, the purpose of this review of literature was to provide the reader with an examination of SLD identification using the RtI method, SLD enrollment trends after implementation of NCLB (2002) and IDEA (2004), and effectiveness of the RtI model on early reading achievement for at-risk readers. However, it remains unclear if students identified with SLD, in an RtI model, demonstrate differential achievement growth compared to their peers without special education identification in an RtI framework. This remains an important question to investigate, considering the graduation rate and test score achievement gap between students with disabilities and students without disabilities. As previously described, federal policy holds districts and schools accountable for the progress of their students with disabilities. In a district with full implementation of RtI both for SLD identification and reading intervention, it is important to evaluate what achievement can be expected for students with SLD. These gaps in the literature led me to the following research questions for my study:

1. Is there a significant difference between third-grade easyCBM© Fall, Winter, and Spring Passage Reading Fluency benchmark scores for (a) students identified for special education with a *high risk* Fall designation (SpEd High) – the SLD-R Tier 4 Group, (b) non-special education students with a *high risk* Fall designation (Non-SpEd High) – the Non-SpEd High Risk Tier 3 Group, (c) non-special education students with a *some risk* Fall designation (Non-SpEd Some) – The Non-SpEd Some Risk Tier 2 Group, and/or (d) non-special education students with a *low risk* Fall designation (Non-SpEd Low) – the Non-SpEd Low Risk Tier 1 Group?
2. Is there a significant difference between third-grade easyCBM© Fall, Winter, and Spring Vocabulary benchmark scores for (a) students identified for special education with a *high risk* Fall designation (SpEd High) – the SLD-R Tier 4 Group, (b) non-special education students with a *high risk* Fall designation (Non-SpEd High) – the Non-SpEd High Risk Tier 3 Group, (c) non-special education students designated with a *some risk* Fall designation (Non-SpEd Some) – The Non-SpEd Some Risk Tier 2 Group, and/or (d) non-special education students with a *low risk* Fall designation (Non-SpEd Low) – the Non-SpEd Low Risk Tier 1 Group?

CHAPTER II

METHODS

In my study, I analyzed deidentified extant student achievement data from a convenience sample obtained from a local school district. Formative reading data came from the district's third-grade easyCBM© assessment scores on Passage Reading Fluency and Vocabulary, collected during the 2015-2016, 2016-2017 and 2017-2018 school years. Basic student demographic variables also were collected from the district's student information system, Synergy, and were combined with easyCBM© data to provide demographic information for the research groups identified above.

Research Design

I employed a quantitative quasi-experimental, longitudinal design for this study (Babbie, 2013). Three formative reading achievement data points collected during Fall, Winter, and Spring trimesters were used to compare four groups of students over the course of a single school year. The primary benefit of using a year-long longitudinal design was the ability to study a phenomenon over a period of time rather than obtaining a single data point, which allowed me to examine the change in student achievement over the course of an entire school year. During this study I examined deidentified extant data of students who had already been assigned to treatment conditions by the district rather than randomly assigning students to groups myself, resulting in a quasi-experimental design and not a true experimental design for this study (Babbie, 2013).

Setting

This study was conducted in a suburban school district in the Pacific Northwest. In 2018, the city had an estimated population of 61,525, however it is a part of a larger

community with its close neighbor that is a larger city with an estimated population of 171,245. Residents are approximately 79.8% white, not Hispanic or Latino, 11.6% Latino, 5.7% two or more races, and 4.4% other (United States Census Bureau, 2020). In 2018 the median household income was $43,157, which is substantially lower than the state average of $59,393. Persons in poverty was reported as 20.4%, compared to a state average of 12.6% (U.S. Census Bureau, 2020). The district serves about 10,854 students in 21 schools, including 12 elementary schools, four middle schools, four high schools, and one Grade 6-12 charter. Of those students, 1,995 (18.4%) have an IEP, and 606 are English language learners (5.6%; National Center for Education Statistics, 2020).

The school district was selected for three reasons. The first was due to convenience and its willingness to participate in university-sponsored research. The district has a long history of being involved in various educational research projects with major universities in Oregon and across states. The second reason the school district was selected for participation was that it uses the RtI model for identification of students with SLD with district-wide implementation. This was an essential quality of the participating district in order to answer the identified research questions. To support the RtI process, the district has adopted a benchmark assessment system to monitor the progress of every student, grades K-8. Implementation of progress monitoring is overseen at the district level, with assessment administration the responsibility of each school. All schools are expected to participate, with the expectation that all students are assessed unless alternative assessment is stipulated on an IEP. The established assessment protocols for this school district provided a data set that was as complete as possible for this study.

Reading benchmark assessments using easyCBM© are conducted for Fall, Winter, and Spring within a districtwide timeline and three-week window. Risk criteria are established at the district level. For this district, students scoring above 50th percentile are considered *low risk*, between the 25th to 50th percentiles are considered *some risk*, and below the 25th percentile are considered *high risk*. School-based multi-disciplinary grade level RtI teams meet regularly to review data and make intervention decisions. This team may also refer students for special education (SpEd) pre-referral process.

The third reason the district was selected was based upon the structure of SpEd programming. The district follows a *community model* of providing SpEd services to elementary and middle school students. The community model for SpEd resulted in elementary schools with a distribution of SpEd students and services throughout the district. The majority of students with IEPs are supported in their assigned school as determined by the district drawn school boundaries, and each school is responsible for identifying students for SpEd services following district procedures and providing services using district-allocated resources. The district-wide distribution of students combined with district-wide implementation of RtI procedures allow analysis across schools, resulting in a large sample of participants using the complete sample frame.

Participants

The sample frame for my study included all third-grade students in the school district during four school years: 2014-2015, 2015-2016, 2016-17 and 2017-18. I selected third grade because the third-grade students with SLD will have been recently found eligible through the RtI process, but will have not yet spent many years receiving special education services. Four school years were selected to increase the number of participants

identified as SLD in reading, and to limit potential contextual factors that could influence data from selecting only a single school year. During the selected school years, a total of 3,210 third-grade students had complete Fall, Winter, and Spring easyCBM© scores. There were two *a priori* rules or conditions established that excluded students from participation: (a) a special education eligibility other than SLD in reading, and (b) incomplete assessment data (lacking a Fall, Winter, & Spring CBM score). Students who were identified with either or both of those conditions were excluded from the research sample, and their achievement data were not included in the statistical analysis because their scores would add construct irrelevant error to the analysis.

All remaining students were placed into one of four groups: (a) students not in special education with a Fall designation of *low risk* (non-sped low risk), (b) students not in special education with a Fall designation of *some risk* (non-sped some risk), (c) students not in special education with a Fall designation of *high risk* (non-sped high risk), and (d) students in special education with a Fall designation of *high risk* (SLD high risk). My final analytic sample included 90 students in the SLD high risk group that met my criteria. To create equal-sized groups for statistical comparison, I used stratified random sampling. I randomly selected 90 students from each of the other three groups as participants, resulting in a total sample of 360 students. Table 1 displays descriptive statistics for (a) Free and Reduced-Price Meals (FARMs), (b) sex, (c) Limited English Proficiency (LEP), (d) migrant education, and (e) Talented and Gifted (TAG). Ethnicity and race data for all student sub-groups are displayed in Table 2. Attendance data, by SpEd group and season are displayed in Table 3.

Table 1

Partial Demographics by Group

Category	Group				Total	%
	SpEd High	Non-SpEd High	Non-SpEd Some	Non-SpEd Low		
FARMs						
No	5	17	25	28	75	20.83
Yes	85	73	65	62	285	79.17
Sex						
Girls	32	40	41	43	156	43.33
Boys	58	50	49	47	204	56.67
LEP						
No	83	72	85	85	325	90.28
Yes	7	18	5	5	35	9.72
Migrant Ed						
No	88	87	86	90	351	97.50
Yes	2	3	4	0	9	2.50
TAG						
No	90	90	90	82	352	97.78
Yes	0	0	0	8	8	2.22

Note. FARMs = Economically disadvantaged; LEP = Limited English proficiency status; TAG = Talented and Gifted status.

Table 2

Race/Ethnicity Demographics by Group

Category	Group				Total	%
	SpEd High	Non-SpEd High	Non-SpEd Some	Non-SpEd Low		
Ethnicity						
Hispanic/Latino	18	28	24	16	86	23.89
Not Hispanic/ Latino	72	62	66	74	274	76.11
Race						
American Indian/Alaskan Native	4	3	0	1	8	2.22
Asian	0	1	0	1	2	0.56
Black	2	1	2	0	5	1.39
Multi-Race	2	6	7	6	21	5.83
Non-US Native American	0	2	0	0	2	0.56
White	82	77	81	82	322	89.44

Table 3

Average Percentage (%) of School Days Attended by Group and Season

Group	Fall			Winter			Spring		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
SpEd High	90	93.09	0.06	90	93.59	0.05	90	94.48	0.05
Non-SpEd High	90	93.71	0.06	90	94.37	0.04	90	94.25	0.04
Non-SpEd Some	90	94.77	0.05	90	94.86	0.05	90	94.90	0.05
Non-SpEd Low	90	94.08	0.07	90	94.12	0.04	90	94.46	0.04
Total	360	93.91	0.06	360	94.23	0.05	360	94.52	0.05

Instrumentation

Two instruments were used to measure reading achievement, and demographic data was also collected for descriptive statistics. Synergy Student Information System provided the demographic data and easyCBM© supplied the reading achievement data.

easyCBM© Curriculum-based System

The formative reading assessment instrument used to measure student achievement was easyCBM© curriculum-based system (CBM; Alonzo, Tindal, Ulmer, & Glasgow, 2006). easyCBM© is a benchmark, progressing monitoring and formative reporting assessment tool for grades K-8. It was designed for use in measuring student achievement in math and reading, and contains assessments that are aligned to Common Core State Standards. Two reading easyCBM© measures from third grade were utilized for statistical analysis: (a) passage reading fluency and (b) vocabulary.

Passage Reading Fluency. The easyCBM© passage reading fluency (PRF) assessment is administered beginning Spring of grade 1, through grade 8. A district-trained test assessor administers the assessment to students individually, following a 13-step process. The assessor provides the student with a grade-level reading passage and asks the student to read it aloud. The assessor times the student for one minute, and while the student is reading, records errors. After the minute is finished, the assessor calculates the number of words correctly read. Thus, the total score for PRF is the number of words correct per minute (wcpm).

Alternate form and test-retest reliability for the easyCBM© PRF measures was investigated in six separate studies (Anderson et al., 2014). Each study, with the exception of one, also conducted generalizability and decision studies to explore how

various facets of the measurement process impacted the reliability of the PRF measures. Grade 3 alternate form reliability was reported at .94-.95. Test-retest reliability at grade 3 ranged from .84-.94, with a median of .90. Reliability studies also included generalizability analyses primarily to examine how specific facets of the measurement process related to the reliability of the measures. Grade 3 G-coefficients ranged from .94-.98. These results largely confirm the results of the test-retest and alternate form reliability analyses, suggesting the measures are highly stable (Anderson et al., 2014).

Vocabulary. The Vocabulary measures (grades 2–8) are intended to assess vocabulary proficiency. The words included in the Vocabulary measures were selected from a variety of content materials and were extensively field-tested. The bank of items represents a wide range of difficulty all aligned to grade-level content standards. The test can be administered via paper-and-pencil or online and takes approximately 10-to-15 minutes to complete. The total score is the number of correct responses that the student provides. During the assessment, the student is presented with a sentence with a bolded single word and asked to select the meaning of the bolded word out of three options. A total of 20 items are presented to the student. The student does not receive correct or incorrect answer feedback as he or she progresses through the assessment.

The vocabulary portion of easyCBM© is newer than the other measures and was implemented in 2011-2012. Wray, Alonzo, and Tindal (2014) investigated the internal consistency of the easyCBM© Vocabulary measures for the Fall and Winter benchmark across all grades (2-8). Cronbach's Alpha and Split-half reliabilities (first half / second half) as well as item level statistics of the top and bottom 27th percentiles, were estimated. Cronbach's Alpha ranged from .76 to .84 and had a median of .81 for all

vocabulary measures in both the Fall and Winter. Split-half reliabilities ranged from .61 to .75 for the first and second half of the measures, with a median of .66 and .69, respectively. The correlation between the two halves ranged from .58 to .72, with a median correlation of .64. Relative to the top / bottom reliability, all items performed as expected, with higher percentile students getting the items correct more often than the lower percentile groups (Anderson et al., 2014).

Synergy Student Information System

Synergy is the student information system used by the participating district. Upon student enrollment, demographic and special education information is collected from parents and entered by each school. Student-level variables including sex/gender, race/ethnicity, English Learner status, free and reduced meal status, special education identification, attendance, and grade level were retrieved from Synergy for this study.

Data Collection

Deidentified extant data were obtained directly from the school district. Within his usual job duties, the district data analyst is responsible for data downloads and is designated to provide Synergy Special Education support to district staff. The deidentified data set was obtained by the data analyst via download from easyCBM© and Synergy. The data were merged into a single file, matching students from each database to create a complete academic and demographic profile. The data analyst removed information that individually identified students, including name, identification number, and classroom information. Before beginning statistical analysis, I applied *a priori* rules to remove students who (a) had a special education eligibility other than SLD in reading, (b) received English Language Development services, and/or (c) had incomplete

formative reading assessment data across the Fall – Winter – Spring trimesters. I then sorted remaining students into the four groups previously described: (a) non-sped low risk, (b) non-sped some risk, (c) non-sped high risk, and (d) SLD high risk.

Data Analysis and Interpretation

A mixed effects model was fit to the data, with season and instructional group estimated as fixed effects (a single coefficient estimated and assumed to a have a constant effect across participants), along with their interaction, and participant estimated as a random effect (i.e., each student had their own estimated intercept). This model was equivalent to a within-subjects analysis of variance with a subject-level random effect. The model intercept represented the average score for the reference group, students in the Fall who were in the non-SpEd high risk instructional group.

CHAPTER III

RESULTS

In this section, I present results for both research questions. A mixed effects model was used to examine third grade easyCBM© PRF and Vocabulary scores for Fall, Winter, and Spring across four different groups. I provide descriptive statistics, then provide the results of the statistical analysis for research questions one and two.

Descriptive Counts, Means and Standard Deviations by Season

Table 4 provides mean PRF and Vocabulary scores and percentiles for by season. The Fall overall mean score was the lowest for PRF. Winter and Spring PRF means were 104.79 and 102.58 wcpm, respectively. It is interesting to note that the overall PRF mean score was lower for the Spring measure than the Winter measure. Fall Vocabulary had the lowest mean, followed by Winter, then Spring.

Table 4

Means and Standard Deviations by Season and Measure

Season	*n*	PRF		Vocabulary	
		M (*SD*)	*M* %tile	*M* (*SD*)	*M* %tile
Fall	360	62.94 (38.44)	29.82	12.31 (4.57)	30.39
Winter	360	104.79 (48.55)	39.90	14.62 (4.40)	35.51
Spring	360	102.58 (47.21)	40.23	15.76 (3.84)	32.30

Boxplots in Figure 1 visually display means and standard deviations by season for PRF and Vocabulary measures. Besides the means and standard deviations by season for each measure, the boxplots also provide an indication of what the overall distribution looks like for each seasonal time point. PRF scores increased from Fall to Winter, and then decreased in Spring, although they remained higher than Fall. Vocabulary scores

also increased from Fall to Winter. For both PRF and Vocabulary, the Spring interquartile range was higher than the Fall median.

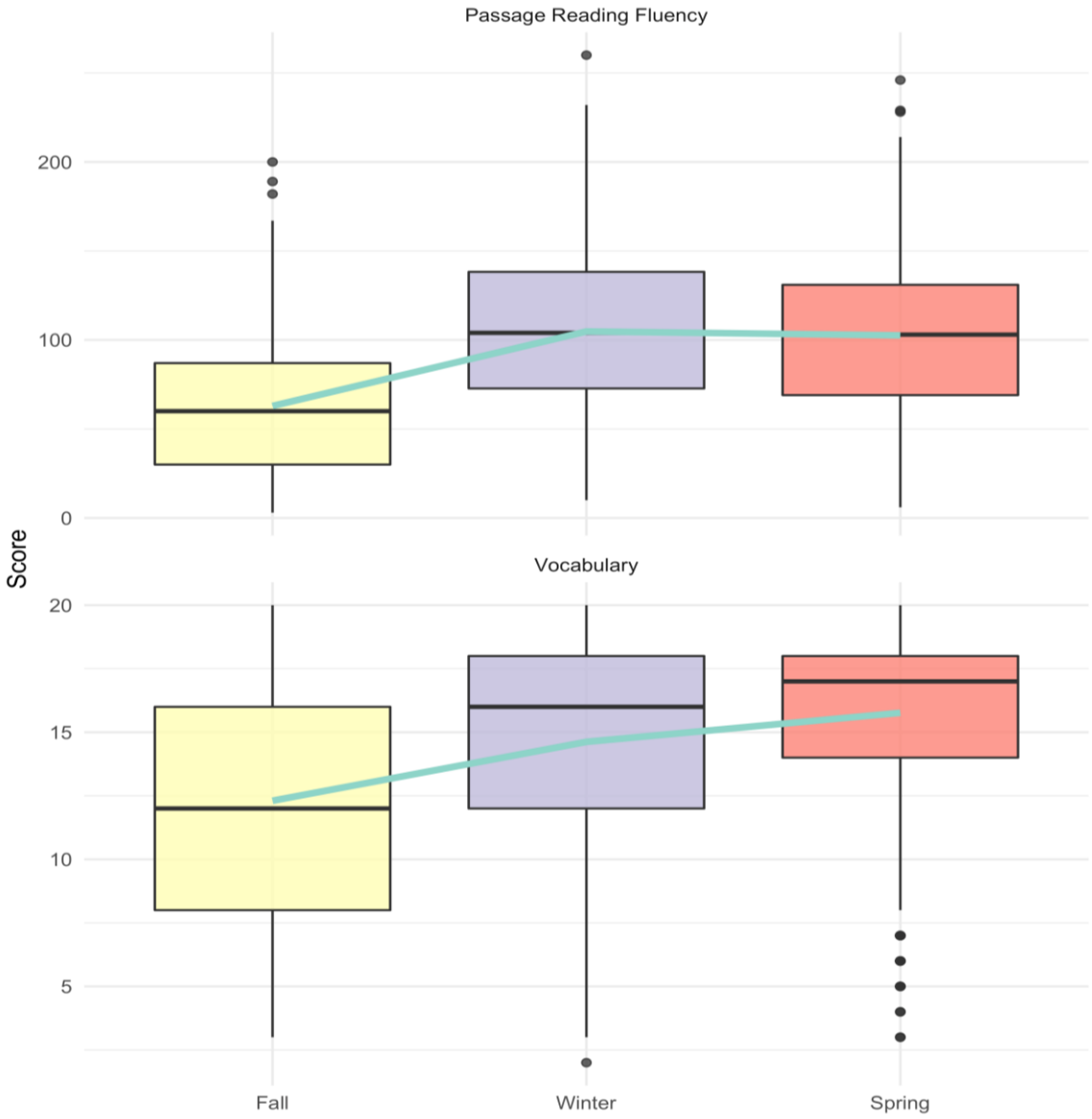

Figure 1. Distribution of scores by measure and season.

Descriptive Counts, Means and Standard Deviations by Group

Table 5 displays separate PRF means, standard deviations (in parenthesis), and percentile for each group. Of the four groups, SpEd High had the lowest overall Fall PRF mean score of 25.63 wcpm, whereas Non-SpEd High had a score of 41.28 wcpm. In Winter, the SpEd High mean score increased to 75.83 wcpm, but remained the lowest of all groups. In contrast, Non-SpEd High had a mean score of 98.20 wcpm. The trend continued to Spring, with SpEd High mean score of 75.10 wcpm and Non-SpEd High mean score of 94.28 wcpm. For complete PRF mean scores, see Table 5.

The Vocabulary mean scores mirror PRF Scores. Of the four groups, SpEd High had the lowest overall Fall vocabulary mean score of 8.11 wcpm, whereas Non-SpEd High had a score of 8.78 wcpm. In Winter, the SpEd High mean score increased to 11.64 wcpm, but remained the lowest of all groups. In contrast, Non-SpEd High had a mean score of 14.36 wcpm. The trend continued to Spring, with SpEd High mean score of 13.33 wcpm and Non-SpEd High mean score of 15.34 wcpm. For complete vocabulary mean scores, see Table 5.

The mean Vocabulary percentile scores displayed in Table 5 are of particular note. The SpEd High group had a Fall mean PRF percentile of 4.98, which increased to 24.51 in Winter and 24.72 in Spring. The Non-SpEd High group had a Fall PRF mean percentile of 11.58, which increased to 33.73 in Winter and 33.98 in Spring. Non-SpEd Some had a Fall percentile of 35.18 in Fall that increased to 43.71 in Winter, and then decreased to 42.41 in Spring. The Non-SpEd Low group had a Fall percentile of 67.53 in Fall that decreased to 57.63 in Winter and then increased to 59.79 in Spring.

As with PRF, the mean Vocabulary percentiles are noteworthy. Students in the SpEd High group had a mean Fall vocabulary percentile of 8.93, which increased to 21.58 for Winter and 21.71 for Spring. The Non-SpEd High group also demonstrated mean percentile growth, beginning with 8.78 for Fall, increasing to 21.58 for Winter and 32.70 for Spring. This is in contrast to the Non-SpEd Some and Non-SpEd Low groups, both of which had higher Fall mean vocabulary percentiles than Spring. For complete PRF and Vocabulary mean percentiles, see Table 5.

Table 5

Descriptive Statistics by Group, Season, and Measure

Group	*n*	PRF		Vocabulary	
		M (*SD*)	%tile	*M* (*SD*)	%tile
Fall					
SpEd High	90	25.63 (14.05)	4.98	8.11 (2.23)	8.93
Non-SpEd High	90	41.28 (17.45)	11.58	8.78 (2.7)	11.23
Non-SpEd Some	90	73.83 (16.68)	35.18	14.59 (2.15)	37.19
Non-SpEd Low	90	111.00 (29.04)	67.53	17.74 (1.37)	64.22
Winter					
SpEd High	90	75.83 (52.86)	24.51	11.64 (4.92)	21.58
Non-SpEd High	90	98.20 (45.27)	33.73	14.36 (4.24)	21.58
Non-SpEd Some	90	112.42 (33.25)	43.71	15.84 (3.06)	38.93
Non-SpEd Low	90	132.71 (42.68)	57.63	16.64 (3.44)	48.83
Spring					
SpEd High	90	75.10 (49.9)	24.72	13.33 (4.73)	21.71
Non-SpEd High	90	94.28 (45.52)	33.98	15.34 (4.09)	32.70
Non-SpEd Some	90	108.14 (29.79)	42.41	16.86 (2.22)	34.32
Non-SpEd Low	90	132.79 (41.87)	59.79	17.52 (2.2)	43.32

Boxplots in Figure 2 display means and standard deviations by season and group and also provide an indication of the overall distribution for each season. SpEd High was

the lowest group for each measure seasonally and had scores distinct from Non-SpEd High. From Fall to Spring, the distribution of scores widened for SpEd High.

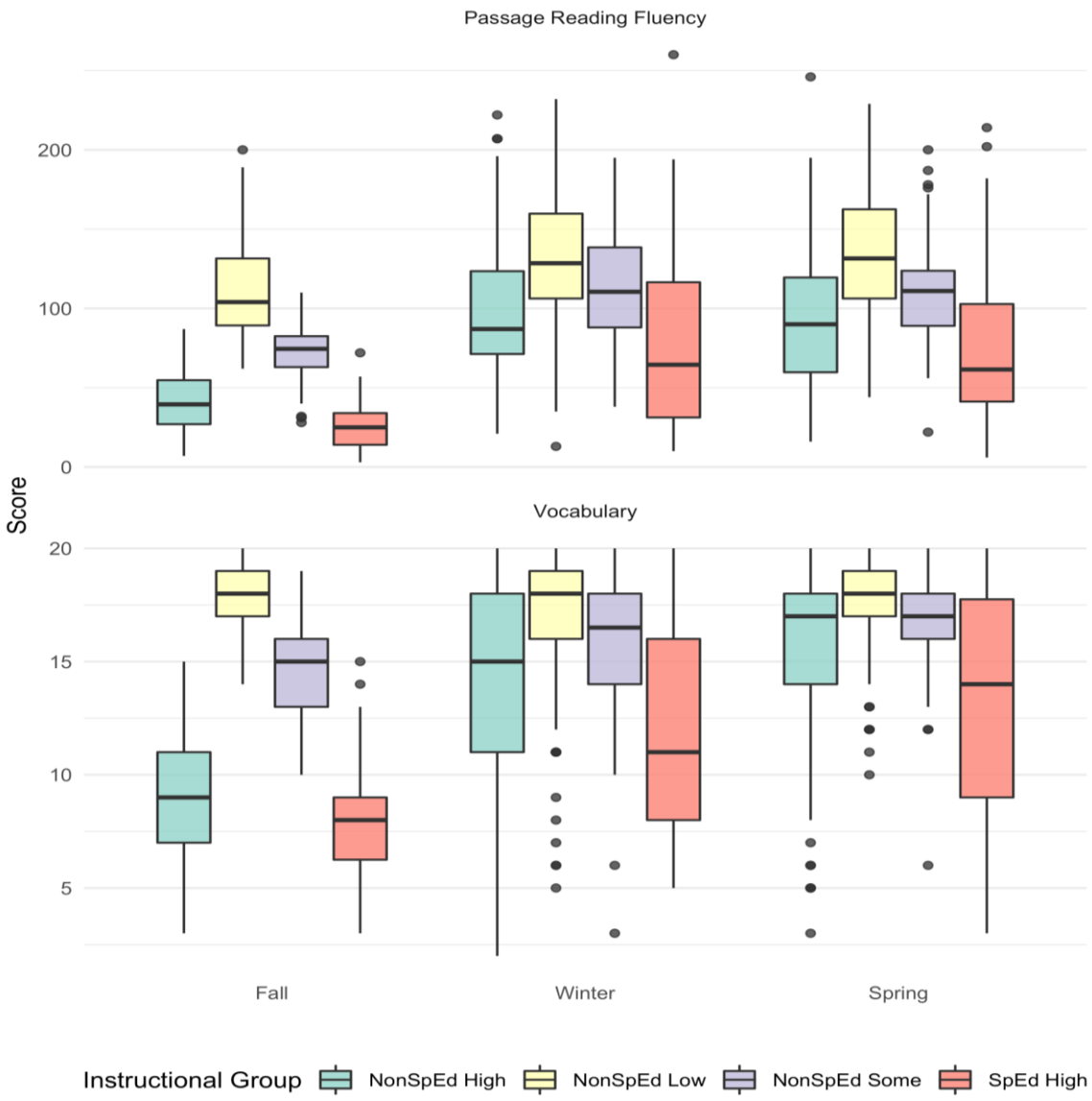

Figure 2. Distribution of scores by season and group.

Finally, Tables 6 and 7 display omnibus tests of significance for *Instructional Grouping*, *Season*, and their interaction. For both PRF and Vocabulary, each of these variables significantly related to students' test scores. Table 6 shows that the PRF scores had a significant main effect for both group and season. That meant that (a) at least two

of the four groups (non-SpEd low risk, non-SpEd some risk, non-SpEd high risk, and SpEd High) and (b) two of the three testing times (Fall – Winter – Spring) were significantly different. More importantly, Table 6 displays a statistically significant interaction between group and season on PRF. Thus, pairwise comparisons between the four groups were conducted to find where significant interactions occurred.

Table 6

Omnibus Tests of Significance: Passage Reading Fluency (PRF)

Group	*Sum Sq*	*Mean Sq*	*NumDF*	*DenDF*	*F*	*Pr (>F)*
Group	347156.78	115718.93	3	581.93	117.81	0
Season	399390.07	199695.03	2	752.58	203.30	0
Group x Season	40495.14	6749.19	6	752.58	6.87	0

Vocabulary scores also had a significant main effect for both group and season (see Table 7). More importantly, there was a statistically significant interaction between group and season on Vocabulary scores. Therefore, pairwise comparisons between the four groups were conducted to find where the interactions occurred.

Table 7

Omnibus Tests of Significance: Vocabulary

Group	*Sum Sq*	*Mean Sq*	*NumDF*	*DenDF*	*F*	*Pr (>F)*
Group	3897.71	1299.24	3	518.33	149.22	0
Season	2235.65	1117.82	2	753.18	128.39	0
Group x Season	1591.91	265.32	6	753.18	30.47	0

Research Questions One and Two

Research Question One asked if there was a significant difference between easyCBM© Fall, Winter, and Spring scores across the four groups, and Research Question Two asked the same but for vocabulary. The parameter estimates from the mixed effects model are displayed for both PRF and vocabulary in Table 8.

PRF

For PRF, students in the Non-SpEd High group scored an average of 41.44 wcpm in the Fall (the model intercept), 98.36 wcpm in the Winter, and 94.44 wcpm in the Spring. Students in the Non-SpEd Low group scored an average of 112.71 wcpm in the Fall, 134.42 wcpm in the Winter, and 134.50 wcpm in the Spring, while students in the Non-SpEd Some group scored an average of 75.10, 113.69, and 109.41 wcpm across season, respectively. Finally, the SpEd High group scored an average of 24.60 wcpm in the Fall, 74.8 wcpm in the Winter, and 74.07 wcpm in the Spring. Students' progress throughout the year was generally, highly non-linear, and with multiple student groups actually regressing slightly from Winter to Spring.

Vocabulary

In Vocabulary, students in the Non-SpEd High group scored an average of 8.79 correct words in the Fall (the model intercept), 14.37 words in Winter, and 15.36 words in Spring, on average. Students in the Non-SpEd Low group scored an average of 17.75 words in Fall, 16.65 words in Winter, and 17.53 words in Spring, while students in the Non-SpEd Some group scored an average of 14.68, 15.94, and 16.95 words across seasons, respectively. Finally, the SpEd High group scored an average of 8.07 words in Fall, 11.61 words in Winter, and 13.30 words in Spring.

Approximately 30% and 21% of the variability in scores was between students for PRF and Vocabulary, respectively. After accounting for instructional grouping, students' initial achievement in the Fall varied between students with a standard deviation of 20.74 wcpm for PRF, and 1.53 words for Vocabulary.

Pairwise Comparisons for PRF and Vocabulary

PRF. Follow-up pairwise comparisons are displayed for PRF in Table 9 and Vocabulary in Table 10. Note that these tables display both the estimated difference between each instructional group within each season and the corresponding *t* statistic with a Tukey correction factor denoting whether the mean difference between groups was larger than would be expected by random sampling variability (i.e., *p*-value). For PRF, these follow-up comparisons indicated that all instructional groups had significantly different mean scores within each season. Further, the magnitude of these differences was generally in the expected direction. For example, the difference between the Non-SpEd Low group and the SpEd High groups was consistently the largest across seasons.

Vocabulary. For Vocabulary, the mean differences between instructional groups were generally, but not universally, significantly different. In the Fall, the Non-SpEd High and SpEd High groups were not statistically different, with a mean difference of 0.72 words. In the Winter, the Non-SpEd Low and the Non-SpEd Some groups were also not statistically different, with only 0.71 words separating the mean scores for each group. This same comparison of those two student groups was also not statistically different in the Spring ($\mu\Delta$= 0.58, *SE* = 0.48, *p* = 0.63). However, all other comparisons were significantly different.

Table 8

Mixed Effects Model Parameter Estimates

Predictors	PRF			Vocabulary		
	Est	*CI*	*p*	*Est*	*CI*	*p*
Fixed effects						
Intercept	41.44	[33.81, 49.06]	<0.001	8.79	[8.11, 9.47]	<0.001
season [Winter]	56.92	[47.77, 66.08]	<0.001	5.58	[4.72, 6.44]	<0.001
season [Spring]	53.00	[43.84, 62.16]	<0.001	6.57	[5.70, 7.43]	<0.001
grouping [Non-SpEd Low]	71.27	[60.73, 81.81]	<0.001	8.96	[8.01, 9.91]	<0.001
grouping [Non-SpEd Some]	33.66	[23.03, 44.29]	<0.001	5.89	[4.94, 6.85]	<0.001
grouping [SpEd High]	-16.84	[-27.42, -6.25]	0.002	-0.72	[-1.68, 0.23]	0.137
grouping [Non-SpEd Low] * season [Winter]	-35.21	[-48.16, -22.26]	<0.001	-6.68	[-7.90, -5.46]	<0.001
grouping [Non-SpEd Some] * season [Winter]	-18.33	[-31.28, -5.38]	0.006	-4.32	[-5.54, -3.10]	<0.001
grouping [SpEd High] * season [Winter]	-6.72	[-19.67, 6.23]	0.309	-2.04	[-3.26, -0.83]	0.001
grouping [Non-SpEd Low] * season [Spring]	-31.21	[-44.16, -18.26]	<0.001	-6.79	[-8.01, -5.57]	<0.001
grouping [Non-SpEd Some] * season [Spring]	-18.69	[-31.64, -5.74]	0.005	-4.30	[-5.52, -3.08]	<0.001
grouping [SpEd High] * season [Spring]	-3.53	[-16.48, 9.42]	0.593	-1.34	[-2.56, -0.13]	0.031
Random effects	*SD*	*SD CI*	Variance	*SD*	*SD CI*	Variance
Student	20.74	[17.73, 23.70]	430.00	1.53	[1.24, 1.80]	2.33
Residual	31.34	[29.66, 32.79]	982.28	2.95	[2.79, 3.09]	8.71

Note. *Est* = Parameter estimate; *CI* = 95% Confidence interval; and *SD* = Standard deviation.

Table 9

Pairwise Comparisons for PRF by Season and Instructional Group

Contrast	Season	*Est*	*SE*	*df*	*t*	*p*
Non-SpEd High - Non-SpEd Low	F	-71.27	5.38	1056.83	-13.24	< 0.001
Non-SpEd High - Non-SpEd Some	F	-33.66	5.43	1046.36	-6.20	< 0.001
Non-SpEd High - SpEd High	F	16.84	5.41	1047.34	3.11	0.01
Non-SpEd Low - Non-SpEd Some	F	37.60	5.37	1062.99	7.01	< 0.001
Non-SpEd Low - SpEd High	F	88.10	5.44	1034.61	16.20	< 0.001
Non-SpEd Some - SpEd High	F	50.50	5.41	1050.30	9.34	< 0.001
Non-SpEd High - Non-SpEd Low	W	-36.05	5.38	1056.83	-6.70	< 0.001
Non-SpEd High - Non-SpEd Some	W	-15.33	5.43	1046.36	-2.82	0.02
Non-SpEd High - SpEd High	W	23.56	5.41	1047.34	4.36	< 0.001
Non-SpEd Low - Non-SpEd Some	W	20.73	5.37	1062.99	3.86	< 0.001
Non-SpEd Low - SpEd High	W	59.61	5.44	1034.61	10.96	< 0.001
Non-SpEd Some - SpEd High	W	38.89	5.41	1050.30	7.19	< 0.001
Non-SpEd High - Non-SpEd Low	S	-40.05	5.38	1056.83	-7.44	< 0.001
Non-SpEd High - Non-SpEd Some	S	-14.97	5.43	1046.36	-2.76	0.03
Non-SpEd High - SpEd High	S	20.37	5.41	1047.34	3.77	< 0.001
Non-SpEd Low - Non-SpEd Some	S	25.08	5.37	1062.99	4.67	< 0.001
Non-SpEd Low - SpEd High	S	60.42	5.44	1034.61	11.11	< 0.001
Non-SpEd Some - SpEd High	S	35.34	5.41	1050.30	6.54	< 0.001

Note. F = Fall; W = Winter; S = Spring.

Table 10

Pairwise comparisons for Vocabulary by season and instructional group

Contrast	Season	*Est*	*SE*	*df*	*t*	*p*
Non-SpEd High - Non-SpEd Low	F	-8.96	0.49	1059.74	-18.45	< 0.001
Non-SpEd High - Non-SpEd Some	F	-5.89	0.49	1050.78	-12.06	< 0.001
Non-SpEd High - SpEd High	F	0.72	0.49	1053.95	1.48	0.45
Non-SpEd Low - Non-SpEd Some	F	3.07	0.48	1063.51	6.32	< 0.001
Non-SpEd Low - SpEd High	F	9.68	0.49	1044.83	19.81	< 0.001
Non-SpEd Some - SpEd High	F	6.62	0.49	1054.95	13.59	< 0.001
Non-SpEd High - Non-SpEd Low	W	-2.28	0.49	1059.74	-4.70	< 0.001
Non-SpEd High - Non-SpEd Some	W	-1.57	0.49	1050.78	-3.22	0.01
Non-SpEd High - SpEd High	W	2.77	0.49	1053.95	5.69	< 0.001
Non-SpEd Low - Non-SpEd Some	W	0.71	0.48	1063.51	1.46	0.46
Non-SpEd Low - SpEd High	W	5.05	0.49	1044.83	10.33	< 0.001
Non-SpEd Some - SpEd High	W	4.34	0.49	1054.95	8.91	< 0.001
Non-SpEd High - Non-SpEd Low	S	-2.17	0.49	1059.74	-4.47	< 0.001
Non-SpEd High - Non-SpEd Some	S	-1.59	0.49	1050.78	-3.26	0.01
Non-SpEd High - SpEd High	S	2.07	0.49	1053.95	4.25	< 0.001
Non-SpEd Low - Non-SpEd Some	S	0.58	0.48	1063.51	1.19	0.63
Non-SpEd Low - SpEd High	S	4.24	0.49	1044.83	8.67	< 0.001
Non-SpEd Some - SpEd High	S	3.66	0.49	1054.95	7.52	< 0.001

Note. F = Fall; W = Winter; S = Spring.

Results Summary

For both PRF and Vocabulary, SpEd High group mean scores across Fall, Winter, and Spring were significantly different than other groups on most of the pairwise comparisons. Specifically, pairwise comparisons found SpEd High mean PRF scores

were significantly different (lower) than all other groups and, importantly, significantly lower than the Non-SpEd High group in Fall ($p = .01$), Winter ($p < .001$), and Spring ($p < .001$). Vocabulary pairwise comparisons for the SpEd High versus the Non-SpEd High found no significant difference in Fall ($p = .45$), but significant differences in Winter ($p < .001$) and Spring ($p < .001$).

CHAPTER IV

DISCUSSION

Results indicate that PRF and Vocabulary scores of third-grade students with SLD (SpEd High) differ from those of their grade-level peers, including students identified as high risk, but who are not receiving special education services. The SpEd High group in my sample had PRF group mean scores for Fall, Winter, and Spring that were significantly different than other groups, and vocabulary scores that were significantly different in Winter and Spring. In the following sections, I (a) review and summarize the statistical analyses presented in the previous chapter, (b) address limitations of my study, (c) connect my findings to previous research, (d) discuss the practical and policy implications of my findings, and (e) provide suggestions for future research.

Review of Findings

Below, I review results and implications for each research question, including (a) use of CBMs for applying risk labels, (b) effectiveness of RtI as a framework for preventing and responding to reading difficulties, and (c) reading achievement growth for students with SLD. My study contributes information about how students identified with SLD, using an RtI model, perform compared to their peers in an RtI framework.

CBMs for Risk Labels

One of the primary aims of an RtI framework is for identification of readers who are at risk, and further, to identify those who have SLD versus students who are at risk, but should not be identified as having SLD. Universal screening and progress monitoring have been established as viable means of identifying students with SLD, and for intervention decision-making (Ardoin, Witt, Connell, & Koenig 2005; Ihori & Olvera,

2015; Vaughn, Linan-Thompson, & Hickman 2003). However, a concern raised by Fuchs et al. (2004) was that the method of assessing responsiveness to instruction in these contexts could produce different prevalence rates of disability and different subsets of unresponsive children. My study relied on the easyCBM© overall risk score to sort students without SLD into three groups (high risk, some risk, and low risk) and compared their mean Fall – Winter – Spring scores to students identified with SLD who were identified as having a high risk of reading difficulties. My research showed that mean PRF and Vocabulary scores and percentiles for each group support that CBMs can likely be a reliable indicator of risk and can be used to accurately group (classify) students into tiered interventions for individualized instruction.

Specifically, as shown in Table 5 the SpEd High and Non-SpEd High groups can be differentiated by Fall mean scores and percentile, with the SpEd High group having a mean PRF score of 25.63 wcpm and a mean percentile rank of 4.98 wcpm while the Non-SpEd High group had a mean PRF score of 41.28 wcpm and a mean percentile rank of 11.58 wcpm. On Vocabulary, the SpEd High group had a mean score of 8.11 words and a mean percentile rank of 8.93 words compared to the Non-SpEd High group that had a mean score of 8.78 words and a mean percentile rank of 11.23 words. As expected, Non-SpEd Some had higher PRF and Vocabulary mean scores than Non-SpEd High, and Non-SpEd Low had the highest mean scores of all groups. To summarize, the easyCBM© designated groups performed as expected in relation to each other. Importantly, easyCBM© measures can meet multiple RtI framework objectives including identification and progress monitoring for students at risk for reading difficulties (Alonzo & Anderson,

2018), which can inform instructional decision-making and contribute to identification of students with SLD.

Fuchs and Fuchs (2017) found that although most RtI models anticipate that 20-25% of students require Tier 2 intervention, if classrooms incorporate (a) strong instructional strategies, (b) accommodations to make instruction accessible to virtually all students, and (c) behavior management systems that promote motivation and engagement, then Tier 2 intervention should be necessary for only 10-15% of students. In addition, Tier 3 should be clearly distinguishable and should be necessary for only 4-8% of students (Fuchs & Fuchs, 2017). These models follow the inoculation versus insulin argument made by Coyne et al. (2004). High-quality, targeted intervention at Tier 2 can act as an immunization against future reading difficulties, leaving only a small group of students who need the ongoing insulin-like effect of sustained intervention. These "insulin-dependent" students ultimately may require a referral for special education evaluation and possible SLD identification (Ardoin et al., 2005; Vaughn et al., 2003).

My findings support the tiered intervention recommendations described above. While all groups demonstrated growth in PRF and Vocabulary, the SpEd High group narrowed the gap in achievement between itself and the other groups. From Fall to Spring, the difference between SpEd High and Non-SpEd Low in PRF mean score dropped from 85.37 wcpm to 57.69 wcpm. The mean PRF percentile rank difference dropped from 62.55 wcpm to 35.07 wcpm. Figures 3 and 4 show those changes. The students in the SpEd High group experienced a condition better than an inoculation effect (an intervention that temporarily boosts achievement), and instead received an insulin effect that allowed for growth over the course of the school year that narrowed the

achievement gap. Figure 3 shows PRF scores by student group. Students in the SpEd High group demonstrated growth that narrowed the gap to the Non-SpEd Low group. Figure 4 shows Vocabulary scores by student group. Students in the SpEd High group demonstrated growth that narrowed the gap to the Non-SpEd Low group.

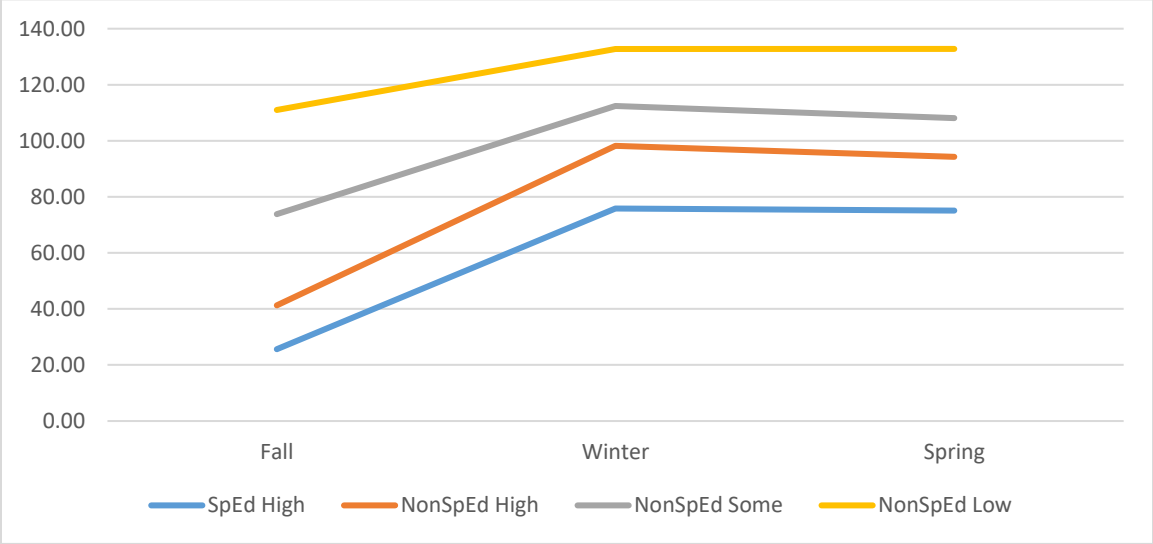

Figure 3. PRF mean scores by season and group.

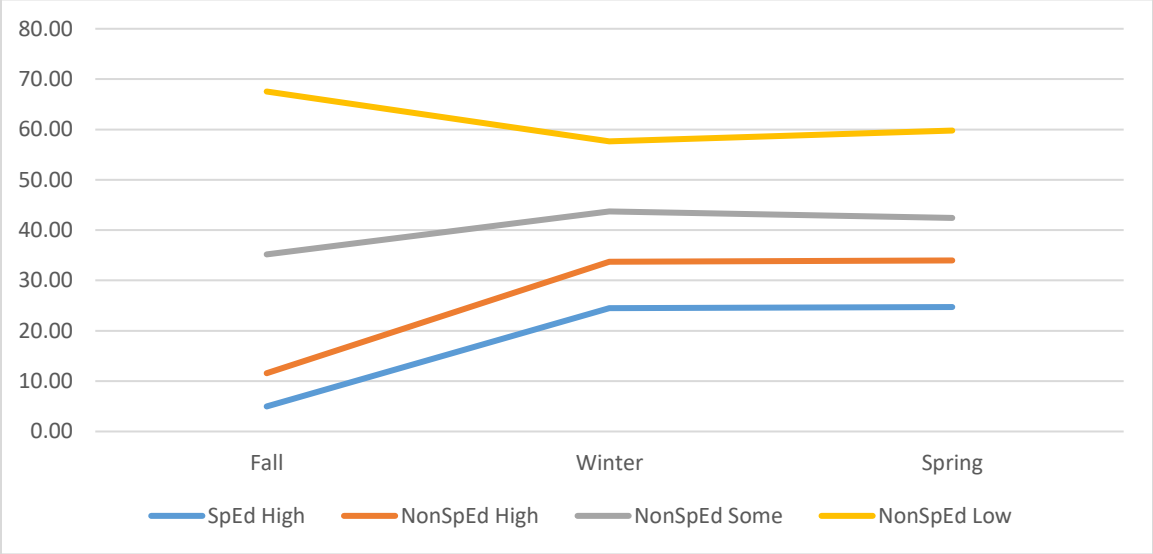

Figure 4. Vocabulary mean percentiles by season and group.

Figures 3 and 4 also speak to a need for a high-quality general curriculum that allows accommodations of individual need. Classroom teachers weak in reading instruction or without strong practices of differentiation and accommodations may seek support from the RtI model for students who may not otherwise need it, creating a bloated system (Fuchs & Fuchs, 2017). High-quality general education instruction is the first step in preventing a need for intervention. Schools must develop decision-making rules that allow access for students who need intervention and prevent students who do not need intervention from being placed inappropriately. In addition, Fuchs and Fuchs reminded us that high-quality general curriculum contributes to the success of students receiving Tier 2 and Tier 3 instruction. Students who are high-risk must still have access to the core reading instruction, as intervention is supplemental to the core instruction.

RtI as an Effective Framework

My results support that RtI can be an effective districtwide framework from an assessment perspective. As a reminder, The National Evaluation (2015) found that first-grade students assigned to receive RtI intervention on the basis of their screening performance achieved statistically significantly lower reading scores than students not assigned to receive intervention. In second grade, the impact was positive but not statistically significant, and at third grade, the impact was close to zero. These results suggested that RtI not only had no effect, but possibly a negative effect, raising the question of the overall effectiveness of an RtI model.

Importantly, my results may counter The National Evaluation (Balu et al., 2015) findings, as the SpEd High students in my study *did* demonstrate progress across time. In addition, mean PRF and Vocabulary scores and percentiles support that overall, the right

students were identified for each group. My study did not include observation or qualitative data related to the RtI framework or decision-making processes, so I can only surmise as to the reason why my results may counter Balu and colleagues' National Evaluation.

Fuchs and Fuchs (2017) cautioned that rather than interpreting The National Evaluation (Balu et al., 2015) results to suggest that RtI is an ineffective framework altogether, we should consider weak fidelity of implementation of RtI as the real problem. Despite the impact schools self-assessing full implementation of RtI, The National Evaluation noted multiple practices that differed from recommendations in previous RtI research literature. In contrast, my study found that an RtI framework *can* produce positive assessment results for all student groups, including students with SLD and students at high risk, as demonstrated by this district-wide examination. The results of my study add support to the argument shared by Fuchs and Fuchs (2017).

Districts and schools are responsible for ensuring the academic growth of all of their students, including those with SLD (NCLB, 2002). RtI has been put forth, with over two decades of research support, as an effective framework for monitoring progress and making data-based decisions for intervention. Fuchs and Fuchs (2017) identified four necessary components to a comprehensive and effective RtI framework:

1. Universal screening to identify students accessing only Tier 1 instruction, likely to experience poor academic outcomes, with progress monitoring for those students.

2. Tier 2 instruction involving academic programs proven in validation studies to be generally effective and delivered to small groups of students with similar needs.
3. Assessment of instructional responsiveness, conducted at each instruction tier, to assess whether a given student has responded adequately.
4. Tier 3 instruction that is intensive and individualized, with ongoing progress monitoring to gauge student responsiveness so that program adjustments can be introduced in a timely manner, and determine when students may move to a less intensive tier.

My research did not attempt to evaluate the district's full implementation of RtI as specified by Fuchs and Fuchs (2017). I only explicitly analyzed the first condition (universal screening results) as part of my study. For my study, third-grade students were identified and categorized into four groups based upon Fall Reading Risk: (a) SpEd High, (b) Non-SpEd High, (c) Non-SpEd Some, and (d) Non-SpEd Low. The Fall – Winter – Spring universal screening data showed that the district's tiered reading interventions seemed to benefit all students.

Importantly my study did not (a) evaluate Tier 2 academic programs (Condition 2), (b) assess individual students' instructional responsiveness, or (c) specifically assess student responsiveness to Tier 3 interventions. Fuchs and Fuchs (2017) stated that "these four RtI components are meant to represent a comprehensive and tightly woven framework of service delivery held together by an iterative test-teach-test process that starts with identification of an at-risk student in the general education classroom and ideally ends with the child back in the classroom but better prepared to do well" (p. 258).

For a more complete program evaluation, these components to high quality RtI frameworks must be examined. The limited focus of my research is a limitation.

Reading Achievement Growth

Third grade reading achievement is established as a predictor of academic achievement, emphasizing the critical work of early intervention in reading difficulties (Fiester, 2010). Of particular interest to this researcher, The National Evaluation (Balu et al., 2015) found that students with an IEP appear to have been affected by the treatment more negatively than other students. If, as Fuchs and Fuchs (2017) state, Tier 3 instruction should primarily be aimed at improving achievement for students with SLD, I was alarmed to discover the opposite being reported in the National Evaluation. My study examined easyCBM© scores for students with SLD across their third-grade school year and found students in the SpEd High group showed statistically significant positive effect from Fall to Spring. These findings confirm Fuchs and Fuchs belief that students with SLD can improve within a district that claims full implementation of the RtI framework. Of lingering concern is that while third grade students with SLD did, in fact, improve on easyCBM© measures and narrowed the gap in achievement, a discrepancy in achievement between students with SLD and other groups of students persists.

Fuchs and Fuchs (2017) claimed that students with SLD should be included throughout the entirety of the RtI framework. Including students with SLD in the universal screening measures and progress monitoring is not enough: Intervention for students with SLD should be considered as part of the RtI framework, not separate from it. This approach requires strong communication between regular education staff, intervention staff, and special education staff. Students with SLD benefit from the

frequent progress monitoring, instructional decision-making, and access to targeted intervention that at-risk students without SLD benefit from. Special education policy at the state and federal level allows for students with IEPs to receive their specially designed instruction within the framework of the RtI system. They do not have to be removed to a special education classroom specifically, as special education is meant to be a service rather than a place. Conversely, special education teachers should be part of the collaboration within the RtI framework, as their expertise can inform intervention strategies for students who are identified as at-risk but without the SLD label.

Limitations

In this section I discuss threats to validity and reliability of my study relative to study design, sample, instrumentation, data, and generalizability.

Internal Validity Concerns

Threats to the internal validity of my study exist including selection and instrumentation (Babbie, 2013). Selection is problematic because the participants in my study were selected through a convenience sample and are limited to one local school district. In an effort to reduce threats to selection validity, used stratified random sampling. I selected all qualifying participants from the special education group (n = 90), and randomly selected the same number of participants from the three non-special education groups to create equal-sized groups. In addition, participants were selected across four consecutive school years, to limit potential influence from initiatives or characteristics unique to a single school year.

Instrumentation is problematic and represents another threat to internal validity because of easyCBM© administration. First, staff members across the district, different

between schools and even within schools, administered easyCBM© assessments and manually entered results into the system. No checks for interrater reliability for assessment or data entry were available. Students might have been assessed by different administrators for each of the benchmarking periods. The data I analyzed was extant, and I was not involved in the administration or oversight of the assessments. Thus, I did not have personal knowledge or control over the fidelity of assessment administration or data entry. Secondly, the district's *overall risk designations* on easyCBM© included three measures: (a) PRF, (b) Vocabulary, and (c) Multiple Choice Reading Comprehension (MCRC). However, my statistical analysis *did not include MCRC* because of (a) the sub-measure's lower reliability and validity concerns (Nese, Anderson, Irvin, Alonzo, 2018) and its reported lower diagnostic utility and variance amongst students' groups (Anderson et al., 2014) and (b) because the district's third grade reading goal focusing on fluency. Had this MCRC measure been included, the results might have been different.

An additional threat to internal validity may be incomplete or inaccurate student demographic data. Demographic data were collected when parents enrolled their children in school, and then entered into Synergy by school staff. Importantly, demographic data were self-reported by parents, and as such may be subject to error through omission or mistake. In addition, data were subject to data entry errors by school staff.

External Validity Concerns

External validity refers to the ability of a study to be generalized to other settings and participants (Babbie, 2013). I remind readers that I utilized a non-randomized convenience sample. Other school districts who have similar policies regarding identification of students and implementation of RtI may find my results informative.

However, the results will not directly transfer, as relevant factors not identified in this study will not be the same, such as curriculum, teacher training, special education supports, and intervention schedules. The results of this study are specific to the school district that provided the data, the population of students served, and district policies. The use of standardized assessment and collection of self-reported demographic characteristics, however, does allow other school districts to look for possible similarities between themselves and the participating school district and examine their own data for similar achievement results relative to their own RtI framework.

The results of the study were reliant on a data set from a single school district and may be an artifact of the identification and assessment policies specific to that district. Results cannot be used as the final word in SLD identification practices but are specific to the district. The study and results may be of interest to school systems designing a similar analysis of achievement trajectories, but the results from this specific study are unique to the district. Further, data were analyzed at a district level, and results should not be applied to the individual student level.

Future Research

My study provides evidence in support of the RtI approach for using CBMs to identify students as SLD, apply risk-labels, and monitor student progress; however, it cannot be considered the final word. Further research is needed for the results of this study to be applied in other school districts, to improve the quality of RtI implementation, to understand the long-term effects of the RtI framework on achievement of at-risk students (both with and without SLD), and to understand how individual student characteristics interact with the RtI framework to influence achievement results.

Replication in Other Districts

Other districts may find the design and results of my study useful when considering the achievement outcomes of their own students within their district's RtI framework, or when considering adopting RtI and planning initial implementation. The participating school district has demonstrated growth with at-risk students by using CBMs to determine risk, monitor progress, and for SLD identification, but similar results cannot be assumed by other school districts without replicating the study with data from those districts. Replication would both inform individual school districts and, also, add to the body of research available in understanding the National Evaluation's results in contrast to previous research and the Fuchs and Fuchs' (2017) critique.

Qualitative Examination

When considering discrepant findings between the National Evaluation (Balu et al., 2015) and my study, it is important to heed Fuchs and Fuchs' (2017) assertion that success or failure of RtI ultimately relies on fidelity of implementation. Thus, future evaluation of qualitative elements of RtI are suggested. My study examined student data at the district-level. The inclusion of rating scales at each of the district schools to identify level of implementation of RtI could further clarify the results of my study or some other study. Balu et al. (2015) identified practices of the impact schools they studied that did not follow common RtI practices from previous literature, including scheduling of intervention, staff allocation for delivering intervention, and use of cut scores for determining intervention participants. The degree to which the schools within the district participating in my study did or did not have similar practices could clarify how those practices may ultimately affect student achievement. Other qualitative factors

not considered in my study that could inform RtI framework success are the intervention curriculum used, level of teacher training and/or experience, how the RtI framework fits into multi-tiered systems of support, generally, professional development offered, and the frequency and extent of collaboration between regular and special education professionals. By understanding the quality of these aspects of RtI implementation, the district could understand practices that may be affecting student achievement and identify leverage points for improvement and professional development opportunities.

Fuchs and Fuchs (2017) stated that an effective RtI framework relies on high quality Tier 1 instruction and supports. Further qualitative examination could evaluate the fidelity of implementation and efficacy of the Tier 1 program of the participating school district. Simmons et al. (2002) described eight tenets of a schoolwide beginning reading improvement model:

1. Addresses reading success and reading failure from a schoolwide *systemic* perspective;
2. Embraces a *prevention framework* by intervening early and strategically during the critical window of instructional opportunity (National Research Council, 1998);
3. Recognizes and responds to the multiple contexts of reading achievement and includes carefully articulated goals, research-based programs, dynamic assessment, adequate and protected instructional time, quality instructional delivery, differentiated instruction, and effective organization and grouping (Editor, 1998);

4. Develops and promotes a *comprehensive system of instruction* based on a research-based core curriculum and enhancement programs (Editor, 1998);
5. Anchors instruction and practices to the *converging knowledge* base of effective reading practice (National Research Council, 1998);
6. Builds capacity in the school by *using school-based teams* to customize interventions to the host environment;
7. Relies on and fosters *the school principal* to serve as instructional leader;
8. Uses ongoing tests sensitive to changes of student performance to identify students at risk, plan instruction groups, and modify instruction according to levels and rates of learning (Good, Simmons, & Smith, 1998); and
9. A qualitative analysis of these eight conditions at the district elementary schools could provide a granular analysis of the academic achievement growth students experienced within the districtwide RtI framework.

Longitudinal Studies

My study was limited in scope to only third graders and easyCBM© achievement monitoring. Students at third grade with SLD eligibility have only experienced a small amount of time with special education services relative to their whole school career. A longitudinal study on the academic achievement of students with SLD and other non-special education high-risk readers, including both easyCBM© and other achievement indicators such as state assessment scores, credit completion in high school, graduation rates, and post-secondary success would provide information on the potential long-term effects of early grade RtI frameworks. This would allow for developing strong policy for implementation of RtI frameworks across all grades of the K-12 system.

Student Specific Characteristics

Left out of my study is a fine grain look at student-level factors and characteristics that may interact with the RtI framework to predict achievement. While I collected demographic data, I did not conduct statistical analysis to understand how factors such as race/ethnicity, sex, poverty, attendance, and English Language Proficiency may or may not influence achievement. These factors, if found to be significant, could inform how educators plan for student specific characteristics within their core instruction or RtI frameworks. Understanding interactions between demographic factors and instructional group would allow the district to proactive address issues of equity and disproportionality through policy and practice.

Conclusion

Over the past 15 years, students with disabilities, including those with SLD, have seen an increase in participation in the regular education environment. Simultaneously their achievement results are under closely scrutiny due to federal policy that holds schools, districts, and states accountable for achievement outcomes for students with disabilities. RtI was intended to be an early warning system for identifying students at risk for reading struggles. It provides intervention for those struggling readers, and when decision-making rules are established, it can determine when students move in and out of intensive intervention support. Finally, it can be used to identify students with SLD. Despite the 20-year history of RtI, recent studies found mixed results on its achievement outcomes for students (Balu et al., 2015; Fuchs & Fuchs, 2017).

My research sought to clarify previously conflicting results by examining CBM scores for students with SLD in a district that has adopted and implemented systematic

RtI processes. It provides evidence that RtI Frameworks can support a system through which students with SLD experience academic growth compared to their peers. The district studied demonstrated that overall appropriate groups of students were identified through RtI as having SLD. Once identified, those same students demonstrated growth in standardized reading measures over the course of a full school year.

Practitioners and policy makers at the district level should find my study useful for two reasons. First, it provides evidence that an RtI system should include students with SLD in progress monitoring and intervention decision-making processes. General and special education professionals should work together to meet the needs of all learners, rather than operating as two separate entities. Second, students with SLD still need access to the core reading curriculum used in the general education classroom. Intervention should supplement the core reading program by targeting specific skill deficits, rather than replace it.

REFERENCES CITED

Alonzo, J. & Anderson, D. (2018). Classification accuracy of the easyCBM kindergarten – Grade 2 reading measures (Technical Report 1801). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Tindal, G., Ulmer, K., & Glasgow, A. (2006). easyCBM online progress monitoring assessment system. Eugene, OR: Behavioral Research and Teaching.

Anderson, D., Alonzo, J., Tindal, G., Farley, D., Irvin, P. S., Lai, C. F., & Wray, K. A. (2014). Technical manual: easyCBM (Technical Report 1408). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Ardoin, S.P., Witt, J.C., Connell, J. E., & Koenig, J. L. (2005). Application of a three-tiered response to intervention model for instructional planning, decision making, and the identification of children in need of services. *Journal of Psychoeducational Assessment, 23*, 362-380. doi:10.1177/073428290502300405

Babbie, E. (2013). *The Practice of Social Research (13th ed.)*. Belmond, CA: Wadsworth Cengage Learning.

Balu, R., Zhu, P., Doolitle, F., Schiller E., Jenkins J., & Gersten, R. (2015). Evaluation of response to intervention practices for elementary school reading (NCEE 2016-4000). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Catabay, R. (2017). Individuals with Disabilities in Education Act. Retrieved March 18, 2017, from http://www.2.ed.gov/about/offices/list/osers/osep/osep-idea.html

Cortiella, C. & Horowitz, S. (2014). The state of learning disabilities. New York: National Center for Learning Disabilities (NCLD).

Coyne, M., Kame'enui, E. J., Simmons, D. C., & Harn, B. A. (2004). Beginning reading intervention as inoculation or insulin: First-grade reading performance of strong responders to kindergarten intervention. *Journal of Learning Disabilities, 37*, 90-104. doi:10.1177/00222194040370020101

Editor. (1998). Every child reading: An action plan of the Learning First Alliance. *American Educator, Spring/Summer*, 52-63.

Fiester, L. (2010). Early warning! Why reading by the end of third grade matters: Kids Count Special Report. Baltimore: Annie. E. Casey Foundation.

Fuchs, D. & Fuchs L. S. (2017) Critique of the national evaluation of response to intervention: A case for simpler frameworks. *Exceptional Children, 83*, 255-268. doi:10.1177/0014402917693580

Fuchs, D., Fuchs, L. S., & Compton, D. L. (2004). Identifying reading disabilities by responsiveness-to-instruction: Specifying measures and criteria. *Learning Disability Quarterly, 27*, 216-227. doi:10.2307/1593674

Fuchs, D., Fuchs, L. S., & Stecker, P. M. (2010). The "blurring" of special education in a new continuum of general education placements and services. *Exceptional Children, 76*, 301-323. doi:10.1177/001440291007600304

Good, R., III, Simmons, D. C., & Smith, S. (1998). Effective academic interventions in the United States; Evaluating and enhancing the acquisition of early reading skills. *School Psychology Review*, *27*, 45-56.

Harn, B. A., Chard D. J., Biancarosa, G., & Kame'enui, E. J. (2011). Coordinating instruction supports to accelerate at-risk first-grade readers' performance: An essential mechanism for effective RtI. *The Elementary School Journal*, *122*, 332-355.

Ihori, D. & Olvera, P. (2015). Discrepancies, responses, and patterns: Selecting a method of assessment for specific learning disabilities. *Contemporary School Psychology, 19*, 1-11. doi:10.1007/s40688-014-0042-6

Individuals with Disabilities in Education Act, 20 U.S.C. § 1400 (2004).

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

McLeskey, J., Landers, E., Willimson, P., & Hoppey, D. (2012). Are we moving toward educating students with disabilities in less restrictive settings? *The Journal of Special Education, 46*, 131-140. doi:10.1177/0022466910376670

McLeskey, J. & Waldron, N. L. (April, 2011). *Full inclusion programs for elementary students with learning disabilities: Can they meet student needs in an era of high stakes accountability?* Paper presented at the Council for Exceptional Children Convention, National Harbor, MD.

National Center for Education Statistics. (2020). Education demographic and geographic estimates. Retrieved January 1, 2020, from https://nces.ed.gov/programs/edge/acsdashboard

National Center on Response to Intervention. (2010). Essential components of RtI – A closer look at response to intervention. Washington, DC: U.S. Department of Education, Office of Special Education Programs, National Center on Response to Intervention.

National Research Council. (1998). Preventing reading difficulties in young children. Washington, DC: National Academy Press.

Nese, J. F. T., Anderson, D., Irvin, P. S., & Alonzo, J. (2018). In-Brief: Reliability of the slope of the easyCBM reading measures (Technical Report No. 1803). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

No Child Left Behind Act of 2002, P.L. 107-110, 20 U.S.C. § 6319 (2002).

Oregon Department of Education. (2016). Oregon Statewide Report Card. Retrieved from http://www.oregon.gov/ode/schools-and-districts/reportcards/Documents/rptcard2016.pdf

Oregon Department of Education. (2017). Oregon Statewide Report Card. Retrieved from https://www.oregon.gov/ode/schools-and-districts/reportcards/Documents/rptcard2017.pdf

Oregon Department of Education. (2018). FFY 2016 Part B State Performance Plan (SPP)/Annual Performance Report (APR). Retrieved from https://www.oregon.gov/ode/reports-and-data/SpEdReports/Documents/2016partbffy.pdf

Oregon Rules for Special Education, OAR 581-015-2120 (2013).

Simmons, D. C., Coyne M. D., Oi-man K., McDonagh, S., Harn, B. A., & Kame'enui, E. J. (2008). Indexing response to intervention: A longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilities*, *41*, 158-173.

Simmons, D. C., Kame'enui, E. J., Good, R. H., Harn, B. A., Cole, C., & Braun, D. (2002). Building, implementing, and sustaining a beginning reading model: Lessons learned school by school. In M. R Shinn, H. M. Walker, & g. Stoner (Eds.), *Interventions for academic and behavioral problems II: Preventive and remedial approaches* (pp. 403-432). Bethesda, MD: National Association of School Psychologist.

Spencer, M., Wagner, R. K., Schatschneider, C., Quinn, J. M., Lopez, D., & Petscher, Y. (2014). Incorporating RTI in a hybrid model of reading disaiblity. *Learning Disability Quarterly*, *37*, 161-171. doi:10.1177/0731948714530967

U.S. Census Bureau. (2020). Quick facts. Retrieved from https://www.census.gov/quickfacts/fact/table/US/PST045219

U.S. Department of Education. (2010). Free and appropriate public education for students with disabilities: Requirements under Section 504 of the Rehabilitation Act of 1973. Washington, DC: U.S. Department of Education, Office for Civil Rights.

VanDerHeyden, A. M., Witt, J. C., & Gilbertson, D. (2007). A multi-year evaluation of the effects of a Response to Intervention (RTI) model on identification of children for special education. *Journal of School Psychology*, *45*, 225–256. doi:10.1016/j.jsp.2006.11.004

Vaughn, S., Linan-Thompson, S., & Hickman, P. (2003). Response to instruction as a means of identifying students with reading/learning disabilities. *Exceptional Children, 69,* 391-409.

Wanzek, J., & Vaughn, S. (2011). Is a three-tier reading intervention model associated with reduced placement in special education? *Remedial and Special Education, 32,* 167–175. doi:10.1177/0741932510361267

Wray, K. A., Alonzo, J., & Tindal, G. (2014). Internal consistency of the easyCBM vocabulary measures Grades 2-8 (Technical Report 1406). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Zirkel, P. A. (2013). The trend in SLD enrollments and the role of RTI. *Journal of Learning Disabilities, 46,* 473–479. doi:10.1177/0022219413495297