THE IDAHO READING INDICATOR AS A PREDICTOR OF SUBSEQUENT DIAGNOSIS OF SPECIFIC LEARNING DISABILITIES

by

Kimberly A. Ennis

A dissertation

submitted in partial fulfillment

of the requirements for the degree of

Doctor of Education in Curriculum and Instruction

Boise State University

May 2015

© 2015

Kimberly A. Ennis

ALL RIGHTS RESERVED

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the dissertation submitted by

Kimberly A. Ennis

Dissertation Title: The Idaho Reading Indicator as a Predictor of Subsequent Diagnosis of Specific Learning Disabilities

Date of Final Oral Examination: 19 March 2015

The following individuals read and discussed the dissertation submitted by student Kimberly A. Ennis, and they evaluated her presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Jack Joseph Hourcade, Ph.D.	Chair, Supervisory Committee
Roger Stewart, Ph.D.	Member, Supervisory Committee
Michael John Humphrey, Ed.D.	Member, Supervisory Committee
Keith W. Allred, Ph.D.	Member, Supervisory Committee

The final reading approval of the dissertation was granted by Jack Joseph Hourcade, Ph.D., Chair of the Supervisory Committee. The dissertation was approved for the Graduate College by John R. Pelton, Ph.D., Dean of the Graduate College.

ACKNOWLEDGEMENT

My sincerest appreciation is given to my wonderful advisor, Dr. Jack Hourcade. Dr. Hourcade was patient with me. More importantly, he kept me on track and moving forward. I am thankful that he was willing to serve as my advisor.

I am also very grateful to Dr. Roger Stewart. Dr. Stewart was kind and generous with his time. I was inspired by his courses on assessments.

I valued the suggestions and knowledge that Dr. Humphrey and Dr. Keith Allred contributed to this dissertation. I thank Dr. Humphrey and Dr. Allred for the time and effort that they dedicated to the supervisory committee.

The dissertation would not have been completed without the help of Stephanie Lee, Assessment Specialist, Idaho State Department of Education, and Steve Tyree, Testing and Assessment Administrator, Boise School District. I hope that the information collected will be a benefit. Finally, thank you Jodi Chilson, Coordinator of Theses, Dissertations, & Fellowship Programs for putting the dissertation in the proper format.

Special thanks to my friends and family. Your encouragement and faith in me kept me motivated.

ABSTRACT

The Idaho Reading Indicator (IRI) is a universal screen used to quickly assess the reading fluency of all Idaho public school students in grades kindergarten through third. The IRI is a tool that aids in the early identification of children with potential reading difficulties. Within the group of children with reading difficulties, some students have specific learning disabilities (SLD). Early identification of reading difficulties is critical to the success of all students who struggle with reading, especially students with SLD. To understand the problem, it is important to understand the nature of reading and the consequences of reading success and failure. The identification of SLD is not an easy task and differences in diagnostic approaches exist. The diagnostic value of the fluency assessment techniques employed by the IRI is examined.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BSD ……………………………….Boise School District

CBM ………………………………Curriculum Based Measurement

ESEA ……………………………..Elementary and Secondary Education Act

IDEA ………………………………Individuals with Disabilities Act

IRI …………………………………Idaho Reading Indicator

IQ ……………………………….....Intelligence Quotient

NCLB ……………………………..No Child Left Behind

NICHHD …………………………National Institute of Child Health and Human Development

OR …………………………………Odds Ratio

R-CBM ……………………………Reading Curriculum Based Measurement

SDE ……………………………….State Department of Education

SLD ……………………………….Specific Learning Disability

CHAPTER 1: INTRODUCTION

The acquisition of reading skills is one of the most fundamental achievements of students in K-6 school programs (Committee on the Prevention of Reading Difficulties in Young Children, 1998). Students who are proficient readers are more likely to experience academic success (Committee on the Prevention of Reading Difficulties in Young Children, 1998; Feitelson, Goldstein, Iraqi, & Share, 1993; Paris, 2005; Pretorius, 2000). In contrast, students who experience reading difficulty are more likely to struggle in school. The “Matthew Effect” in reading refers to the Biblical concept of the “poor getting poorer” and the “rich getting richer” (Stanovich, 1986). The Matthew Effect illustrates the sharp contrast between proficient readers and poor readers. Good readers read more and become even more proficient while poor readers continue to lag far behind (Cunningham & Stanovich, 2001; Taylor, Frye & Maruyama, 1990). Poor readers rarely are able to make enough gains in reading to become academically proficient or successful (Torgesen, 1998).

It is critical that students with potential problems in reading are identified at an early age. In order to turn poor readers into competent readers, students who are identified as having difficulty in reading should receive interventions sooner rather than later in their academic careers (Vellutino, Scanlon, Small, & Fanuele, 2006). Students with specific learning disabilities (SLD) are of particular concern because they typically are likely to require specialized reading instruction. Students with SLD often are not identified until they are ages 8 or 9 (Lyon, 1996). It is often challenging to separate students with SLD from other types of students who are poor readers (Fletcher et al., 1994; Fuchs, Fuchs, Mathes, Lipsey, & Roberts, 2002). This is due to in part to the fact that poor reading performance may stem from a number of causes, including (a) speaking a primary language that is not English, (b) low intelligence, and (c) SLD (Fletcher et al., 1994; Fuchs et al., 2002). Another cause of reading failure may be the result of poor instruction (Committee on the Prevention of Reading Difficulties in Young Children, 1998; Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998).

Our society and political leaders have long recognized the importance of reading. Nationally, with the passage of the No Child Left Behind Act of 2001 (NCLB), attention has been drawn to school performance as measured by student achievement. States have systems of measuring student performance. In Idaho, the assessment used to measure student performance in reading for grades one through three is the Idaho Reading Indicator (IRI). It is surprising how little research has been done on the IRI. It is of specific interest to consider the IRI in relation to SLD. One potential way to understand this relationship is to compare IRI scores to the subsequent identification of SLD within a group of students. To date, there does not appear to be a study that has examined the relationship of statewide reading achievement scores and the identification of SLD. This proposed study seeks to examine the relationship between student performance on the IRI and the subsequent identification of SLD within a specific population of students. Since the IRI is designed to identify students at-risk for reading failure, one hypothesis is that low IRI scores will have a higher correlation with SLD. Conversely students with higher IRI scores are less likely to be identified as SLD.

Background

As stated earlier, the study is concerned with the relationship between the IRI and the subsequent identification of SLD. To understand the key variables of this study, a brief overview of the IRI is given to provide historical context as well as a description of the IRI. It is also necessary to explain the different approaches to identifying students as SLD.

The first variable in the study is the IRI scores of a cohort of first grade students in 2009 and their subsequent IRI scores through third grade. The IRI is state-mandated reading assessments for students in the state of Idaho who are in grades one through three (Idaho Comprehensive Literacy Act, Idaho Code 33-1614). The IRI was first used in 1999. The IRI is a universal screen designed to assist in the identification of students with potential reading problems and an outcomes measure because the IRI is used to determine if students are meeting grade-level goals as compared with peers (Fisk, n.d.). The kindergarten IRI measures Letter Name Fluency and Letter Sound Fluency. The first grade IRI contains Letter Sound Fluency and oral reading fluency measurements (R-CBM). The second and third grade IRIs contain only R-CBMs. The administration of the IRI is required by law to take less than 10 minutes (Idaho State Department of Education, 2013). The IRI is also a high-stakes assessment. The results of the IRI are used as a measurement of student reading achievement and gains in reading proficiency. There have been three versions of the IRI.

The second variable is the subsequent identification of students as SLD. Within the time period from 2009 to 2013, this study proposes to examine the relationship between IRI scores and later identification of SLD. The method of identifying students as

SLD changed in 2010. In general, two methods of labeling students as SLD are currently used across the country. They are (a) the traditional approach (the intelligence quotient (IQ)-discrepancy model) and (b) the contemporary Response to Intervention (RTI). The reauthorization of the Individuals with Disabilities Education Improvement Act (IDEA) of 2004 included RTI as an alternative means of identifying students with SLD (Berkeley, Bender, Gregg Peaster, & Saunders, 2009). Now states may use the IQ-discrepancy model, RTI, or a combination of them. However, clearly there is a move nationally as well as in the state of Idaho to adopt the RTI model in the identification of SLD.

The traditionally used IQ-discrepancy model analyzes the gap between a student's IQ and academic achievement scores on standardized assessments. A student is labeled SLD when her performance on an IQ test is within the average or above average range and her performance on an academic achievement measure is significantly below the IQ score (Fletcher, Denton & Francis, 2005; Fletcher et al., 1994; Mercer, Jordan, Allsopp, & Mercer et al., 1996; Stanovich, 1991a). In Idaho, a 15 point discrepancy was required to apply the label of SLD to students referred for consideration of special education services. Under the traditional approach, a student with an IQ of 100 would have to have a score of 85 or below on an academic achievement test to meet eligibility requirements for a diagnosis of SLD.

The contemporary approach towards the diagnostic process in SLD is known as response to intervention, or RTI. The philosophical foundation of RTI is to ascertain whether the student's deficits in reading are due to (a) the quality or nature of the reading instruction being provided, or (b) some inherent issue with the student's ability to acquire the unique neurological skills required in effective reading. To this end, RTI begins in the general education classroom, with students of concern being provided scientific researched-based interventions in reading over an adequate period of time (eight to twelve weeks). Before students may be referred for a SLD evaluation, the student must have demonstrated significant resistance to general education interventions during this period, with an actual rate of learning lower than reasonably expected.

Both the IQ-discrepancy and RTI models of SLD identification have advocates and critics. Another emergent diagnostic approach, the blended approach, combines elements of RTI and an assessment of cognitive processes (Barth et al., 2008; Baskette, Ulmer, & Bender, 2006; Batsche, Kavale & Kovaleski, 2006; Bradley, Danielson, & Hallahan, 2002; Fletcher, Coulter, Reschly, & Vaughn, 2004; Fuchs, Fuchs, & Stecker, 2010; Kavale & Spaulding, 2008; Keogh, 2005; McKenzie, 2009). The cognitive processes related to reading are working memory, processing speed, executive function, and receptive and expressive language (Johnson, Humphrey, Mellard, Woods, & Swanson, 2010). Students with SLD performed significantly lower on cognitive processing tasks than their typically achieving peers (Johnson et al., 2010).

Many have proposed that such a blended approach will lead to more accurate and effective identification of students with SLD (Fletcher et al., 2005; Fuchs, Deshler, & Reschly, 2004; Ofiesh, 2006). In 2010, the Idaho State Department of Education (SDE) implemented such a blended approach (RTI plus assessment of cognitive processes) for identifying students as SLD. Prior to August of 2010, the IQ-Discrepancy model was in effect.

Problem Statement

The need to identify poor readers and to provide interventions is urgent. The stakes are high for students, schools, and for society. The consequences of reading failure cannot be ignored. Failure to reach reading proficiency levels may have devastating results on a student's life. NCLB ushered in an age of accountability testing. The failure to meet proficiency levels may result in penalties against schools. With the importance placed on testing, stakeholders need to be more informed about the assessments that are used in high stakes testing.

In Idaho, the test that is used to measure the reading proficiency of students in grades kindergarten through third is the IRI. The IRI has been used for approximately fourteen years. The IRI is designed to be quick to administer and easy to score (Idaho State Department of Education, 2013). The IRI is a universal screen that can only provide a "snapshot" of a particular student's reading achievement.

The first grade IRI combines Letter Sound Fluency with oral reading fluency (R-CBM) measures. In the measure of Letter Sound Fluency, students make the sounds for as many of the letters displayed on a sheet of paper in one minute. Oral fluency is measured by counting the number of words in a passage read correctly in one minute. The same three R-CBM passages are given in the fall and spring. The middle score (not the highest nor the lowest reading rate) is used to determine reading proficiency levels. The only method of assessment on the second and third grade IRI is the R-CBMs. The second and third grade R-CBMs are administered and scored with the same methods described in the first grade IRI.

There are few available empirical studies of the IRI. One study considered how well the IRI could predict performance on Idaho Standards Achievement Tests (ISATS; Stewart, 2009). The ISATS is the high stakes academic achievement assessment of reading, language usage, and mathematics for grades three through tenth grade. Stewart (2009) found that the IRI correctly predicted student performance for 84% of the students. The strongest correlation between the IRI and the ISATS was among white students who were not enrolled in special education or Title I schools.

Two studies (Nave & Burke, 2007a, 2007b) were commissioned by the state to compare the first version of the IRI (1999 – 2007) with the proposed replacement version of the IRI (2007 – 2009). The studies compared proficiency ratios between students on both versions of the IRI in the fall and winter. The two studies found similar patterns of student performance on both versions of the IRI.

A more recent study by Santi and Francis (2012) was commissioned by the Idaho State Department of Education to review the IRI and to assess current practices related to the identification of students at risk for reading problems in the early grades. The authors concluded that the IRI's use as a universal screen to identify students at risk for reading failure and as a measurement of teacher accountability were at cross purposes. The IRI as a measure of teacher performance hinders the use of the IRI as a means to identify students with potential reading difficulties. The limited psychometric information associated with the IRI was also a concern.

The IRI is a measure of reading proficiency. The IRI identifies students who are potentially at risk for reading failure. Within the population of students at risk for reading failure, how many of those students are found to be SLD? To date, there has not been a study that examines the correlation between IRI scores and the subsequent identification of students as SLD.

Purpose of the Study

The purpose of this study is to examine the relationship between IRI scores and the subsequent identification of SLD. The fall and winter IRI scores of students who entered first grade in 2009 will be collected from grades one through three. The IRI scores will be correlated to the subsequent identification of SLD by the end of the first semester in fourth grade within the group of these students.

Research Questions and Hypothesis

The following questions will guide the research as the relationship between IRI scores and the subsequent identification of SLD is explored.

Question 1: What is the correlation between Letter Sound Fluency scores by students on the first grade fall IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

Question 2: What is the correlation between Letter Sound Fluency scores by students on the first grade spring IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

Question 3: What is the correlation between R-CBM scores by students on the first grade fall IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

Question 4: What is the correlation between R-CBM scores by students on the first grade spring IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

Question 5: What is the correlation between R-CBM scores by students on the second grade fall IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

Question 6: What is the correlation between R-CBM scores by students on the second grade spring IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

Question 7: What is the correlation between R-CBM scores by students on the third grade fall IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

Question 8: What is the correlation between R-CBM scores by students on the third grade spring IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

Procedures

The most recent edition of the IRI was introduced in 2009. The performance of students who entered first grade in 2009 will be the first variable in this study. The IRI scores of the students who were in first grade in 2009 will be collected across grades one through three. The second variable is the subsequent identification of SLD within the group of students who entered first grade in 2009. The data will be collected from the

Boise School District (BSD). The BSD is the second largest school district in the state and it is considered an urban school district.

The first grade IRI from the fall of 2009 and spring 2010 data will be collected. The first grade IRI contains two separate measures: (a) Letter Sound Fluency and (b) Reading Curriculum Based Measures (R-CBM). The second grade 2010 fall IRI scores and 2011 spring IRI scores will be collected. The third grade 2011 fall IRI and the 2012 spring IRI scores will be collected. All second and third grades IRIs are R-CBMs.

Student identification numbers and IRI scores will be paired. The lists of student identification numbers of students identified as SLD from first through fourth grade will be collected from 2009 through 2013. Students identified as SLD will be coded as “one.” Students not identified as SLD will be coded as “zero.” Three types of statistical analyses were conducted for this study. First, a logistic regression analysis was used to determine the relationship between IRI scores and the identification of SLD at any point between the years of 2009 to 2013. Secondly, a Pearson chi-square was used to further examine phenomenon that might exist in the data. The Cramer’s V statistic was used to study the magnitude of the relationships between the variables. The expected and actual count was used to measure the gap between the expected count (the number if there was no relationship between IRI scores and the subsequent identification of SLD) and the actual count for each cell in the 2 X 3 table. Finally, an odds ratio was calculated to demonstrate the relationship between IRI scores and the subsequent identification of SLD.

Limitations of the Study

There are several potential limitations of this study. To begin with, the sample may be too small to make extensive generalizations. Next, the Boise School District student body may not mirror the national population in potentially significant ways, including ethnic diversity. Another potential limitation to the study is that there was a shift in the way students are identified as SLD during the time period in question. Specifically, in 2010 Idaho changed from the IQ-discrepancy model to a blended approach. The ways that students with SLD are identified is given in Chapter 2: Review of the Literature.

Furthermore, the time frame of this study may be a limitation. The study ends when the students finish fourth grade. The potential to be identified as SLD is present until graduation (or even beyond). These "late diagnoses" would not be identifiable in this study as presently structured. Finally, at the district level, students are listed as having SLD but the type of SLD is not specified. The IRI is a measure of reading achievement. There are students who are labeled as SLD who have academic difficulties in other areas (e.g., math) but not reading. However, approximately 80% to 85% of students with SLD have deficits in reading (Lyon, 1996).

Significance of the Study

The acquisition of reading skills is arguably the most important academic function of schools. Given this, it is critical that students with potential reading difficulties, including those with SLD, be identified and provided appropriate interventions as early as possible.

Unfortunately, historically diagnostic decisions concerning the presence of a learning disability were made only after the student had exhibited several academic years of literacy struggles and were then referred for possible evaluation for SLD. Earlier identification of reading difficulties can result in more favorable outcomes for students with reading difficulties (Boscardin, Muthen, Baker, & Francis, 2008; Snowling, 2012; Torgesen, 1998). This study will examine the predictive capability of an already existing and universally administered (in the state of Idaho) assessment instrument in identifying students who subsequently are found to have SLD. The Institutional Research of Board Boise State University granted approval for this research project and the protocol was considered exempt on April 7th, 2014.

CHAPTER 2: REVIEW OF THE LITERATURE

The ability to read is a fundamental skill necessary for Americans to fully participate in all aspects of our society (National Center for Family Literacy, 2008; Committee on the Prevention of Reading Difficulties in Young Children, 1998). Reading skills are invaluable in almost every arena in society. It can be argued that the necessity of functional reading skills is higher now than before. The ability to read is linked to social and economic advancement of the individual; thus, the success of our society depends on the reading ability of our citizens. The driving forces pushing the demand for expanded reading abilities include the international marketplace and rapid evolution of technology (Carnevale, 1991; Pretorius, 2000). Learning to read is arguably one of the most important expectations of students in primary elementary education (Committee on the Prevention of Reading Difficulties in Young Children, 1998; Johnson, Pool, & Carter, n.d.).

Reading, simply stated, is making meaning from print (Leipzig, January, 2001). Another way to describe the act of reading is by thinking of reading as turning written symbols into internal (and perhaps external) language (Walcutt, 1967). Defining reading is difficult because reading is a complex behavior utilizing word recognition, reading comprehension, reading fluency, and motivation (Bormuth, 1973-74; Leipzig, January, 2001). The act of reading requires the coordination of several cognitive processes. Reading is a psycholinguistic process, meaning that cognitive, psychological, neurobiological, and language development are needed to learn how to read (Committee on the Prevention of Reading Difficulties in Young Children, 1998).

Skillful readers have a firm language foundation, which enables them to understand the meaning of the text. Acquiring the skills required for reading begins with the child's language development. The roots of reading begin with the child's exposure to and acquisition of oral language, stories, and print awareness (Bear, Invernizzi, Templeton, Johnston, 1996; Committee on the Prevention of Reading Difficulties in Young Children, 1998). Oral language, stories, and print awareness are "threads" that bind together as children learn how to read (Bear et al., 1996). Good readers have strong receptive language, expressive language, phonological awareness, print awareness, decoding skills, and large vocabularies (Johnson et al., n.d.; Scarborough, 2009).

To read, individuals must use visual perception, visual discrimination, visual memory, recall, and directional orientation (Traub & Bloom, 1990). Reading combines attention, memory, language, and motivation (Adams, 1994; Committee on the Prevention of Reading Difficulties in Young Children, 1998). When we read, we map letters to sounds with visual-auditory integration (Traub & Bloom, 1990). Once mapped, the letter-sounds are synthesized into syllables, and the syllables are combined into words. Orthographic knowledge is the information stored in our memory and retrieved so that we can represent spoken language in written form (Apel, 2011). As orthographic knowledge builds, the braided bond becomes stronger and thickens as the reader's vocabulary expands (Bear et al., 1996).

A critical stage occurs when the reader shifts from sounding out words and word recognition to "reading fluency" or the ability to gain linguistic meaning from printed symbols quickly and accurately (Kuhn & Stahl, 2003; National Institute of Child Health and Human Development [NICHHD], 2000; Rasinski & Hoffman, 2003; Spear-Swerling, 2006; Wise et al., 2010). After developing word recognition, improving reading fluency is the next significant step for beginning readers (NICHHD, 2000). Skillful readers are able to comprehend words quickly, accurately, and read with expression (Adams, 1994; NICHHD, 2000; Committee on the Prevention of Reading Difficulties in Young Children, 1998; Wolf & Katzir-Cohen, 2001). Measuring reading fluency, or the ability to quickly read words in text with few errors, is a means to assess reading competency and achievement (Fuchs, Fuchs, Hosp, & Jenkins, 2001; Good, Simmons, & Kame'enui, 2001; Spear-Swerling, 2006; Wise et al., 2010; Wolf & Katzir-Cohen, 2001).

Reading comprehension is the ability to perceive, understand, and construct meaning from text (Wilhelm, 2015). In some ways, reading comprehension may present a greater challenge for learners than do word recognition and decoding, as comprehension requires more complex integration of word recognition and language skills (Kuhn & Stahl, 2003; Rasinski & Hoffman, 2003; Traub & Bloom, 1990). Interestingly, despite the fact that fluency and comprehension are different skills, many researchers have reported a strong relationship between a student's fluency and comprehension (e.g., Fuchs et al., 2001; Good et al., 2001; Spear-Swerling, 2006; Thurlow & van den Broek, 1997; Wise et al., 2010; Wolf & Katzir-Cohen, 2001). Reading comprehension and reading fluency both require accurate and automatic word identification, as well as oral language comprehension (Spear-Swirling, 2006; Wolf & Katzir-Cohen, 2001). Reading comprehension and reading fluency appear to have a reciprocal relationship (Pikulski & Chard, 2005; Spear-Swerling, 2006). The relationship is reciprocal because fluency may advance reading comprehension, and in turn reading comprehension may foster fluency (Pikulski & Chard, 2005; Pinnell et al., 1995).

Learning to read typically follows developmental stages, starting at the beginning of a child's life (Chall, 1967, 1983). In the pre-reading stage ("Emergent Reading" or Stage Zero), oral language, phonological awareness, and letter naming are developed during early childhood and kindergarten. Stage One ("Initial Reading or Decoding") spans grades one through two. In Stage One children learn the letters of the alphabet and develop sound-symbol associations. In Stage Two ("Confirmation, Fluency, or Ungluing from Print") occurs in grades two through three. In Stage Two, children learn more difficult words and read more elaborate stories. At the end of this period, students are "ungluing from print." They are more fluent readers and rely less on decoding. Stages One and Two are the "learning to read" period in a child's life.

The later stages are the "reading to learn" period in a student's education. In Stage Three, "Reading for Learning the New," children read to learn new information. Stage Three covers grades four through junior high. Stage Four, "Multiple Viewpoints," occurs approximately during high school. Students are able to reflect on what they read from multiple points of view. Stage Five starts around age 18 or older. The "Construction and Reconstruction" or Stage Five, is the most advanced stage in reading development. Readers in Stage Five analyze, synthesize, and evaluate what they read. A summary of Chall's (1967, 1983) Developmental Stages of Reading is shown in Table 2.1.

Table 2.1

Chall's (1967, 1983) Developmental Stages of Reading

Stage	Name	Age and Grade Span
Zero	Emergent Reading	birth to kindergarten
One	Initial Reading and Decoding	grades one through two
Two	Confirmation, Fluency, Ungluing from Print	grades two through three
Three	Reading for Learning the New	grades four through Jr high
Four	Multiple Viewpoints	high school
Five	Construction and Reconstruction	age 18 or older

Chall's (1967, 1983) work does not directly address automaticity, a critical issue of reading. Automaticity is defined as the ability to read without spending mental energy on decoding text (Rasinski, 2004; Samuels, 2002; Stanovich, 1991b). Automaticity enables readers to *think* about what they read while they are reading. In the schooling of elementary students, automaticity is the shift from "learning to read" to "reading to learn." As students develop automaticity, they progress through four early phases (Ehri, 1998, 2005; Kuhn & Stahl, 2003).

Ehri's (1998, 2005) work illustrates the four developmental phases of automaticity. In the first, "pre-alphabetic phase," students read words through memorization of sight words or use context cues to guess an unfamiliar written word. In the next phase, the "partial-alphabetic phase," students are able to read sight words because they know some letters of the alphabet and they can put them together to form words. In the third or "full-alphabetic phase," readers store sight words in memory and can decode novel words by making connections between the letters seen in written words and phonemes in the spoken equivalents. In the fourth or "consolidated-alphabetic phase," students have a bank of sight words in memory. Their print lexicons rapidly expand as they read many more different words. As more words are stored into memory, letter pattern knowledge becomes stronger.

Impact of Reading

As stated previously, learning to read fluently is one of the most significant accomplishments of an elementary school student (Johnson et al., n.d.). Students who achieve reading fluency experience abundant benefits. Students who fail to achieve reading fluency too often experience serious consequences to their academic growth and subsequent life successes.

Growth in reading achievement brings increased vocabulary and greater general knowledge (Cunningham & Stanovich, 1990, 1991, 1997, 1998, 2001, 2003; Stanovich & Cunningham, 1992, 1993). Time spent reading correlates to higher reading achievement (Cunningham & Stanovich, 2001; Taylor et al., 1990). The difference in the amount of time spent reading among students in grades third through fourth can predict with some accuracy, subsequent growth in reading comprehension (Taylor et al., 1990). The correlation between minutes reading at school and reading achievement was .37 (Taylor et al., 1990).

Cipielewski and Stanovich (1992) used regression analysis to determine whether print exposure predicted individual differences in reading ability growth as measured by the Iowa Test of Reading Comprehension (Hieronymus, Hoover, & Lindquist, 1982), the Stanford Reading Comprehension, Stanford Reading Rate, and Stanford Phonetic

Analysis (Karlsen, Madden, & Gardner, 1984). Print exposure was measured by title and author recognition. In five out of six analyses, the print exposure measure was able to account for the variance in fifth grade reading ability, after adjustments were made for third grade reading ability. It was concluded by the authors that print exposure plays a role in both developing reading ability and contributes to the growth of reading ability.

Cipielewski and Stanovich (1992) examined fifth grade reading comprehension scores as measured by their Stanford Reading Comprehension scores with title and author recognition. It was found that in the fifth grade students, title recognition has a unique variance of 11.0%, when third grade reading comprehension was partialed from the analysis. Author recognition had an 8.1% unique variance when third grade reading comprehension was partialed. The authors concluded that title and author recognition accurately predicted performance on Stanford Reading Comprehension scores for 89% of the students and author recognition accurately predicted approximately 92%.

Cipielewski and Stanovich (1992) also considered the relationship between title and author recognition and the Stanford reading rate subtest. Significant unique variances were found for both measures when third grade reading comprehension was partialed. Unique variances of 10.7% were found for title recognition and 18.5% for author recognition. Again, the two measures were found to accurately predict reading rate scores with approximately 80% accuracy for title recognition and 81.5% accuracy for author recognition.

Title recognition had a significant and unique variance of 7.4% as a predictor of Iowa reading comprehension scores, after third grade comprehension skill and decoding were partialed (Cipielewski and Stanovich, 1992). Author recognition was found not to predict Iowa reading comprehension scores. Student performance on the Stanford Phonetic Analysis was used to determine the decoding ability that was partialed in this analysis.

Cunningham and Stanovich's (1997) study showed that first grade students who develop reading skills at a faster rate than their peers are more likely to read more over time. Reading more words further improved their reading comprehension, vocabulary, and general knowledge (Cunningham & Stanovich, 1997). In a long-term follow up within the population, Cunningham and Stanovich (1997) found 11th grade students who experienced early reading success read more as high school students. Children who struggled with reading in first grade but subsequently caught up at grade level in the third and fifth grades also demonstrate more print exposure in 11th grade (Cunningham & Stanovich, 1997).

Early reading success produces a synergy that contributes to reading achievement (Cunningham & Stanovich, 1998, 2001, 2003). Specifically, readers with early reading success read more, develop larger vocabularies, and expand their general knowledge. This is known as the Matthew Effect (Stanovich, 1986). Early reading difficulties are the academic equivalent of the "poor-get-poorer" phenomenon. Young readers who do not learn to read are likely to fall further behind. The students who fall behind are likely to read less and fail to obtain many of the rewards of successful reading.

Successful Readers and Academic Success

The link between successful readers and academic success begins even before the primary school years (Bear et al., 1996; Chall, 1967, 1983; Elley, 1991; Feitelson et al., 1993; Pretorius, 2000; Committee on the Prevention of Reading Difficulties in Young

Children, 1998). Preschool aged children who are read to have larger vocabularies, greater general knowledge, and better development of ideas (Bus, van IJzendoorn, & Pellegrini, 1995; Cunningham & Stanovich, 1997; Neuman, Copple, & Brendekamp, 1998). At school entry, children who are read to in preschool learn to read and write faster than their peers who were not read storybooks (Bus et al., 1995). Students who are successful readers tend to have better vocabularies and cognitive skills (Cunningham & Stanovich, 1990, 1991, 1997, 1998, 2001, 2003; Stanovich & Cunningham, 1992, 1993). Reading ability and academic success correlate across all grades and university levels (Cunningham & Stanovich, 1997; Elley, 1991; Feitelson et al., 1993; Pretorius, 2000). At all grade levels, academically successful students read more frequently and comprehend more of what they read (Pretorius, 2000). Strong readers experience the "rich-get-richer" side of the Matthew Effect (Stanovich, 1986).

Poor Readers and Academic Struggles

Conversely, most poor readers continue to struggle throughout their school careers (Stanovich, 1986). Poor readers tend to read less print (Cunningham & Stanovich, 1998, 2001, 2003). Students who read less do not develop vocabulary and reading comprehension skills as quickly as their peers who read more print (Cunningham & Stanovich, 1990, 1991, 1998, 2001, 2003; Stanovich & Cunningham, 1992, 1993). The result for most struggling readers is poor reading comprehension (Cipielewski & Stanovich, 1992). First grade students who struggle with reading are less likely to improve their vocabulary and general knowledge. First graders who are poor readers are also less likely to become avid and lifelong readers. Most students who have difficulty reading experience poor academic achievement (Feitelson et al., 1993; Pretorius, 2000;

Committee on the Prevention of Reading Difficulties in Young Children, 1998). Especially significant, children who begin school with poor reading ability rarely “catch up” to their peers (Torgesen, 1998).

This paper focuses on a subgroup of lower achieving readers, specifically children identified with specific learning disabilities (SLD). There are many reasons why students may experience reading difficulties. These include (a) speaking a primary language that is not English, (b) low intelligence, (c) poor reading instruction and (d) SLD (Committee on the Prevention of Reading Difficulties in Young Children, 1998; Fletcher et al., 1994; Foorman et al., 1998; Fuchs et al., 2002). The variety of potential causes for reading difficulties can make it difficult to identify students with SLD. Researchers are concerned with the ability to identify students with SLD and separate SLD from other causes of low reading achievement. Students described as “low achievers” have expected reading difficulties because of low intelligence. Students with SLD are unique from other student groups with academic difficulties because students with SLD have “unexpected” learning difficulties. The learning challenges of students with SLD are unexpected because of their average to superior intelligence.

Specific Learning Disabilities

To understand the full nature of SLD, it is imperative to investigate the definition, as well as the diagnostic practices typically used to identify SLD. Diagnostic approaches attempt to capture the uniqueness of SLD, while separating SLD from other causes of reading difficultly. Currently, our American education system is experiencing a shift from the traditional approach of identifying SLD to a contemporary one. The strengths and weaknesses of each diagnostic approach are examined. Finally, the diagnostic procedures used to identify students with SLD in Idaho are presented.

Definition and Identification of Specific Learning Disabilities (SLD)

SLD diagnosis involves educational and psycho-educational assessments, observations, and review of the individual's medical, developmental, family, and educational history (American Psychiatric Association, 2012). SLD excludes students whose academic difficulties are primarily caused by (a) visual, hearing, or motor impairments, (b) cognitive impairment, (c) emotional disturbance, (d) environmental, (e) cultural or economic disadvantages, and (f) limited English speaking skills (American Psychiatric Association, 2012; Idaho State Department of Education, 2009; Individuals with Disabilities Improvement Education Act [IDEA], 2004; Stanovich, 1991a). Furthermore, it must be documented that the academic deficits are not the result of a lack of appropriate instruction (American Psychiatric Association, 2012; Idaho State Department of Education, 2009; IDEA, 2004; Stanovich, 1991a). Students with SLD have academic achievement that is well below their measured intellectual ability (American Psychiatric Association, 2012). The intellectual ability of students with SLD is typically within the average to superior range.

An individual with a SLD performs poorly in one or more of the following academic domains: (a) oral expression, (b) listening comprehension, (c) written expression, (d) basic reading skills, (e) reading fluency, (f) reading comprehension, (g) mathematic calculation, and (h) mathematics problem solving (Idaho State Department of Education, 2009). The impact of the SLD must significantly interfere with academic achievement, occupational performance, or daily living activities (American Psychiatric Association, 2012).

The identification of students with SLD is a relatively new and challenging practice in education, despite the fact that students with these academic characteristics have existed since the beginning of educational systems (Reschly, 1996; Siegel, 1999a; Vaughn & Fuchs, 2003). Unlike many other disabilities, SLD does not have a clear medical cause (Reschly, 1996). SLD is identified later in childhood. SLD is not discernible in infants, toddlers, and preschool children. The majority of students with SLD are referred for an evaluation only after demonstrating significant and long-term academic and behavioral problems. Students with SLD are difficult to separate from students whose academic deficits are due to ineffective instruction (Baskette et al., 2006; Vellutino et al., 1996).

Traditional Identification of Specific Learning Disabilities

Since 1977 until very recently, the identification of students with SLD relied on the IQ-achievement discrepancy model (Berkeley et al., 2009; Kavale, 2002; Mercer et al., 1996). To be eligible for special education under the IQ-achievement discrepancy model, students must display a significant discrepancy between intelligence quotient (IQ) scores and academic achievement, with the latter falling far below what would be predicted by the former (Fletcher et al., 2005; Fletcher et al., 1994; Mercer et al., 1996; Stanovich, 1991a). Students with IQs within the average range or higher are labeled as SLD when their achievement scores are significantly below the IQ score. Until August of 2010, the typical diagnostic process in Idaho required a significant discrepancy of 15 points between IQ and academic achievement as measured on standardized assessments

(Idaho State Department of Education, 2009). For example, a student might have a standard score of 100 points (within the average range) on a standardized IQ test. To be labeled as having a SLD, he or she would have to have a standard score of 85 or lower on an achievement test. This 15 point gap between the IQ score historically has been a primary diagnostic criterion for the label of SLD.

Criticisms of the Traditional Model to Identify Specific Learning Disabilities

Over the past two decades the efficacy of the IQ-discrepancy increasingly has come under question. There are several arguments against the use of the IQ-achievement discrepancy model. These include (a) expansion of the number of students labeled as SLD, (b) the variability of eligibility formulas from state to state for SLD, (c) the questionable ability of the discrepancy model to identify "true" SLD, and (d) the consequences of "waiting-to-fail." "Waiting-to-fail" means that a student must fall significantly behind in academics before he or she can receive special education (Bradley, Danielson & Doolittle, 2007; Gersten & Dimino, 2006; Lau et al., 2006; Lyon, 1996; Vaughn & Fuchs, 2003; Vaughn, Linan-Thompson, & Hickman, 2003).

Following the introduction of the discrepancy model in 1977, the nation quickly experienced more than a 200% increase in the number of students labeled with SLD (Baskette et al., 2006; Berkeley et al., 2009; Bradley, Danielson, & Doolittle, 2005; Scruggs & Mastropieri, 2002; Shaw, Cullen, McGuire, & Brincherhoff, 1995; U.S. Department of Education, 2000; Vaughn et al., 2003). Critics attribute this rapid growth of SLD to an over-identification of SLD that was directly attributed the IQ-achievement discrepancy model (Gersten & Dimino, 2006; Vaughn et. al., 2003). The increasing number of students labeled SLD costs school districts money and resources (Gersten &

Dimino, 2006; Wong, Graham, Hoskyn, & Berman, 2011). If there is an over-identification of students as SLD, finding a way to more accurately identify students with SLD while minimizing over-identification issues would be of benefit to students as well as school systems (Valas, 1999; Gersten & Dimino, 2006).

Fundamental conceptual and procedural problems may have contributed to the growing numbers of students diagnosed with SLD. The lack of consensus on what constitutes SLD and vague definitions may have contributed to the challenge of accurate identification (Lyon, 1996; Ysseldyke et al., 1983). The numbers of SLD students may have grown because of inconstant state criteria and failure to apply strict adherence to criteria at the local level (Berkeley et al., 2009; Johnson, Mellard, & Byrd, 2005; Mercer, King-Sears, & Mercer, 1990; Ofiesh, 2006; Reschly, 1996; Scruggs & Mastropieri, 2002). The SLD label may be incorrectly applied to students who more accurately fit the criteria of an intellectual disability (MacMillan, Gresham, Siperstein, & Bocian, 1996). The failure of school practitioners to accurately apply federal and state definitions may be a factor in the rising numbers of students labeled as SLD (Lyon, 1996; Mercer et al., 1996; Scruggs & Mastropieri, 2002). In a study review from 1978 to 1999, it was estimated that 52% to 70% of students identified by the schools as SLD did not meet the eligibility criteria for SLD by federal and state definitions (MacMillan & Speece, 1999).

Critics of the IQ-achievement discrepancy model are dissatisfied with the dissimilar SLD eligibility criteria between states (Mercer et al., 1990; Wong et al., 2011). The lack of consensus as to the size of the gap between achievement and IQ is a problem (Aaron, 1991). For example, in 1990, two states required a discrepancy between IQ and achievement of two standard deviations, six states required a standard deviation of 1.5, and six states required one standard deviation (Mercer et al., 1990). Three states permitted a range of one to 1.5 standard deviations. Five states varied the discrepancy based upon the student's age, grade, or both age and grade. The problem with the inconsistency of formulas in the discrepancy model is that a student could be considered as SLD in one state but not in another (Mercer et al., 1990; Moats & Lyon, 1993).

Another argument against the IQ-achievement discrepancy model is the variability in the number of students identified as SLD among states. The prevalence of SLD between the states ranged from 2% to nearly 9% (Finlan, 1992). In 1992 to 1993, Wisconsin reported 2% of their student population as SLD while Massachusetts identified 7% (Coutinho, 1995). The variance has been attributed to the differences between state IQ-achievement discrepancy formulas.

Opponents of the IQ-achievement discrepancy model argue that the approach may fail to accomplish the one thing it was created to achieve, to accurately identify students with SLD. The use of IQ scores may be immaterial to the diagnosis of SLD (Siegel, 1999b). Many have argued that the discrepancy model is unable to differentiate between reading problems caused by reading-related cognitive deficits found in students with SLD or instructional failure (e.g., Vellutino et al., 2006). Siegel (1989) found that students with and without IQ-achievement discrepancy had similar performances on reading achievement and cognitive assessments associated with reading ability (phonological awareness, short-term verbal memory, word retrieval, orthographic coding, and visual measures). Siegel's (1989) findings were later supported by Fletcher et al. (1994). The discrepancy model may over identify students with high IQs and average achievement while missing students with lower IQs and below-average academic achievement (Birch, 2004; Semrud-Clikeman, 2005; Semrud-Clikeman et al., 1992). The IQ-achievement discrepancy approach is likely to inevitably result in misclassification of students with reading difficulties (Vaughn et al., 2003). Students who were not SLD are mislabeled SLD, students with other disabilities are misidentified as SLD, and students who are genuinely SLD can go unnoticed.

The standardized assessments used to identify SLD in the discrepancy model may not effectively separate students who are under-achieving from students with SLD (Fuchs & Young, 2006; Lau et al., 2006; Vaughn & Fuchs, 2003; Vaughn et al., 2003; Vellutino, Scanlon, & Lyon, 2000; Ysseldyke, 2005; Ysseldyke et al., 1983). Defects in standardized assessments may underestimate the intelligence of students with reading disabilities (Fuchs & Young, 2006; Siegel, 1999a; Vellutino et al., 2000; Ysseldyke, 2005). Standardized assessments cannot make a clear distinction between students who have mild SLD and those who do not have SLD but have difficultly reading for other reasons (Reschly, 1996). The result is that students with mild SLD may not be identified under the discrepancy model. Inherent linguistic and cultural biases in standardized assessments may over-identify poor, minorities, and culturally and linguistically diverse students as SLD (Artiles & Trent, 1994; Garland & Strosnider, 2005; Lau et al., 2006; Reschly, 1996; Semrud-Clikeman, 2005; Ysseldyke, 2005). Finally, national normative data of standardized assessments may not reflect the population of the students within a school district (Siegel, 1999a; Vaughn & Fuchs, 2003; Vellutino et al., 2000; Ysseldyke, 2005).

Further problems have been identified with the discrepancy model linked to instruction. The IQ-achievement discrepancy does not consider a student’s instructional history, most notably the quality of past reading instruction (Velluntino et al., 2006). Standardized academic achievement tests used in the discrepancy model may not match the curriculum taught (Fuchs & Young, 2006; Lau et al., 2006; Vaughn & Fuchs, 2003; Vellutino et al., 2000; Ysseldyke, 2005).

Significantly, the IQ-achievement discrepancy calculation does not inform remediation (Aaron, 1991; Semrud-Clikeman, 2005; Siegel, 1999b; Vaughn & Fuchs, 2003). When given the standard scores from IQ and achievement assessments, educators may underestimate the capabilities of students, causing teachers to reduce expectations for reading achievement (Ysseldyke, 2005). The result perpetuates the Matthew Effect by setting low expectations for students that are capable of achieving reading proficiency (Cunningham & Stanovich, 1998, 2001, 2003; Morgan, Fuchs, Compton, Cordray, Fuchs, 2008; Ysseldyke, 2005).

Finally, the discrepancy model or "waiting-to-fail" requires students to fall too far behind in academics before they can receive interventions (Bradley et al., 2007; Gersten, & Dimino, 2006; Lau et al., 2006; Lyon, 1996; Vaughn & Fuchs, 2003; Vaughn et al., 2003). In general, remediation is given to students only after they meet eligibility criteria as SLD under the IQ-achievement discrepancy model. One reason for the delayed identification emerges from the psychometric properties of standardized assessments (Vaughn et al., 2003). Usually first grade students do not meet eligibility criteria because there is not a wide enough gap between IQ and achievement at this young age (Stage, Abbott, Jenkins, & Berninger, 2003). It commonly takes one to three years of continued reading failure before the gap between intellectual ability and demonstrated academic achievement grows wide enough to label students as SLD under the discrepancy model

(Lyon, 1996; Stage et al., 2003; Stuebing, Fletcher, & LeDeux, 2002). Seventy-four percent of the students with SLD were identified at ages 8 or 9 (Lyon, 1996). The reading problems of the students typically were known early in their academic career, but they did not meet the eligibility criteria until they were in the third or fourth grade. By then, the students experienced common problems associated with reading difficultly. The common problems associated with poor reading achievement are low self-esteem, decreased motivation, and never fully acquiring basic reading skills. The cost of "waiting-to fail" is the time that could have been spent on remediation (Lyon, 1996; Vaughn et al., 2003).

Early intervention may prevent long-term reading difficulties (Vellutino et al., 2006). Struggling students who received reading interventions only during their kindergarten year or interventions in kindergarten and first grade performed better on emergent literacy skill assessments than those who did not receive any interventions in reading. The majority of the students who received early intervention were no longer at risk. The discrepancy model may make it difficult to identify younger students with SLD and prevent the benefits of early interventions (Fletcher et al., 1994; Gersten & Dimino, 2006; Semrud-Clikeman, 2005; Ysseldyke, 2005).

To be fair, not everyone agrees that there is an over representation of students with SLD. Others have proposed alternative explanations for the increase in numbers (Lyon, 1996). Several reasons besides an over-identification may account for the growth of SLD. First, it may be that increased awareness of SLD has led to more assessments and identification. More students with milder forms of SLD have been added to the numbers and profited from specialized instruction. It is important to provide special education to students with mild SLD because they may experience academic challenges and face limited career opportunities as a consequence (Reschly, 1996). An awareness of the importance of phonological awareness in reading development has also contributed to increased identification rates (Lyon, 1996; Wong et al., 2011). Phonological awareness assessments may help identify younger students in late kindergarten and early first grade with SLD.

Contemporary Identification of Specific Learning Disabilities

Response to Intervention (RTI) is a general education approach that has emerged in the last decade to systematically address students who struggle with academic and behavioral challenges. RTI uses problem-solving and student-centered models to provide interventions for students who have trouble learning to read (Berkeley et al., 2009; Johnson, Mellard, Fuchs, & McKnight, 2006). RTI requires targeted and researched-based interventions. RTI may add more instructionally relevant data than is collected in the IQ-achievement discrepancy model because progress monitoring probes and interventions are similar to the core curriculum (VanDerHeyden, Witt, & Gilbertson, 2007; Lau et al., 2006). RTI puts more emphasis on student outcomes instead of IQ and achievement tests (Kavale, 2005). The reauthorization of the Individuals with Disabilities Education Improvement Act of 2004 allows for RTI to be an alternative means of identifying students with SLD (Berkeley et al., 2009). Under RTI, students with SLD conceptually are being redefined as those students with reading difficulties who fail to progress despite the provision of empirically documented and research-based effective reading instruction.

Within the RTI model are two approaches to intervention implementation. They are (a) the problem-solving approach, and (b) the protocol system approach (Buffum, Mattos, & Weber, 2009; Marston, 2005). The problem-solving approach has four stages (problem identification, problem analysis, plan implementation, and plan evaluation). To illustrate how the problem-solving approach works, consider a first grade student who has difficulty with letter sound fluency. A problem-solving team moves systematically through each of the aforementioned four stages to design an individualized intervention.

The second RTI approach is the standard treatment protocol system (Buffum et al., 2009; Marston, 2005; Wong et al., 2011). The standard treatment protocol has been more researched than the problem-solving approach (Wong et al., 2011). In the standard treatment protocol system, the student firsts meets established criteria for intervention (Buffum et al., 2009). The intervention is matched to address the student's skill deficit. To understand how this approach works, consider the first grade student who has difficulty with letter sound fluency. The first grade student's performance on letter sound fluency universal screen fell below the 25th percentile. This student would quickly be placed in a supplemental reading program along with other students with the same profile.

Both the problem-solving approach and standard protocol approach have tiers or levels of intervention. The primary components of both RTI systems are (a) tiers of varying instructional intensity targeted to student needs and (b) frequent progress monitoring to determine the effectiveness of the instruction and to make adjustments as needed (Buffum, Mattos, & Weber, 2009). The tiers vary in RTI models currently used in the United States (Fuchs & Fuchs, 2007; National Association of State Directors of

Special Education & Council of Administrators of Special Education, NASDSE & CASE, 2006). The tiers differ in frequency and intensity of interventions (Buffum et al., 2009). The most common RTI model has three tiers, although some models have four or more tiers (Buffum et al., 2009; Mellard & Johnson, 2008; Shapiro, n.d.). For the purposes of this paper, the three-tier model will be described to illustrate the RTI process of SLD identification.

Tier 1 is also known as the core, base, primary, or universal program (Buffum et al., 2009; Marston, 2005). Tier 1 is the initial level instructional practices. The teaching and school experiences in Tier 1 are what all students encounter every day. At the Tier 1 level, general education teachers implement scientifically based instruction and core curriculum while closely monitoring their students' academic progress, especially the students who are found to be at risk on universal screening assessments (Fuchs & Deshler, 2007; Hughes & Dexter, n.d. a; Mellard & Johnson, 2008; NASDSE & CASE, 2006). Approximately 75% to 80% of all students are should be able to demonstrate adequate educational progress when receiving instruction at the Tier 1 (Buffum et al., 2009; NASDSE & CASE, 2006). Universal screenings of academic skills are given to all students several times per school year to identify students who are potentially at risk for academic failure (Mellard & Johnson, 2008).

Students who are struggling academically at the Tier I level then move on to receive research-based interventions in the general education setting in Tier 2. Tier 2 is known as "supplemental level" or "secondary intervention" because the interventions are in addition to Tier 1 instruction (Buffum et al., 2009; Marston, 2005; Swanson & Vaughn, 2011). Tier 2 usually would target approximately 15% - 20% of the students in a school (NASDSE & CASE, 2006; Swanson & Vaughn, 2011). Tier 2 interventions generally include small group instruction, along with instruction in the general education classroom (Fuchs, Compton, Fuchs, Bryant, & Davis, 2007; Mellard & Johnson, 2008; Swanson & Vaughn, 2011). The small groups of three to five students are formed based on the homogeneous needs of the students (Mellard & Johnson, 2008; Swanson & Vaughn, 2011). Interventions in Tier 2 usually include increased intensity, frequency, and duration of interventions. The short-term interventions last from nine to 12 weeks and are often provided three to five days per week for 30 to 60 minutes per session. Tier 2 interventions are typically implemented by the general education classroom teacher, specialists, or supervised paraprofessionals (Swanson & Vaughn, 2011). In Idaho, Tier 2 interventions are legally required to be different from and in addition to the core instruction (Idaho State Department of Education, 2009).

Tier 3 is considered the “intensive level” of intervention (Buffum et al., 2009; Marston, 2005). Tier 3 targets approximately the 5% - 10% of the remaining student population who present the greatest educational challenges (Buffum et al., 2009; NASDSE & CASE, 2006). Tier 3 may offer (a) different interventions than Tier 2, (b) offer the same interventions but increases intervention time, and/or (c) increased frequency of progress monitoring (Buffum et al., 2009; Fuchs et al., 2007). Tier 3 interventions are more intense and include small group and individual instruction (Marston, 2005; Mellard & Johnson, 2008; Swanson & Vaughn, 2011). Tier 3 interventions are typically implemented by reading specialists, supervised paraprofessionals, and/or special education teachers (Swanson & Vaughn, 2011). Some of these students may be enrolled in special education. Tier 3 typically requires the collaboration with a special educator. If the interventions do not work, the student may be referred to special education. In some models, Tier 3 is seen as synonymous with traditional special education placement (Fuchs et al., 2007; Mellard & Johnson, 2008). In the models that equate Tier 3 with special education, students enrolled in special education receive specially designed instruction from a special education teacher or supervised special education paraprofessional.

A visual representation of the RTI model is presented in Figure 2.1. The RTI model depicted below illustrates the structure for academic and behavioral systems. The bulk of the student population is in Tier 1 are at the bottom of the pyramid. A small percentage of students are in Tier 2. At the top of the pyramid, the fewest number of students are in Tier 3.

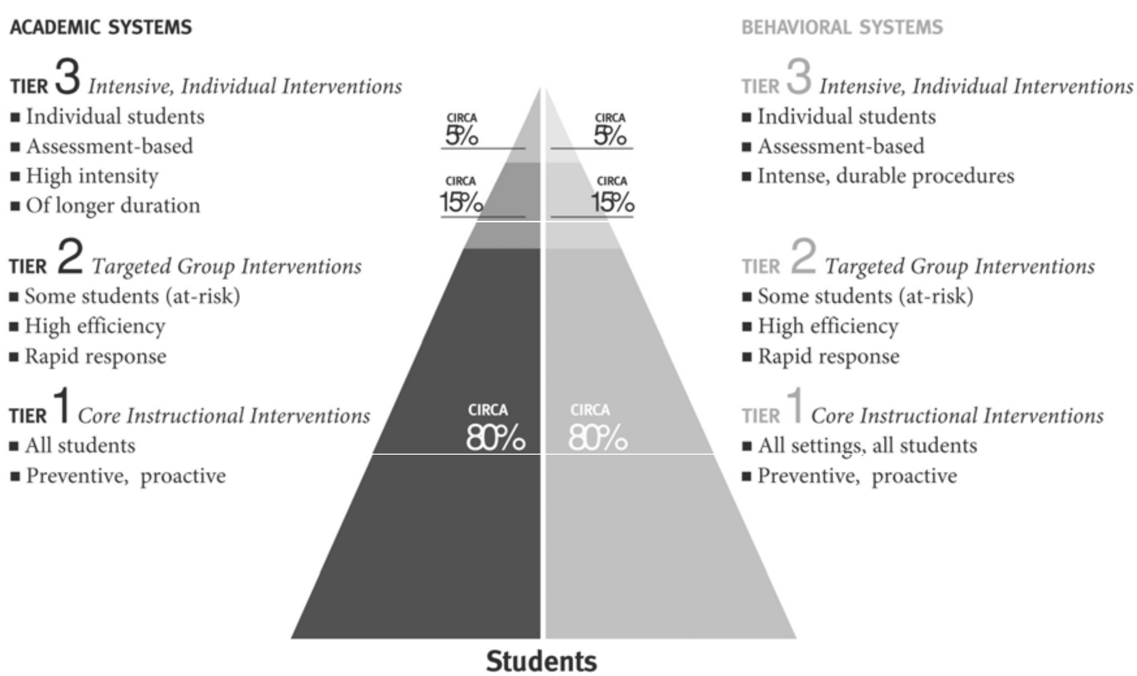

Figure 2.1. The RTI Three-Tier Model (NASDSE & CASE, 2006).

Criticism of the Contemporary Model to Identify Specific Learning Disabilities

There are several arguments against RTI, including (a) lack of research documenting its educational efficacy, (b) role changes of educational staff, (c) funding, and (d) the time it takes to identify students with SLD. Some critics have claimed that RTI ignores the well-established body of research that has emerged since the 1970s to describe the SLD construct (Kavale, 2005). The RTI model may be unable to sort students with SLD from students with other types of disabilities (Kavale, 2005). When students are labeled SLD through the RTI approach without assessments of cognitive processes, students who do not truly have SLD may be inappropriately included in the category (Baskette et al., 2006; Goodman & Webb, 2006; Kavale, 2005; Mastropieri & Scruggs, 2005).

The second issue is concerned with the potential to over identify students as having SLD. Under the IQ-discrepancy model, the majority of the students were identified as SLD in grades three or four (Baskette et al., 2006). If RTI is used to identify students as SLD, too many kindergarten through fourth grade students could be labeled as SLD, thus significantly increasing the numbers of students in special education. The potential to over identify students with SLD may be difficult to avoid because RTI targets interventions for students who score at the 25th percentile rank or lower on reading assessments (Gerber, 2005; Kavale, 2005).

To counter the concern by critics of RTI over the potential for the rising number of students labeled as SLD, RTI supporters have noted exactly the opposite outcome has emerged. Over the last decade (2000-2009), while the total number of students requiring special education grew 2%, the number of students identified as SLD declined by 14%

(Cortiella, 2011). There were several reasons for the decline in the percentages of students identified as SLD, including number of struggling students receiving preventative interventions because of RTI.

The third argument against RTI is in the inconsistent implementation of RTI across the states. RTI models, procedures, and practices differ from state to state, making RTI as inconsistent as the IQ-discrepancy model (Berkeley et al., 2009). States vary in the RTI approach that they choose (Hallahan et al., 2007). For example, some states require the exclusive use of phonologically based reading interventions, while other states allow any scientifically based curriculum for RTI interventions (Baskette et al., 2006). The differences in RTI practices between states may result in variability of SLD prevalence among states. There is a wide range of RTI interventions and this might cause a higher variability in prevalence rates of SLD students in states that use RTI to establish eligibility for special education (Baskette et al., 2006; Hallahan et al., 2007).

One of the strongest arguments against RTI is the limited amount of research on the educational impact of the implementation of RTI (Garland & Strosnider, 2005; Gerber, 2005; Kavale, 2005). The lack of large-scale studies of reading interventions is another weakness of RTI (Gersten & Dimino, 2006). Although the RTI model calls for research-based interventions, many states implementing RTI include interventions that have a “modest empirical validation” (Kavale, 2005). Furthermore, some of the RTI interventions are designed to address one deficit despite the fact that many struggling readers present a “multivariate problem” of reading deficits. Most of the available RTI research is limited to reading interventions for grades kindergarten through third grade (Baskette et al., 2006). More research studying systematic phonics programs, especially those targeting culturally and linguistically diverse students is needed (CLDS; Ehri et al., 2001). In addition, more research is needed to find research-based interventions that can be “realistically and reliably” implemented by general education teachers (Speece, Case, & Molloy, 2003).

Another criticism of RTI is its potential to change the roles of teachers, special educators, and school psychologists (Gerber, 2005; Gersten & Dimino, 2006; Mastropieri & Scruggs, 2005). In RTI, teachers may have the primary responsibility to provide instruction at all three tier levels, while continuing to monitor progress of all students, and move students between the tiers (Gerber, 2005; Gersten & Dimino, 2006; Mastropieri & Scruggs, 2005). The roles of special educators and psychologists are less clear until the point of referral to special education (Gerber, 2005; Gersten & Dimino, 2006; Mastropieri & Scruggs, 2005). General education classroom teachers will be responsible for the delivery of scientifically based instruction in Tier 1, but may not have the knowledge or tools to accomplish this task (Mastropieri & Scruggs, 2005).

Concern about funding is yet another criticism of RTI. Some oppose RTI because of the provision under the IDEA 2004 that allows local education agencies to use up to 15% of their IDEA Part B funds for early intervention. These critics worry that using the IDEA funding in the general education setting will take money from limited special education funds (Kavale & Spaulding, 2008; Mastropieri & Scruggs, 2005; Scruggs & Mastropieri, 2002).

Finally, critics worry that RTI slows down the process of SLD identification. The time it takes to identify students with SLD and placement into special education under the RTI model is questioned (Kavale, 2005). Although there is not a universally set time, Tier 2 interventions generally last from eight to 12 weeks (Idaho State Department of Education, 2009; Mellard & Johnson, 2008). This could mean that a student with SLD would have to show a failure to respond to interventions for a minimum of two to three months before being referred for a special education evaluation. However, the argument about length of time to identify students with SLD under RTI can also be made against the IQ-discrepancy approach. As discussed earlier, critics call the IQ-achievement discrepancy approach "waiting-to-fail" because the processes to deliver special education to students with SLD takes several academic years (Baskette et al., 2006; Montgomery & Moore-Brown, 2006).

In contrast, students can obtain interventions at an early age with RTI. RTI is considered a "preventative model" because interventions may be provided as soon as a reading problem is identified. Furthermore, RTI interventions may prevent placement in special education because early intervention may help the student overcome the learning difficulty, thus making further special education considerations unnecessary (Bradley et al., 2007; Fletcher et al., 2004; Gersten & Dimino, 2006; Lau et al., 2006; Semrud-Clickeman, 2005; Shaywitz et al., 2004; Vaughn & Fuchs, 2003).

Other criticisms of RTI have not gone unchallenged. Fuchs and Deshler (2007) disagree with the contention that there is little research on effective RTI interventions. A "large handful" of instructional procedures and curricula have been rigorously tested and should be considered "good bets" to use in Tier 1. RTI interventions are available for phonological processing, decoding, and reading fluency, but there are very few researched-based interventions for reading comprehension (Kavale, 2005). School districts and states would be well served with an efficient "vetting" process to identify scientifically based instructional practices and curriculum (Foorman, 2007; Fuchs & Deshler, 2007). Documents about effective reading instruction techniques and tools are available through the National Institute of Child Health and Human Development (NICHD) and the National Research Council (NRC; Foorman, 2007).

A Blended Approach to the Identification of Students with Specific Learning Disabilities

At the present time, RTI as a means to identify students with SLD is not mandated but permitted in the reauthorization of the Individuals with Disabilities Education Improvement Act (IDEA) of 2004 (Berkeley et al., 2009; Bradley et al., 2007). The act continues to allow states to use the traditional IQ-achievement discrepancy criterion to identify students as learning disabled (Kavale & Spaulding, 2008). States may also employ a combination of the approaches as criteria for the category of SLD. Merging parts of RTI and comprehensive psychometric assessments as a means to identify students with SLD may satisfy both camps in the debate (Baskette et al., 2006: Bradley et al., 2007; Barth et al., 2008; Batsche et al., 2006; Kavale & Spaulding, 2008; Fletcher et al., 2004; Fuchs et al., 2010; Keogh, 2005; McKenzie, 2009).

Some advocate for a blended approach of RTI and a comprehensive evaluation because of the criticisms of both the (a) traditional discrepancy model and (b) RTI (Fletcher et al., 1994; Fletcher et al., 2005; Fuchs, Deshler, & Reshley, 2004; Garland & Strosnider, 2005; Mather & Gregg, 2006; Ofiesh, 2006). The data gathered in RTI and information obtained from a comprehensive evaluation has the potential to aid in the accurate identification of students with SLD (Fletcher et al., 2005; Fuchs, Deshler, & Reshley, 2004; Ofiesh, 2006).

Including an assessment of cognitive processes with RTI may be an effective means of identifying students suspected of having SLD (Johnson et al., 2010). The meta-analysis of 32 studies found moderately large to large effect sizes (ES>.80) in cognitive processing differences between students with SLD and typically achieving students. The cognitive processes related to reading are working memory, processing speed, executive function, and receptive and expressive language. Students with SLD performed significantly lower on all cognitive processing areas than their typically achieving peers.

A blended approach is consistent with the historical definition of SLD because the central characteristic of unexpected academic underachievement is maintained (Fletcher et al., 2005; Fletcher, Morris, & Lyon, 2003). Additionally, the discrepancy component in the blended approach is preserved; however, the discrepancy is not the gap between IQ and achievement. The discrepancy in the RTI model occurs when the student demonstrates a significant lag in academic performance and fails to respond adequately to the intervention (Fletcher et al., 2005). Students with (a) average IQs, (b) achievement below the 25th percentile, and (c) lack of response to Tier 2 levels of RTI instruction were more likely to have SLD (Johnson et al., 2010). When a student fails to make the expected gains during reading intervention, in many cases the student may be referred for a SLD special education evaluation (Fletcher et al., 2003, 2005).

Since August 1 of 2010, the state of Idaho adopted a blended approach to determine eligibility for SLD (Henderson & Hopper, 2010; Hyatt, 2010; Idaho State Department of Education, 2009). To meet eligibility for SLD in Idaho, a student must fail to make significant gains during RTI, show low achievement on achievement assessments, and demonstrate a pattern of psychological processing skill strengths and weaknesses (Idaho State Department of Education, 2009). The psychological processes must be related to the area of academic concern (Hyatt, 2010). Phonological processing, processing speed (rapid automatized naming), and memory are some examples of psychological processes used in reading.

There are many reasons for Idaho's change to the blended model (Henderson & Hopper, 2010). First, the new eligibility criteria for SLD more accurately fit the federal definition. The new approach to identify SLD is consistent with the evidence that demonstrates that students may have impairments in specific areas. The blended model allows for the evaluation and intervention process to be more closely linked to the specific needs of the student suspected of having SLD. The decision to change the eligibility criteria reflects research opposed to the IQ discrepancy model. Finally, it is hoped that the change in eligibility criteria will lead to more accurate labeling of students as SLD.

Academic Characteristics of Students Diagnosed with Specific Learning Disabilities

The academic characteristics of students with SLD are similar to those described for all students with poor reading skills. Students with SLD display an assortment of significant skill deficits (Stanovich, 1988). Students with SLD may show more difficultly with decoding text, word recognition, comprehension, and reading fluency. Generally, students with SLD read with less accuracy, less automaticity, and less expression (Rasinski, 2004). Reading with accuracy refers to the correct decoding of words in text. Reading with automaticity is the ability to read without exercising effort into decoding. Freedom from decoding enables the reader to put more energy into comprehension, making connections, making inferences, and evaluating the content of the text (Rasinski, 2004; Samuels, 2002; Stanovich, 1991b). Reading prosody is the ability to read with correct expression and phrasing (Rasinski, 2004). Students with SLD may not read as much as their non-SLD peers (Cunningham & Stanovich, 1998, 2001, 2003). Students with SLD have an increased risk of dropping out of school (Blackorby & Wagner, 1996).

The Need for Early Identification of Reading Difficulties

The ability to identify students with potential reading difficulties early in the school career is critical for students, educators, and schools. A number of studies have documented the impact of early intervention. First grade students with difficulty in reading have been found to improve on standard reading measures after receiving interventions (Berninger et al., 1999). First grade students who were identified as at-risk for reading difficulties and given interventions outperformed their matched peers who did not receive interventions on reading and phonological-processing assessments (Hurford & Johnson, 1994). First grade students with reading difficulties who were tutored by non-certified tutors for 30 minutes, four days a week over one school year significantly outperformed non-tutored peers on reading tasks (Vadasy, Jenkins, & Pool, 2000).

Several studies demonstrate the ability to accurately identify students with potential reading problems at an early age. Catts, Fey, Zhang, and Tomblin (1999) found that over 70% of poor readers in second grade have a history of language impairments in kindergarten. Within this group of second grade students, most demonstrated poor phonological processing skills in kindergarten. A student's ability to read in first grade turns out to be an accurate predictor of how well the student will academically perform in high school (Cunningham & Stanovich, 1997; Davis, Lindo, & Compton, 2007; Fletcher et al., 2002).

The risk of "waiting-to-fail" before intervention may cost the student with reading difficulties the opportunity to compensate (Semrud-Clickeman, 2005; Shaywitz et al., 2004). Early intervention is important to combat low motivation and the fear of failure in students with reading difficulties (Hudson, High, & Otaiba, 2007; Semrud-Clickeman, 2005; Shaywitz et al., 2004).

Reading Achievement and Accountability

Under current educational policy, educators are held accountable for student achievement (Castillo, 2005; Schilling, Carlisle, Scott, & Zeng, 2007). The NCLB reauthorization and made substantial changes to the Elementary and Secondary Education Act (EASA; 1965; Jorgensen & Hoffmann, 2003). NCLB requires all students in grades three through eight and grade ten to have proficiency rates of 100% in reading, language arts, and math, and meet graduation requirements by 2014 (Davidson, Reback, Rockoff, & Schwartz, 2013; Linn, Baker, & Betebenner, 2002).

Under the Obama administration, there has been an easing of the NCLB achievement requirements. Some states have been granted waivers to allow more flexibility in NCLB requirements (U.S. Department of Education, 2012a). Instead of measuring student achievement by proficiency targets, the waivers permit states to measure student achievement based on student growth (Idaho Department of Education, 2015). Waivers may be granted to states that develop a plan to raise standards, enhance accountability, and improve teacher effectiveness in instruction (U.S. Department of Education, 2012a). Twelve states received waivers (US. Department of Education, 2012b). Twenty-six additional states, including Idaho and the District of Columbia applied for waivers (U.S. Department of Education, 2012b). Idaho received a waiver in October of 2012 (U.S. Department of Education, 2012a).

NCLB requires that states create an assessment system that tracks the academic achievement of students (Jorgensen & Hoffmann, 2003; U.S. Department of Education, 2003). States must test all students in grades three through eight and grade ten each year in reading and mathematics (Jorgensen & Hoffmann, 2003; U.S. Department of Education, 2003). The tests are based on the state standards, with the results made public (Jorgensen & Hoffmann, 2003; U.S. Department of Education, 2003). In response, the state of Idaho developed a comprehensive assessment system that goes beyond the requirements of NCLB (Barr, Joyner, Parrett, Willison, 2004). Idaho's assessment system begins with kindergarten and continues through high school. The assessment system includes reading, math, and language arts. The Idaho Reading Indicator (IRI) is the assessment tool used as a universal reading screen to identify students at-risk for reading failure for grades kindergarten through third. The additional purpose of the current IRI is to measure individual student growth (Idaho State Department of Education, 2013).

The Idaho Reading Indicator (IRI)

The IRI is a screening assessment that gives a "snapshot" of student's reading ability (Idaho State Department of Education, 2013). The IRI was not designed as, and does not function as, a comprehensive diagnostic reading assessment. It takes less than 10 minutes to administer the IRI (Idaho State Department of Education, 2013). The IRI is easy to score. The scoring system of the IRI consists of three-point rating, "1," "2," or "3" (Idaho State Department of Education, 2013). A student score of "3" indicates "mastery of the skills." A score of three is considered "proficient." A student score of "2" indicates "partial mastery of some or all skills." A student score of "1" indicates "a lack of mastery of some or all skills."

History of the IRI

The development of the IRI began years before the initial legislation to assess the reading abilities of all primary elementary students was enacted in 1999. In 1997, the Reading Study Committee was formed after resolutions were passed by the Idaho legislature directing the Idaho State Department of Education to conduct a study of reading education in Idaho (Barr, Flachbart, & Stewart, 2002). The Reading Study Committee consisted of the State Superintendent of Public Education, 25 professional committee members, and consultants. The Reading Study Committee commissioned a preliminary reading test to determine the status of reading education in Idaho. The study found that 40% of Idaho fourth grade students were reading below grade level. Initiatives were recommended to guarantee the right of all Idaho school children to read. An Interim Legislative Reading Committee was appointed in 1998 and continued the work of the Reading Study Committee by drafting the legislation for the Comprehensive Literacy Act. In 1999, the Idaho legislature passed the Idaho Comprehensive Literacy Act (Idaho Code 33-1614, 33-1615). Since 1999, all public school students, kindergarten through third grade, have been given the IRI to identify students below grade level in reading (Idaho Comprehensive Literacy Act, Idaho Code 33-1614).

Students identified as below grade level with scores of 1 on the IRI assessment are offered 40 hours of additional instruction time beyond the regular school day (Idaho Comprehensive Literacy Act, Idaho Code 33-1615). Students who score a 2 may be given intervention, providing the parents of a student who scored a 1 the choice not to allow their child to participate in intervention (S. Lee, personal communication, February 4, 2014). Intervention funding is based upon the number of students who receive a score of 1 on the IRI. In 2001, the Idaho Reading Initiative added statutory regulations to require that a minimum of 85 percent of all third grade students read at grade level at the end of third grade (Idaho Comprehensive Literacy Act, Idaho Code 33-1616). The amendment additionally makes test scores public (Idaho Comprehensive Literacy Act, Idaho Code 33-1616).

If schools fail to meet targets after two years, the state may initiate a school intervention program (Idaho Comprehensive Literacy Act, Idaho Code 33-1616). This may include visits from designated personnel from schools that have met state targets or personnel familiar with reading achievement. The intervention team will make recommendations to the district on how to meet and exceed reading achievement targets. The IRI data is used to determine funding for professional development, technical assistance, curriculum, and other unspecified activities by the Idaho State Department of Education (2013). IRI information is also used to fund school interventions.

From 1999 to present, the IRI has been changed three times. The IRI was developed with the assistance of two educational consultative companies, the Waterford Institute and AIMSweb (a Pearson product). For the purposes of this dissertation, each IRI will be referred to by the educational companies' names. The first version of the IRI will be referred to as the Waterford Institute IRI. The second and third IRI editions will be called the AIMSweb IRI.

Development of the Waterford Institute IRI

The Waterford Institute IRI was in use from 1999 through 2007 (Idaho State Department of Education, 2007). The Waterford Institute IRI was developed by a group including the State Superintendent of Public Education, Idaho State Department of Education (SDE) employees, reading teachers from around the state, and Waterford Institute consultants (Barr et al., 2002). National reading expert Marilyn Jager Adams consulted with the Waterford Institute to develop the IRI (Sherman, 2001). The IRI was pilot-tested during the 1999 – 2000 school year (Barr & Flachbart, 2003). The Waterford Institute IRI was administered three times per year, in the fall, winter, and spring. The Waterford Institute IRI was available in English and Spanish (Sherman, 2001). A three point scoring system was used to classify student performance on the Waterford Institute IRI (Idaho State Department of Education, 2004). A score of “3” was “at grade level” or proficient, “2” was “near grade level,” and “1” was “below grade level.”

The Waterford IRI contained multiple subtests per grade level (Idaho State Department of Education, 2004). The fall kindergarten IRI asked students to (a) write their name, (b) detect rhymes, (c) detect syllables, and (d) identify upper case letters. In the winter, the IRI contained subtests that asked students to (a) identify lowercase letters, (b) match first sounds, and (c) generate rhyme. In the spring, kindergarten students had to (a) say letters, (b) produce rhyme, (c) say the first sound in words, (d) identify a letter, a word, and a sentence, and (e) read words on a word list. A list of the subtests found on the Waterford Institute kindergarten IRI can be found on Table 2.2.

Table 2.2

Waterford IRI Kindergarten Subtests

Fall Subtests	Winter Subtests	Spring Subtests
Write Your Name	Identify Lowercase Letters	Say Letters
Detect Rhyme	Match First Sound	Produce Rhyme
Detect Syllable	Generate Rhyme	Say the First Sound
Identify Uppercase Letters		Identify a Letter, a Word, and a Sentence
		Read Word List

The first grade fall Waterford Institute IRI contained subtests similar to the kindergarten edition (Idaho State Department of Education, 2004). Students were required to (a) identify words, (b) produce rhyme, (c) write letters, (d) read a sentence, and (e) say the first word. In the winter, students were asked to (a) blend sounds, (b) read a story, and (c) sound out nonsense words. In the spring, students (a) read a story, (b) answered comprehension questions about the story, and (c) sounded out nonsense words. An outline of the subtests found on the first grade Waterford Institute IRI is on Table 2.3.

Table 2.3

Waterford IRI First Grade Subtests

Fall Subtests	Winter Subtests	Spring Subtests
Identify Words	Blend Sounds	Read a Story
Produce Rhyme	Read a Story	Answer Comprehension Questions
Write Letters	Sound Out Nonsense Words	Sound Out Nonsense Words
Read a Sentence		
Say the First Word		

In the fall, Waterford Institute IRI second grade students had to (a) read a story, (b) answer comprehension questions about the story, and (c) sound out nonsense words. These tasks were repeated in the winter and spring administrations. The third grade IRI was similar to the second grade edition. In the fall, the students (a) read sight words, (b) read a story, and (c) answered comprehension questions. In the winter and spring third grade IRI administrations, a spelling probe was added.

The Waterford Institute IRI was not designed to measure student growth (Idaho State Department of Education, 2004). The purpose was to assess reading proficiency on grade level material. The IRI subtests assessed the skills that were expected to be learned by students at the time of the administration.

Development of the 2007 and 2009 AIMSweb IRI Editions

In 2005, the Idaho State Department of Education decided to replace the Waterford Institute IRI (Idaho State Department of Education, 2014). This decision was based on the recommendation of the IRI Steering Committee, a group that included the State Superintendent of Public Education, SDE staff, reading specialists, test coordinators, curriculum directors, and elementary school principals (Idaho State Department of Education, 2007). The reason for the change was explained as “new reading research on critical skills indicating reading difficulties, as well as issues with test familiarity and test security” (Idaho State Department of Education, 2014). The IRI Steering Committee recommended incorporating “new research that pinpoints specific early literacy skills that are highly predictive of future reading successes into the proposed new IRI” (Idaho State Department of Education, 2007).

The SDE contacted several assessment companies and asked for a short benchmark test to meet the requirements of the Idaho Comprehensive Literacy Plan (Idaho State Department of Education, 2007). The IRI Steering Committee determined that AIMSweb IRI subtests “represent the most current scientifically-based research on effective and efficient identification of at-risk readers” (Idaho State Department of Education, 2007).

Several other possible considerations may have entered into the decision to hire Pearson to develop the new AIMSweb IRI. The Waterford Institute IRI and both versions of the AIMSweb IRI are able to trace reading development from early literacy development to reading grade level passages (Idaho State Department of Education, 2013). There is one significant difference between the earlier Waterford Institute IRI and the two versions of the AIMSweb IRI. As noted previously, the Waterford Institute IRI was not designed to measure individual student growth. The 2007 and 2009 AIMSweb IRI were designed to keep track of individual student growth (Fisk, n.d.; Idaho State Department of Education, 2013).

Another factor that may have influenced the decision to select the AIMSweb IRI may have been cost. The Waterford Institute IRI cost $3.00 per year per student for the fall, winter, and spring assessments (Idaho State Department of Education, 2007). Additional costs were added for Spanish materials. The AIMSweb proposal cost $2.00 per year per student for the three assessments, with no additional costs for Spanish materials. In addition, the package included (a) electronic progress monitoring for the students who are not at benchmark, (b) training materials for administrators and teachers, and (c) the capacity to give immediate IRI results to teachers and schools.

An additional benefit to the AIMSweb product selection is the connection to a developed RTI computer program (Idaho State Department of Education, 2007). AIMSweb developed the progress monitoring system that the SDE adopted for RTI (AIMSweb Pro Complete, 2010; Fisk, n.d.; Idaho State Department of Education, 2013). Pre-intervention and data collection under the RTI model are key components necessary before referral to special education. Students who receive Tier 2 and Tier 3 interventions receive progress monitoring through the AIMSweb progress monitoring probes. AIMSweb offers these options to school districts at additional costs.

More insight into the decision making process may be gained from a memorandum written by the State Reading Coordinator and addressed to all Building Administrators, District Test Coordinators, and School Superintendents (C. Hanson, personal communication, May 16, 2007) and a document prepared for the State Board of Education by the Idaho State Department of Education (Idaho State Department of Education, 2007). Both documents refer to the findings of the Northwest Regional Education Laboratory pilot studies that concluded that the new AIMSweb IRI was more "sensitive in identifying at risk readers in several areas as well as those of Hispanic learners and the special education population." The document titled, "Idaho Reading Indicator 2007-2008," and the Idaho State Department of Education meeting minute notes (Idaho State Department of Education, 2007) reference the findings of the pilot study as the reason for the switch from the Waterford Institute IRI to the 2007 AIMSweb IRI. The IRI Steering Committee made the recommendations after reviewing the pilot study data.

<u>The 2006 -2007 AIMSweb IRI Pilot Test and Studies</u>

In 2006 – 2007, the AIMSweb IRI was pilot tested in 37 schools. The sample represented 15 school districts around the state (Idaho State Department of Education, 2007). In a written proposal to the Idaho State Superintendent of Public School Instruction, the IRI Steering Committee recommended implementing the AIMSweb IRI in the fall of 2007. The proposal was approved unanimously by the Idaho State Board of Education (Idaho State Board of Education, 2007).

The SDE commissioned three studies by the Northwest Regional Lab to compare the Waterford Institute IRI to the AIMSweb pilot IRI (Idaho State Department of Education, 2007; Nave & Burke, 2007a, 2007b). Nave and Burke (2007a, 2007b) compared the percentages of student performance on the Waterford Institute IRI and the pilot AIMSweb IRI. The authors examined the fall, winter, and spring IRIs. Information from the fall and winter reports is available, though information from the spring pilot IRI report is not. The fall report examined two questions, while the winter report included a third question (Nave & Burke, 2007a, 2007b). The first question in each compared the achievement levels of students on the Waterford Institute IRI and the pilot AIMSweb. The second question looked at individual achievement on equivalent subtests. The third question in the winter report compared the fall and winter measures of the Waterford Institute IRI and the AIMSweb pilot IRI (Nave & Burke, 2007b).

The results of question one. The first question in both the fall and winter studies compared total achievement scores between the Waterford Institute IRI and the pilot AIMSweb IRI (Nave & Burke, 2007a, 2007b). Nave and Burke (2007a, 2007b) found that a majority (76% of the students in the fall and 78% in the winter) achieved the same proficiency level on the Waterford IRI and the pilot AIMSweb IRI. The authors compared total achievement by grade level, gender, general education, Title I schools ethnicity, and special education.

In the fall, the pilot AIMSweb IRI and the Waterford Institute IRI identified 20% of kindergarten students in Title I schools as “1” (Nave & Burke, 2007a). In grades first, second, and third, the pilot AIMSweb IRI identified more students as “1.” The winter tests differ in proportion of students reaching proficiency for grades one and two. Sixty-two percent of the students in first grade were proficient on the winter pilot AIMSweb, compared to an 84% proficiency level on the Waterford Institute IRI. In second grade, there was an almost opposite result.

Differences in the way the two measures identified proficiency rates for Hispanic students were found (Nave & Burke, 2007a, 2007b). On the fall pilot AIMSweb IRI, more Hispanic students scored at “1” across all grade levels (Nave & Burke, 2007a). More Hispanic students reached proficiency on the Waterford Institute IRI (72%) compared to 45% on the pilot AIMSweb IRI. The reverse was found for second grade. Forty-six percent of Hispanic students were proficient on the Waterford Institute IRI, compared to 61% on the pilot AIMSweb IRI.

Variations were found in the performance of special education students (Nave & Burke, 2007a, 2007b). In the fall, AIMSweb placed higher proportions of kindergarten, first, and third grade students in special education as “1” (Nave & Burke, 2007a). The proportions reversed for second grade students receiving special education. Forty-eight percent of the first grade students enrolled in special education scored proficient on the pilot AIMSweb IRI, compared to 68% on the Waterford Institute IRI. The proportion changes for second grade special education students. Fifty-one percent of second grade students enrolled in special education were proficient on the pilot AIMSweb IRI, compared to 45% on the Waterford Institute IRI. In third grade, 39% of the students in special education were scored as “1” on the pilot AIMSweb IRI, while 43% scored “1” on the Waterford Institute IRI.

The results of question two. The second question asked in both studies was concerned with performance proportions on equivalent subtests found on the Waterford Institute IRI and the AIMSweb pilot IRI (Nave & Burke, 2007a, 2007b). Comparisons could be made on six assessments across the grade levels. In this section, the subtests compared will be described by grade level. Then the results of the subtest comparison will be reported by fall and winter administrations.

Letter naming fluency was the only measure that could be compared on the fall and winter kindergarten assessments. On the Waterford Institute IRI, the fall subtest “Identify Uppercase Letters” and the winter subtest “Identify Lowercase Letters” asked the students to name as many letters on the page in one minute. The fall and winter pilot AIMSweb IRI used Letter Name Fluency to assess letter recognition and was measured in the same way as the Waterford Institute IRI.

In the fall, first grade students could be compared on letter sound fluency measurements. “Say the Letter Sound” on the Waterford Institute IRI and the “Letter Sound Fluency” on the AIMSweb required students to name as many sounds of the letters presented on a page in one minute. In the winter, two first grade reading assessments were compared. First grade oral reading fluency was compared on the “Read a Story” (Waterford Institute IRI) and “R-CBM” (pilot AIMSweb IRI). The oral reading fluency assessments measure how many words per minute are read in a passage. The second comparison made in the winter on nonsense word fluency measurements. Nonsense word fluency is assessed by having students decode made-up words. The nonsense word fluency subtests on the first grade assessments were “Sound Out Words” (Waterford Institute IRI) and “Nonsense Word Fluency” (AIMSweb IRI).

One comparison could be made between second grade students in the fall on oral reading fluency measures. “Read a Story” on the Waterford Institute IRI and the R-CBM on the AIMSweb IRI were evaluated. In the winter, two measures could be compared on second grade assessments. Oral reading fluency assessments, as mentioned earlier, were again compared. Nonsense word fluency was assessed by the subtests called “Sound Out Words” (Waterford Institute IRI) and “Nonsense Word Fluency” (AIMSweb).

In the fall, two third grade subtests were compared. They were oral reading fluency and reading comprehension measures. Oral reading fluency was assessed by comparing “Read a Story” (Waterford IRI) and R-CBM (AIMSweb). Reading comprehension was assessed on the Waterford IRI by having students answer questions related to the reading passage, called “Answer Comprehension Questions.” A Maze reading CBM was used to assess reading comprehension on the pilot AIMSweb IRI. On the Maze CBM, every seventh word in the passage is replaced with a choice of the correct word and two foils (Consortium on Reading Excellence, 2008). Three measures were matched in the winter. Oral reading fluency and the reading comprehension subtests were administered again in the winter. The Spelling subtest of the Waterford Institute IRI and the Spelling CBM of the pilot AIMSweb IRI were also examined.

The results of the comparison of the Waterford Institute IRI and the pilot AIMSweb indicated that in the fall the majority of students achieved at the same levels on both versions of the IRI (Nave & Burke, 2007a). In spite of this finding, differences were found on some sets of comparable subtests. The agreement between the Waterford Institute IRI “Identify Uppercase Letters” and AIMSweb “Letter naming Fluency” was 79%. An agreement of 69% was found between the Waterford Institute IRI “Say the First Sound” and the AIMSweb “Letter Sound Fluency for first grade. There was an 85% agreement in the performance of second grade students on the Waterford Institute IRI “Read a Story” and the AIMSweb “R-CBM.” In contrast, there was a 64% agreement on the Waterford Institute IRI “Sound Out Words” and AIMSweb “Nonsense Word Fluency.” In third grade, there was a 77% agreement between “Read a Story” (Waterford Institute IRI) and the R-CBM (AIMSweb pilot IRI). The weakest relationship of 57% was between the Waterford Institute IRI “Answer the Questions” and the AIMSweb “Maze.” Table 2.4 provides a side-by-side list of the subtests compared in the fall.

Table 2.4

Fall 2006, Matching Skill Assessments on Waterford Institute IRI and Pilot AIMSweb IRI

Grade	Pilot AIMSweb Subtest	Waterford Institute Subtest	Agreement
K	Letter Naming Fluency	Identify Upper Letters	79%
1	Letter Sound Fluency	Say the First Sound	69%
2	R-CBM	Read a Story	85%
2	Nonsense Word Fluency	Sound Out Words	64%
3	R-CBM	Read a Story	77%
3	Maze	Answer Comprehension Questions	57%

In the winter, the comparison of the subtests ranged from 54% agreement to 88% (Nave & Burke, 2007b). In kindergarten, there was a 76% agreement between the Waterford Institute IRI “Identify Lowercase Letters” and AIMSweb “Letter Naming Fluency.” There was a 77% agreement between the Waterford Institute IRI “Read a Story” and AIMSweb “R-CBM” measures for first grade students. The agreement between the Waterford Institute IRI “Sound Out Words” and AIMSweb “Nonsense Word Fluency” for first grade students was 71%. In second grade, the agreement between the Waterford Institute IRI “Read a Story” and AIMSweb “R-CBM” was 88%. The agreement between the Waterford Institute IRI “Sound Out Words” and AIMSweb “Nonsense Word Fluency” was 57%. Three subtests were compared in the winter for third grade. The agreement between the Waterford Institute IRI “Read a Story” and AIMSweb “R-CBM” was 81%. On the subtests Answer Comprehension Questions (Waterford Institute IRI) and Maze (AIMSweb pilot IRI) was 54%. On third grade spelling measures, the agreement was 62%. Table 2.5 provides a summary of subtest comparisons.

Table 2.5

Winter 2007, Matching Skill Assessments on Waterford Institute IRI and Pilot AIMSweb IRI

Grade	Pilot AIMSweb Subtest	Waterford Institute Subtest	Agreement
K	Letter Naming Fluency	Identify Lowercase Letters	76%
1	R-CBM	Read a Story	77%
1	Nonsense Word Fluency	Sound out Words	71%
2	R-CBM	Read a Story	88%
2	Nonsense Word Fluency	Sound out Words	57%
3	R-CBM	Read a Story	81%
3	Maze	Answer Comprehension Questions	54%
3	Spelling-CBM	Spelling	62%

The results of question three. The third question compared the fall and winter Waterford Institute IRI and the AIMSweb pilot IRI (Nave & Burke, 2007b). Students on both the Waterford Institute IRI and the AIMSweb pilot IRI reached proficiency of an “overwhelming majority” of 92%. The rates of students moving from “2” in the fall to proficient in the winter on the Waterford Institute IRI and the AIMSweb pilot IRI were similar at 45% and 47%, respectively. Sixty percent of the students who had “1” on the fall Waterford Institute IRI obtained a score of “1” in the winter. Thirty percent of the students who obtained a score of “1” on the fall Waterford Institute IRI grew to a score of “2” in the winter. Ten percent of the students who had a “1” on the fall Waterford Institute IRI obtained the proficient level of achievement in the winter. On the AIMSweb pilot IRI, 54% of the students who scored “1” in the fall remained unchanged in the winter. Students who had a “1” on the fall AIMSweb pilot IRI showed growth to a score of “2” at a rate of 37% in the winter. Students with scores of “1” in the fall on the AIMSweb pilot IRI reached proficiency at a rate of 9% in the winter.

The 2007 to 2009 AIMSweb IRI

When the AIMSweb IRI was first administered during the 2007 – 2008 school year, it had fewer subtests than the Waterford Institute IRI. Two subtests (the Maze and Spelling CBMs) that were on the pilot AIMSweb IRI were removed (S. Lee, personal communication, August 27, 2012). The reasons for the removal of these subtests were to cut down the amount of testing time and eliminate assessments that measure similar skills (M. Shinn, personal communication, August 2, 2013). Although the R-CBM (oral reading fluency measurement) is not a measure of comprehension, it has been found to correlate well with reading fluency measures (Muyskens, Betts, Lau, & Marston, 2009; Roehring, Petscher, Nettles, Hudson, & Torgesen, 2008; Shinn, Good, Knutson, Tilly, & Collins, 1992; Spear-Swerling, 2006). Since R-CBM can be used as an indicator of reading comprehension, the Maze was dropped (M. Shinn, personal communication, August 2, 2013).

The AIMSweb IRI uses a scoring system of “1,” “2,” or “3” (Idaho State Department of Education, 2013). The three point scoring system remained in place, however the classification terms changed. The new classification terms were “benchmark,” “strategic,” and “intensive” instead of “at grade level,” “near grade level,” and “below grade level” (Idaho State Department of Education, 2013). This three level classification terminology is consistent with the RTI three tier model.

Benchmark is defined as the expected level of achievement to be rated as proficient at grade level (Idaho State Department of Education, 2013). A student score of "3" indicates "mastery of the skills" and is considered "proficient." In the RTI model, students who fall within the "proficient" category receive core instruction in general education settings without additional interventions. The Waterford Institute IRI classified students with a score of "3" as "at grade level."

A student score of "2" indicates "partial mastery of some or all skills" and is referred to as "strategic" (Idaho State Department of Education, 2013). In the RTI model, students who fall within the "strategic" category may need additional interventions to the core instruction. A score of two was called "near grade level" on the Waterford Institute IRI.

A student score of "1" indicates "a lack of mastery of some or all skills" and is referred to as "intensive" (Idaho State Department of Education, 2013). In the RTI model, students who fall within the "intensive" category require additional interventions as supplements to the core instruction. Students who receive a score of one on the IRI must be offered extended reading interventions. A score of one was called "below grade level" on the Waterford Institute IRI.

The 2007 AIMSweb IRI offered revisions to the subtests and cut scores for each grade level. In kindergarten, reading skills are assessed with Letter Name Fluency in the fall. In the winter and spring IRI, Letter Name Fluency was retested and included with Letter Sound Fluency and Phoneme Segmentation Fluency. Phoneme Segmentation Fluency is a measure of phonemic awareness (Fisk, n.d.). In this subtest, students are given words with three or four syllables and asked to break them apart at the individual sound level (Kaminski & Good, 1996). Table 2.6 lists the subtests found on the fall, winter, and spring kindergarten IRI. The cut scores and converted IRI scores reported on the "Idaho Indicator Student Report Card" are included on Table 2.6. Cut scores are used to classify student performance (Zieky & Perie, 2006). Cut scores are selected points on a test scale. Cut scores are commonly used on statewide, reading achievement measures. Cut scores are set with the involvement of stakeholders, including policymakers, educators, and consultants who are experts on measurements. Accepted methods and the judgments of the stakeholders are typically used to determine cut scores. Not all of the results for all of the subtests were reported on the "Idaho Indicator Student Report Card."

Table 2.6

2007 AIMSweb IRI Kindergarten Cut Scores

Subtest	Fall Scores	Winter Scores	Spring Scores
Letter Name Fluency	0 - 2 = 1	0 - 18 = 1	0 - 30 = 1
	0 - 10 = 2	19 - 32 = 2	31 - 42 = 2
	11 - 100 = 3	33 - 100 = 3	43 - 100 = 3
Letter Sound Fluency		0 - 6 = 1	0 - 17 = 1
		7 - 16 = 2	18 - 29 = 2
		17 - 100 = 3	30 - 100 = 3
Phoneme Segmentation		0 - 5 = 1	0 - 19 = 1
		6 - 16 = 2	20 - 39 = 2
		17 - 93 = 3	40 - 99 = 3

Note. Benchmark = 3; Strategic = 2; Intensive = 1.

On the first grade 2007 IRI, Letter Name Fluency and Phoneme Segmentation Fluency were assessed again (Fisk, n.d.). Nonsense Word Fluency was added to the battery. Letter Sound fluency was omitted in first grade because the same skills could be assessed with Nonsense Word Fluency, since students were able to respond with isolated sounds of the letters or read the whole word on the Nonsense Word Fluency measure. In the winter and spring, R-CBMs (oral reading fluency) were assessed along with Phoneme Segmentation Fluency and Nonsense Word Fluency. Three R-CBM passages were given to first grade students in the winter and spring. The same three passages were used in the winter and spring to measure growth. Letter Name Fluency was not reassessed in the winter and spring. The medium raw score was used to determine the IRI score of “1,” “2,” or “3” on this subtest. Table 2.7 provides a list of the subtests found on the first grade fall, winter, and spring IRI. The cut scores and converted IRI score reported on the “Idaho Indicator Student Report Card” are included on Table 2.7.

Table 2.7

2007 AIMSweb IRI First Grade Cut Scores

Subtest	Fall Scores	Winter Scores	Spring Scores
Letter Name Fluency	0 - 29 = 1		
	30 - 41 = 2		
	42 - 100 = 3		
Phoneme Segmentation	0 - 16 = 1	0 - 33 = 1	0 - 40 = 1
	17 - 34 = 2	34 - 46 = 2	41 - 50 = 2
	35 - 98 = 3	47 - 98 = 3	51 - 98 = 3
Nonsense Word Fluency	0 - 13 = 1	0 - 31 = 1	0 - 39 = 1
	14 - 24 = 2	32 - 43 = 2	40 - 55 = 2
	25 - 220 = 3	44 - 220 = 3	56 - 220 = 3
R-CBM		0 - 12 = 1	0 - 27 = 1
		13 - 22 = 2	28 - 52 = 2
		23 - 250 = 3	53 - 220 = 3

Note. Benchmark = 3; Strategic = 2; Intensive = 1

The second grade fall IRI measured Nonsense Word Fluency and R-CBM in the fall (Fisk, n.d.). Nonsense Word Fluency was not included in the winter and spring second grade IRI. The second grade (winter, spring) and third grade (fall, winter, spring) IRIs contained R-CBMs. Three R-CBM passages were given to second grade students and three passages were given to third grade students in the fall, winter, and spring. The same three passages were used in the fall, winter, and spring to measure growth. The medium raw score was used to determine the IRI score of “1,” “2,” or “3” on the R-CBM.

The cut scores and converted IRI score reported on the "Idaho Indicator Student Report Card" are included on Table 2.8.

Table 2.8

2007 AIMSweb IRI Second and Third Grade Cut Scores

Grade	Subtest	Fall Scores	Winter Scores	Spring Scores
Second	Nonsense Word Fluency	0 - 32 = 1		
		33 - 48 = 2		
		49 - 220 = 3		
	R-CBM	0 - 26 = 1	0 - 51 = 1	0 - 67 = 1
		27 - 53 = 2	52 - 76 = 2	68 - 91 = 2
		54 - 229 = 3	77 - 229 = 3	92 - 229 = 3
Third	R-CBM	0 - 48 = 1		0 - 81 = 1
		49 - 76 = 2		82 - 109 = 2
		77 - 229 = 3		110 - 298 = 3

Note. Benchmark = 3; Strategic = 2; Intensive = 1. NWF = Nonsense Word Fluency.

2007 AIMSweb IRI Research

One study (Stewart, 2009) examined how well the 2007 fall, winter, and spring AIMSweb IRI versions predicted performance on the Idaho Standards Achievement Tests **(**ISATS). The ISATS were Idaho's high stakes statewide assessments (Idaho Standards Achievement Tests, n.d.). The ISATS were administered to students in third grade through tenth grade. The ISATS measured academic proficiency in reading, language usage, and mathematics.

The study was conducted over the course of the 2008 though 2009 school year (Stewart, 2009). The AIMSweb IRI scores of 17,000 third grade students and ISATS were examined in the study. The AIMSweb IRI scores from three assessments (fall, winter, and spring) were correlated to the ISATS scores (Stewart, 2009). The Pearson Product Moment Correlations between the IRI and ISATS were .66 in the fall, .67 in the winter, and .67 in the spring. It was found that the IRI correctly predicted student performance for all but 15.8% of the students. The strongest correlation between the IRI and the ISATS was found among white students who were not enrolled in special education or attending Title I schools.

Changes to the 2007 AIMSweb IRI

The major change on the 2007 AIMSweb IRI implemented in 2009 was the reduction of the number of subtests on the first and second grade IRI. Specifically, the Phoneme Segmentation Fluency and Nonsense Word Fluency subtests were removed. Two reasons emerge for the changes to the 2007 AIMSweb IRI: (a) time required for administration of the subtests, and (b) psychometric properties of the subtests (M. Shinn, personal communication, August 2, 2013). Although the assessments were short, Idaho educators expressed concerns that they were spending too much time testing. Phoneme Segmentation Fluency and Nonsense Word Fluency subtests were cut to shorten testing time.

Some subtests were found to be better predictors and others duplicated skills assessed by stronger measures. For kindergarten students, Letter Name Fluency was found to be a better predictor of early reading ability than Phoneme Segmentation Fluency (S. Lee, personal communication, August 27, 2012). Phoneme Segmentation had poor reliability and not enough validity on the 2007 AIMSweb IRI. Phoneme Segmentation Fluency is more difficult to administer and score than Letter Sound

Fluency (Garland & Strosnider, 2005). Letter Sound Fluency is the “single best indicator of early reading” for young students who have mastered Letter Name Fluency. Measures of phonological awareness do not seem as sensitive for identifying students who are at risk for reading failure (Scarborough, 1998; Schatschneider, Fletcher, Francis, Carlson, & Foorman, 2004; Speece & Case, 2001). Phonological awareness assessments in kindergarten are good predictors of superior reading ability but not of future reading problems (Scarborough, 1998). Phoneme segmentation fluency measured in the middle and end of first grade was found to be the weakest predictor of reading comprehension performance in first and second grade students (Riedel, 2007).

Nonsense Word Fluency was determined to be unnecessary because it measured similar skills as Letter Sound Fluency (S. Lee, personal communication, August 27, 2012). Ritchey’s (2008) research may have been the basis for this decision. Nonsense word fluency and letter sound fluency were found to be comparable in identifying and predicting concurrent and future reading difficulty in kindergarten students. It was found that students who were at risk for reading failure scored poorly on nonsense word fluency and letter sound fluency assessments. Letter sound fluency was a less complicated means to assess letter sound knowledge. Furthermore, educators criticized using nonsense words because too many teachers practiced reading non-real words with their students at the expense of learning to read real words. Nonsense words were found to only moderately correlate with the end-of-first-grade outcomes; thus. the value of the practice was questioned (Dion, Dube, Roux, Landry, & Bergeron, 2012; L.S. Fuchs, Fuchs, & Compton, 2004; Riedel, 2007). Consequently, the Idaho Reading Indicator Steering Committee, Dr. Mark Shinn, and the AIMSweb researchers recommended dropping

Phoneme Segmentation Fluency and Nonsense Word Fluency from the IRI (S. Lee, personal communication, August 27, 2012). Table 2.9 shows the changes in the subtests used in the 2007 and 2009 AIMSweb IRI for grades kindergarten through second. The third grade IRI continued to use oral reading fluency measures.

Table 2.9

Subtest Changes from 2007 to 2009 for Grades Kindergarten through Second on the AIMSweb IRI

Fall	Winter	Spring
	Kindergarten 2007 AIMSweb	
Letter Name Fluency	Letter Name Fluency	Letter Name Fluency
	Letter Sound Fluency	Letter Sound Fluency
	Phoneme Segmentation Fluency	Phoneme Segmentation Fluency
	Kindergarten 2009 AIMSweb	
Letter Name Fluency	Letter Name Fluency	Letter Name Fluency
Letter Sound Fluency	Letter Sound Fluency	Letter Sound Fluency
	First Grade 2007 AIMSweb	
Letter Name Fluency		
Phoneme Segmentation Fluency	Phoneme Segmentation Fluency	Phoneme Segmentation Fluency
Nonsense Word Fluency	Nonsense Word Fluency	Nonsense Word Fluency
	R-CBM	R-CBM
	First Grade 2009 AIMSweb	
Letter Sound Fluency	Letter Sound Fluency	Letter Sound Fluency
R-CBM	R-CBM	R-CBM

Second Grade 2007 AIMSweb		
Nonsense Word Fluency		
R-CBM	R-CBM	R-CBM
Second Grade 2009 AIMSweb		
R-CBM	R-CBM	R-CBM

Finally, the Letter Name Fluency, Letter Sound Fluency, and R-CBMs for all grades were changed on the 2009 IRI. The 2007 IRI contained benchmark probes found in the AIMSweb system. All authorized school district users of the AIMSweb system had access to the benchmark probes. AIMSweb could not secure the probes within the system (Idaho State Department of Education, 2013). In 2009, the SDE purchased unique benchmark probes and proprietary ownership from AIMSweb to secure the probes, in accordance with the Idaho Comprehensive Literacy Act (Idaho Code 33-1614, Idaho State Department of Education, 2013). The 2009 IRI benchmark probes are now the intellectual property of the Idaho SDE.

Current AIMSweb IRI

As previously reported, the current AIMSweb came into existence in 2009**.** The 2009 AIMSweb IRI simply will be referred to as the IRI from this point. The number of times per year that the IRI was to be administered changed in 2010. Since the beginning, the Waterford Institute IRI and the 2007 AIMSweb IRI were required to be administered three times per year (Stewart, 2010). In the 2010-2011 school year, the requirement to give the winter IRI was dropped because of budget cuts to education (Idaho State Department of Education, 2014b; S. Lee, personal communication, July, 23, 2013).

School districts may still administer the winter IRI for their own purposes and report the results to the state. Unlike the fall and spring IRI results, the winter results are not reported to the public (S. Lee, personal communication, February, 4, 2014). School districts do not receive reimbursement for the winter assessment.

The kindergarten IRI assesses reading with reading readiness and phonological awareness measures (Idaho State Department of Education, 2013). In the first grade, the IRI assesses letter sound and reading fluency (rate and accuracy). The IRI for grades one through three assesses reading fluency using a set of three passages at each grade level (Idaho State Department of Education, 2013). The same passages are given during the fall and spring, with an optional winter test administration. Students are timed for one minute on each task. (Each grade level IRI will be described with more specific information following this section.)

The three point scoring system is retained in the current IRI. The categories correspond to the three RTI tiers. A score of “3” is called **“**benchmark” and defined as reading skills at or above grade level (Idaho State Department of Education, 2013). A student score of “2” is called “strategic.” A student score of “1” is called “intensive.”

The Kindergarten IRI

The kindergarten IRI was briefly described earlier as an assessment of reading readiness and phonological awareness (Idaho State Department of Education, 2013). In the fall, kindergarten students are given the Letter Name Fluency and Letter Sound Fluency measurements. Letter Name Fluency and Letter Sound Fluency are presented with upper and lower case letters (Idaho State Department of Education, 2013). Both assessments are administered again in the spring.

To measure Letter Name Fluency, kindergarten students are asked to name as many letter names on a sheet of letters in one minute. The Letter Name Fluency assessments used in the fall and spring contain 100 mixed upper case and lower case letters. The fall Letter Name Fluency measure is different from the spring version.

In the Letter Sound Fluency subtest, students are given a similar page of printed letters, but instead are asked to produce the sound of each letter, again with a one minute limit. Students are asked to say the “sounds of the letters” on the Letter Sound Fluency subtest. The number of letter-sound correspondences a student makes in one minute is recorded as the fluency rate (University of Oregon, n.d.). The Letter Sound Fluency assessment contains 100 lower case letters. A different version of the Letter Sound Fluency containing all lower case letters is administered in the spring. Table 2.10 lists the subtests cut scores and converted score (intensive, strategic, and benchmark).

Table 2.10

Current Kindergarten IRI Subtests, Cut Scores, and Converted Scores

Subtest	Fall Cut Scores	Spring Cut Scores
Letter Name Fluency	0 - 2 = 1	0 - 30 = 1
	3 - 10 = 2	31 - 42 = 2
	11 - above = 3	43 – above = 3
Letter Sound Fluency	0 - 0 = 1	0 – 17 = 1
	1 - 1 = 2	18 - 29 = 2
	2 - above = 3	30 – above = 3

Note. Benchmark = 3; Strategic = 2; Intensive = 1.

The First Grade IRI

The first grade Idaho Reading Indicator (IRI) has two components, Letter Sound Fluency and Reading Curriculum Based Measure (R-CBM, Idaho Reading Indicator (IRI): Parent Brochure, n.d.). The R-CBM used in the IRI is an oral reading fluency measure. The Letter Sound Fluency assessment is presented with 100 lower case letters. A different version is given in the spring, also presented in lower case letters. The first grade Letter Sound Fluency assessments do not contain the same letter arrangements as the kindergarten Letter Sound Fluency presentations. The same set of three reading passages is used for the R-CBM in the fall, winter, and spring first.

Permission was obtained from Stephanie Lee, Assessment Specialist, Idaho State Department of Education, to conduct a readability analysis of the three reading passages. The passage information and readability assessments are found on Table 2.11. The information is also given in narrative form in the subsequent paragraphs.

The first passage of the first grade IRI R-CBM contains a total of 256 words. Forty-five percent of the words are unique and 54% are repeated words (Scott, n.d.). There is an average of seven words per sentence and a total of 36 sentences. Eight-six percent of the words contain single syllables. Fourteen percent of the words contain two syllables. There are no words with three or more syllables. The Flesch-Kincaid Grade Level readability is .6, at the first grade level. The Flesch Reading Ease rating is 103.5 or “very easy.” On the Fry Graph Readability Formula, the passage falls in the first grade level, with an average of 113 syllables per 100 words and 14.1 sentences per 100 words.

Passage two of the first grade IRI R-CBM contains a total of 260 words. Fifty-one percent of the words are unique and 49 percent are repeated words (Scott, n.d.). There is an average of seven words per sentence and a total of 36 sentences. Eight-five percent of the words contain single syllables. Fifteen percent of the words contained two syllables. Three words have three or more syllables, representing zero percent of the passage. The Flesch-Kincaid Grade Level readability is .9, at the first grade level. The Flesch Reading Ease rating is 101.2 or “very easy.” The passage falls within the first grade range on the Fry Graph. There is an average of 117 syllables per 100 words and 13.8 sentences per 100 words.

The third passage of the first grade IRI R-CBM contains a total of 246 words. Forty-six percent of the words are unique and 54% are repeated words (Scott, n.d.). There is an average of seven words per sentence and a total of 36 sentences. Eight-one percent of the words contain single syllables. Eleven percent of the words contained two syllables. One word has three or more syllables, representing zero percent of the passage. The Flesch-Kincaid Grade Level readability is .3, at the kindergarten grade level. The Flesch Reading Ease rating is 105 or “very easy.” On the Fry Graph Readability Formula, the passage falls in the first grade level, with an average of 115 syllables per 100 words and 14.6 sentences per 100 words.

Table 2.11

Readability of the Three First Grade IRI Passages

Passage Information & Readability	Passage 1	Passage 2	Passage 3
Total Words	256	260	246
Average Words per Sentence	7	7	7
Total Number of Sentences	36	36	36
Percentage of Unique Words	45%	51%	46%
Percentage of Repeated Words	54%	49%	54%
Percentage of Single Syllable Words	86%	85%	81%
Percentage of Two Syllable Words	14%	15%	11%
Percentage of Three Syllable Words	0%	0%	0%
Flesch Reading Ease	103.5	101.2	105.0
Flesch Reading Ease Rating	"very easy"	"very easy"	"very easy"
Flesch-Kincaid Grade Level Readability	.6	.9	.3
Flesch-Kincaid Grade Level Readability Grade Level	First	First	Kindergarten
Fry Graph Readability Formula Grade Level	First	First	First
Average syllables per 100 words	113	117	115
Average sentences per 100 words	14	14	15

The R-CBM is administered in the fall as a baseline measure (Idaho State Department of Education, 2013). The score of the R-CBM is not counted towards reaching proficiency targets in the fall. The R-CBM is used to measure proficiency in the spring. Students are asked to read R-CBMs in one minute. Three passages are used to assess oral reading fluency. The medium score (middle number) of the three reading passages is recorded. The same passages are used in the winter and spring IRIs. Table 2.12 provides a list of subtests and benchmark targets for the fall and spring.

Table 2.12

Current First Grade IRI Subtests, Cut Scores, and Converted Scores

Subtest	Fall Cut Scores	Spring Cut Scores
Letter Sound Fluency	0 - 19 = 1	0 - 51 = 1
	20 - 30 = 2	52 - 71 = 2
	31 - above = 3	72 - above = 3
R-CBM	0 - 0 = 1	0 - 27 = 1
	1 - 1 = 2	28 - 52 = 2
	2 - above = 3	53 - above = 3

Note. Benchmark = 3; Strategic = 2; Intensive = 1.

The Second and Third Grade IRI

Only oral reading fluency is assessed on second and third grade students. Each grade level has a separate set of three R-CBM passages. The administration is the same as described for first grade students. Table 2.13 provides a list of subtests and benchmark targets for the fall and spring.

Table 2.13

Current Second and Third Grade IRI Subtests, Cut Scores, and Converted Scores

Grade	Subtest	Fall Cut Scores	Spring Cut Scores
Second	R-CBM	0 - 26 = 1	0 - 67 = 1
		27 - 53 = 2	68 - 91 = 2
		54 - above = 3	92 - above = 3
Third	R-CBM	0 - 48 = 1	0 - 81 = 1
		49 - 76 = 2	82 -109 = 2
		77 - above = 3	110 - above = 3

Note. Benchmark = 3; Strategic = 2; Intensive = 1.

Permission was obtained to conduct readability assessments of the R-CBM passages of the second and third grade IRI. Passage one of the second grade IRI contains a total of 258 words. Table 2.14 contains the passage information and readability results for the second grade R-CBMs. The information is also presented in the following paragraphs.

The average number of words per sentence is nine in passage one of the second grade IRI (Scott, n.d.). There are a total of 31 sentences. Fifty-five percent of the passage contains unique words. Forty-five percent of the passage contains repeated words. Eighty-one percent of the words are single syllable words. Nineteen percent of the passage has double syllable words. There are no words with three syllables or more. The Flesch-Kincaid Grade Level readability is 1.9, at the second grade level. The Flesch Reading Ease rating is 97 or “very easy.” The passage falls within the second grade range with the Fry Graph Readability Formula. There is an average of 122 syllables per 100 words and 11.7 sentences per 100 words.

The second passage of the second grade IRI contains a total of 250 words. The average number of words per sentence is eight (Scott, n.d.). There are a total of 31 sentences. Fifty-three percent of the passage contains unique words. Forty-seven percent of the passage contains repeated words. Eighty-three percent of the words are single syllable. Sixteen percent of the passage has two syllable words. There are three words with three or more syllables and this represents 1% of the passage. The Flesch-Kincaid Grade Level readability is 1.6, at the second grade level. The Flesch Reading Ease rating is 98.3 or "very easy." On the Fry Graph, the passage falls in the second grade level, with an average of 122 syllables per 100 words and 12.1 sentences per 100 words.

Passage three of the second grade IRI contains a total of 253 words. The average number of words per sentence is eight (Scott, n.d.). There are a total of 31 sentences. Fifty-four percent of the passage contains unique words. Forty-five percent of the passage contains repeated words. Eighty-six percent of the words are single syllable. Fourteen percent of the passage has double syllable words. There are no words with three syllables or more. The Flesch-Kincaid Grade Level readability rating is 1.9 at the second grade level. The Flesch Reading Ease is 102.2, or "very easy." This passage falls within the second grade range on the Fry Graph. There is an average of 121 syllables per 100 words and 12.1 sentences per 100 words.

Table 2.14

Readability of the Three Second Grade IRI Passages

Passage Information & Readability	Passage 1	Passage 2	Passage 3
Total Words	258	250	253
Average Words per Sentence	9	8	8
Total Number of Sentences	31	31	31
Percentage of Unique Words	55%	53%	54%
Percentage of Repeated Words	45%	47%	45%
Percentage of Single Syllable Words	81%	83%	86%
Percentage of Two Syllable Words	19%	16%	14%
Percentage of Three Syllable Words	0%	1%	0%
Flesch Reading Ease	97.0	98.2	102.2
Flesch Reading Ease Rating	"very easy"	"very easy"	"very easy"
Flesch-Kincaid Grade Level Readability	1 .9	1 .6	1 .9
Flesch-Kincaid Grade Level Readability Grade Level	Second	Second	Second
Fry Graph Readability Formula Grade Level	Second	Second	Second
Average syllables per 100 words	122	122	121
Average sentences per 100 words	12	12	12

Table 12.15 summarizes the passage information and readability results for the third grade IRI R-CBM passages. The first passage of the third grade IRI contains a total of 271 words. The average number of words per sentence is 11 (Scott, n.d.). There are a total of 25 sentences. Fifty-five percent of the passage contains unique words. Forty-five

percent of the passage contains repeated words. Seventy-nine percent of the words are single syllable words. Eighteen percent of the passage has two syllable words. Three percent of the passage contains three or more syllables with eight words. The Flesch-Kincaid Grade Level readability is 3.4, at the third grade level. The Flesch Reading Ease rating is 90.5 or "very easy" to read. The passage falls within the fourth grade level on the Fry Graph. There is an average of 129 syllables per 100 words and 9 sentences per 100 words. The passage falls within the third grade level on the Raygor Graph. The average number of six or more characters per 100 words is 18. The average number of sentences per 100 words is 9.

Passage two of the third grade IRI contains a total of 342 words. The average number of words per sentence is eleven. There are a total of 32 sentences (Scott, n.d.). Fifty-five percent of the passage contains unique words. Forty-five percent of the passage contains repeated words. Eighty-five percent of the words are single syllable. Twelve percent of the passage has two syllable words. There are thirteen words with three or more syllables and this represents 4% of the passage. The Flesch-Kincaid Grade Level readability is 2.8, at the third grade level. The Flesch Reading Ease rating is 94.7 or "very easy." The passage falls within the third grade reading range on the Fry Graph. There is an average of 124 syllables per 100 words and 9.1 sentences per 100 words. The passage falls within the third grade level on the Raygor Graph. The average number of six or more characters per 100 words is 14. The average number of sentences per 100 words is 12.1.

The last third grade passage contains a total of 314 words. The average number of words per sentence is eleven (Scott, n.d.). There are a total of 30 sentences. Fifty-one percent of the passage contains unique words. Forty-nine percent of the passage contains repeated words. Eighty-one percent of the words are single syllable. Seventeen percent of the passage has two syllable words. Five words have three syllables or more, representing 2% of the passage. The Flesch-Kincaid Grade Level readability rating is 2.8 at the third grade. The Flesch Reading Ease is 94.2, or “very easy.” The passage falls within the fourth grade level on the Fry Graph. There is an average of 129 syllables per 100 words and 9.4 sentences per 100 words. The passage falls within the third grade level on the Raygor Graph. The average number of six or more characters per 100 words is 18. The average number of sentences per 100 words is 9.4.

Table 2.15

Readability of the Three Third Grade IRI Passages

Passage Information & Readability	Passage 1	Passage 2	Passage 3
Total Words	271	342	314
Average Words per Sentence	11	11	11
Total Number of Sentences	25	32	30
Percentage of Unique Words	55%	55%	51%
Percentage of Repeated Words	45%	45%	49%
Percentage of Single Syllable Words	79%	85%	81%
Percentage of Two Syllable Words	18%	12%	17%
Percentage of Three Syllable Words	3%	4%	2%
Flesch Reading Ease	90.5	94.7	94.2
Flesch Reading Ease Rating	“very easy”	“very easy”	“very easy”
Flesch-Kincaid Grade Level Readability	3.4	2.8	2.8
Flesch-Kincaid Grade Level Readability	Third	Third	Third

Grade Level			
Fry Graph Readability Formula Grade Level	Fourth	Third	Fourth
Average syllables per 100 words	129	124	129
Average sentences per 100 words	9	9	9
Raygor Graph Readability Formula Grade Level	Third	Third	Third
Six or more characters per 100 words	18	14	18
Average sentences per 100 words	9	12	9

The Future of the IRI

Funding to the Idaho Reading Initiative was reduced in 2009-2010, and has remained flat since (S. Lee, personal communication, July, 23, 2013). Savings had to be found. In 2013, the Idaho State Department of Education (SDE) did not renew the contract with AIMSweb. Instead, the SDE created a site to house IRI data, the IRI Web Application. The SDE possesses proprietary ownership of the IRI benchmark probes; therefore, the SDE was able to continue to use the current IRI without violating contract agreements with AIMSweb. The IRI measurements and cut scores were unchanged during the 2011-2012 through 2013-2014 assessment rounds.

The decision to keep the data within the SDE, rather than hire an outside vendor, saved the state more than $145,000. Some of the savings will go to school districts to defray IRI assessment and intervention costs. A portion of the savings is being used to hire consultants to investigate best practices for early literacy reading assessments and to make recommendations to the SDE Assessment Department staff. The recommendations of the consultants may signal significant changes to the IRI in the future.

In 2012, the SDE commissioned a study to review the IRI and the influence of the IRI on current practices in the identification of students at risk for potential reading problems (Santi & Francis, 2012). The authors provided several recommendations. First, the IRI was used for a variety of purposes that may conflict with one another. The IRI is used to identify students at risk for reading difficultly and also to evaluate teacher performance. The IRI fails to deliver the scope and precision required of an outcome assessment for teacher accountability because the IRI is a quick screening of reading ability. The legislative intent of the IRI was to give teachers information about the reading skills of students. To use it to measure student improvement as a means of teacher accountability is in conflict with the original legislative intent because teachers feel pressure to teach to the test.

Santi and Francis (2012) found problems with the administration and scoring of the IRI. The IRI does not contain comprehension questions. The authors question the use of the median score, or a single raw score from the three reading samples. Santi and Francis (2012) suggested having multiple and equated passages and different forms of the IRI. The interpretation of IRI scores for state standards of reading proficiency was not supported. The IRI cannot adequately classify students into proficiency categories without a standard setting process to determine performance level descriptors. Overall, the current proficiency indicators are consistent with the norms from similar assessments with two exceptions. First, the samples used in the reports were not representative of the entire Idaho school populations. Secondly, only one version of the IRI exists for each grade level, which narrows the scope of assessment. The authors expressed concerns over the limited psychometric information of the IRI.

Santi and Francis (2012) recommended that the IRI be used as a reading screener with consideration to modifying test administration. The implementation of systematic data collection was recommended to establish the validity and reliability of decisions based on the IRI. Finally, the key purpose of the IRI must be determined. It was also recommended that a standardized diagnostic assessment be adapted for students who fail the IRI.

The report by Santi and Francis (2012) was given to the Idaho Literacy Technical Advisory Committee for consideration as the committee develops recommendations on how to continue to improve reading proficiency in Idaho (Idaho State Department of Education, 2014a). The Idaho Literacy Technical Advisory Committee met from June of 2014 to November of 2014. The Idaho Literacy Technical Advisory Committee gave recommendations in a report (Idaho State Board of Education, 2014) to the Idaho State Board of Education in November of 2014. The Idaho Literacy Technical Advisory Committee recommended that the IRI be used as a universal screen and not as a teacher accountability measure. Furthermore, the Idaho Literacy Committee recommended that the IRI be reviewed to address concerns about its technical adequacy and explore alternative measures (Idaho State Board of Education, 2014).

Curriculum-Based Measurements of Reading

Curriculum-based measurements (CBM) are standardized procedures designed to accurately and efficiently access academic status and growth in the basic skills of reading, mathematics, written expression, and spelling (Christ & Silberglitt, 2007; Deno, 1985; Shinn, 1989, 1998). Most available CBM research is focused on oral reading fluency (D. Fuchs, Fuchs, & Compton, 2004; L. S. Fuchs, et al., 2004). CBMs have standard administration directions; are timed; and have scoring instructions, including judgment standards, record forms and charts. Examples of CBMs include (a) timed passages for oral reading fluency, (b) lists of letters for letter sound fluency, and (c) letter naming fluency. CBMs are designed to be a quick and efficient means of assessing the current level of student functioning on tasks that match instructional goals (Hosp, Hosp, & Howell, 2007; Jenkins & Fuchs, 2012).

CBMs have many purposes. The CBM was first used by special educators to frequently measure student performance and evaluate instructional effectiveness (Bradley & Ames, 1977; Christ & Silberglitt, 2007; Deno, 1985, 1986, 1989; Deno, Marston, & Tindal, 1985; Shinn, 1989). CBMs are now used in school-based and district-wide assessment systems (Marston, 2012). CBMs are used to monitor student progress within the RTI framework (Mellard & Johnson, 2008; Hosp et al., 2007). CBMs may also be used as universal screens to identify students who are at-risk for school failure (Jenkins & Fuchs, 2012).

The IRI is a collection of CBMs. Grades kindergarten through first employ the early reading CBMs of letter name fluency and letter sound fluency. Three oral reading fluency CBMs are given to students in first through third grades. This section addresses the CBMs used in the IRI: Letter Name Fluency, Letter Sound Fluency, and R-CBM (oral reading fluency).

Letter Name Fluency

Letter name fluency is measured by asking the student to name as many letters as possible on a page in one minute. Letter name fluency is a useful measure of kindergarten and early first grade reading skills until students reach the ceiling when they master letter recognition. In a study of 945 kindergarten students, letter name and letter sound knowledge were found to be uniquely predictive of subsequent reading achievement (Schatschneider et al., 2004). Measures of letter name and letter sound knowledge were not timed in the Schatschneider et al. (2004) study.

Stage, Sheppard, Davidson, and Browning (2001) demonstrated that kindergarten letter name and letter sound fluency were good predictors of first grade students' performance on oral reading fluency CBMs given in first grade. Kindergarten letter name fluency uniquely predicted first grade reading growth. The results of the hierarchical multiple regression analyses correlated kindergarten letter name fluency and first grade oral reading fluency at .628. The authors advocate for the use of letter name fluency measures as an early screen for kindergarten students.

Elliott, Lee, and Tollefson (2001) studied letter naming fluency, letter sound fluency, initial phoneme segmentation, and phonemic segmentation in 75 kindergarten students. Initial phoneme segmentation is the ability to identify the first sound that is heard in a word. Phoneme segmentation is the ability to break words into individual sounds. Letter name fluency was the best predictor of kindergarten achievement on the Broad Reading and Skills Cluster of the *Woodcock-Johnson Psych-Educational Battery-Revised* (WJ-R; Woodcock & Johnson, 2001). Letter name fluency correlated with Broad Reading with .63 and Skills Cluster with.75. Letter sound fluency correlated with Broad Reading with .58 and Skills Cluster with .72. The combination of letter sound fluency and letter name fluency was correlated with the subtests of WJ-R. The correlation of letter name fluency and letter sound fluency with Broad Reading was .64 and Skills Cluster was .79. The results of the analysis are limited because only two (letter name and letter sound) of the four measures studied were fluency measures (Burke, Hagan-Burke, Zou, & Kwok, 2010).

Hintze, Ryan, and Stoner (2003) compared letter name fluency, initial sound fluency, and phoneme segmentation fluency to *The Comprehensive Test of Phonological Processing* (CTOP; Wagner, Torgesen, & Rashotte, 1999). The initial sound fluency task was timed. Kindergarten students were asked to choose the picture that begins with the target sound from a display of four pictures. Phoneme segmentation was also timed. The students had one minute to say the stimulus word in segmented syllables. Letter name fluency and phonemic segmentation fluency correlated to the CTOP Phonologic Awareness Composite with a strong correlation of .53 for each fluency measure. Initial sound fluency correlated with the Phonologic Awareness Composite with .60. Initial sound fluency correlated to the Rapid Naming Composite with .20 and phonemic segmentation fluency correlated with .39. Letter name fluency had the strongest correlation to the Rapid Naming Composite, with a correlation of .58.

Speece, Mills, Ritchey, and Hillman (2003) determined that the combination of letter name fluency and nonsense word fluency measures taken in kindergarten appeared to be a moderate to strong predictor of oral reading fluency. Concurrent and predictive validity coefficients were calculated for letter name fluency and nonsense word fluency with a battery of language, reading-related, and reading assessments given to 39 children as kindergarten and first grade students. Letter name fluency and nonsense word fluency were strong measures most sensitive to identify poor readers in first grade. The findings are limited because of the small sample size (Burke et al., 2010).

Letter name fluency and nonsense word fluency were correlated to two subtests (Letter Word Identification and Word Attack) of the WJ-R and oral reading fluency (Speece et al, 2003). The concurrent validity coefficients for WJ-R Letter Word Identification were .77 for letter name fluency and .91 for nonsense word fluency, for students in kindergarten. The predictive validity coefficients of letter name fluency with first grade reading measures were .55 for Letter Word Identification of the WJ-R, .44 for Word Attack on the WJ-R, and .69 on oral reading fluency. The predictive validity coefficients of nonsense word fluency were .59 for Letter Word Identification of the WJ-R, .59 for Word Attack on the WJ-R, and .71 on oral reading fluency. The authors analyzed the data with the Pedhazur (1997) predictive approach to regression.

Letter name fluency and nonsense word fluency, taken together, contributed a range of 1.4% to 34% unique variance across the reading measures (Speece et al., 2003). Neither letter name fluency nor nonsense word fluency contributed a unique variance to the prediction of first grade WJ-R subtests. The authors speculated that the WJ-R is norm-referenced test designed to differentiate children within a wide developmental period. Oral reading fluency is an outcome measure. There were students in the study who scored above the 25th percentile on the WJ-R subtests, but did not perform well on oral reading fluency. The authors recommended that reading outcome measures such as oral reading fluency be included in a reading assessment system.

The kindergarten nonsense word fluency was the only significant predictor of kindergarten WJ-R Letter Word Identification (Speece et al., 2003). Nonsense word fluency also was a significant unique variance predictor of first grade WJ-R Word Attack, oral reading fluency, and nonsense word fluency. Kindergarten nonsense word fluency was not a unique variance predictor of first grade WJ-R Letter Word Identification. Letter name fluency did not contribute a unique variance to any of the regression models. The authors discovered that letter name fluency demonstrated reasonable criterion-related validity, however the construct validly for nonsense word fluency is stronger when the regression analysis is taken into account. The authors demonstrated that nonsense word fluency is a valid measure when taken in the spring of kindergarten. Prior to the study by Speece et al. (2003), it was assumed that nonsense word fluency was only appropriate for first grade students. The combination of kindergarten letter name fluency and nonsense word fluency was 87.5% sensitive to the identification of poor readers.

In a study of 330 students, Rouse and Fantuzzo (2006) examined the predictive validity of letter name fluency, phoneme segmentation fluency, and nonsense word fluency from the end of kindergarten to the end of first grade. The authors compared the measures of fluency with standardized reading assessments. Although letter name fluency, phoneme segmentation fluency, and nonsense word fluency had a significant relationship to the *Test of Early Reading Ability-Third Edition* (TERA-3; Reid, Hresko, & Hammill, 2001), letter name fluency had the strongest canonical correlation. The combination of letter name fluency, phoneme segmentation fluency, and nonsense word fluency explained 51.9% of the variance of the performance of first grade students on the *Developmental Reading Assessment* (DRA; Beaver, 1997). Letter name fluency was found to be the strongest predictor of DRA performance, with a standard regression coefficient of .45. Nonsense word fluency had a standard regression coefficient of .20 and phoneme segmentation fluency was at .17. Letter name fluency and phoneme segmentation fluency were better predictors than nonsense word fluency on the vocabulary, language, and reading subtests of the *TerraNova* (*TerraNova*, 1996). Letter name fluency had a standard regression coefficient of .26 on the Reading, .46 on the Vocabulary, and .36 on the Language subtests of the *TerraNova*. Phoneme segmentation fluency was the next strongest with standard regression coefficients of .24 on Reading, .17 on Vocabulary, and .25 on Language. The authors suggest that nonsense word fluency does not provide unique predictive information about later reading ability.

Clemens, Shapiro, and Thoemmes (2011) investigated word identification fluency, letter naming fluency, phoneme segmentation fluency, and nonsense word fluency administered to 138 first grade students. Word identification fluency is the ability to read high frequency words in one minute (D. Fuchs Fuchs, & Compton, 2004). The early reading CBMs were administered in the fall of first grade. At the end of first grade, the students reading abilities were assessed with oral reading fluency CBMs, Test of Word Reading Efficiency: Sight Word Efficiency and Phonemic Decoding Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999), a Maze CBM, and a latent variable composite. The latent variable composite is a composite of all the raw scores from the oral reading fluency, TOWRE, and the Maze. Except for phoneme segmentation, all of the fall screening measures had moderate to strong relationships to the spring outcome measures. Phoneme segmentation fluency had a weak to moderate relationship. Word identification fluency tended to demonstrate the strongest correlations to the end-of-the year assessments. Word identification fluency was a significant predictor of each outcome measure. When the screening CBMs were combined, models of two or three CBMs were as accurate as all of the screening CBM measures put together. Combining word identification fluency with phoneme segmentation fluency or nonsense word fluency reduced the number of false positives. Used together, word identification fluency and phoneme segmentation fluency *or* word identification fluency and nonsense word fluency made the strongest predictions.

Letter Sound Fluency

As mentioned earlier, the IRI uses letter name fluency and letter sound fluency to assess first grade reading. Letter sound fluency is measured by asking the student to say as many sounds of the letters on a page in one minute. There does not seem to be as much research on letter sound fluency as letter name fluency (Speece et al., 2003). Letter sound fluency was found to be an effective measure of early reading skills in kindergarten (Burke et al., 2009) and first grade students (Speece & Ritchey, 2005).

As discussed in the letter name fluency section earlier, letter name and letter sound fluency taken at the end of kindergarten made a strong prediction of first grade oral reading fluency (Stage et al., 2001). Letter name fluency appeared to be a strong predictor of growth in first grade oral reading fluency. Untimed measures of letter name knowledge and letter sound knowledge were also found to be strong predictors of future reading ability (Schatschneider et al., 2004).

Letter sound fluency was compared to phonological awareness and rapid automatized naming measurements in 276 first and second grade students (Speece & Ritchey, 2005). Letter sound fluency collected in January was the best predictor of first grade reading growth and fluency in May. There was not a strong correlation of letter sound fluency to second grade oral reading fluency. Phonological awareness and rapid automatized naming assessments were not found to be uniquely predictive for first grade reading ability. Rapid automatized naming was calculated by the speed in which students named objects.

Letter sound fluency and nonsense word fluency were investigated by Ritchey (2008). In this study, 91 kindergarten students were given letter sound fluency and nonsense word fluency assessments five times during the second semester of kindergarten. Both assessments correlated to concurrent reading ability and were found to be effective in predicting comparable future reading ability. There was variability in the students' responses to nonsense word fluency tasks. Ritchey (2008) noted that low scores on nonsense word fluency may not be a valid indicator of risk for reading difficulties for all kindergarten students. Kindergarten students who had low scores on nonsense word fluency and who demonstrated reading difficulties, also had low scores on letter sound fluency. Letter sound fluency was found to be a valid and simple means to assess letter sound knowledge and the potential risk for reading failure.

R-CBM (Oral Reading Fluency)

The R-CBM on the IRI is an oral reading fluency assessment. There is a long history of research on the efficacy of oral reading fluency. Oral reading fluency measures correlate with a variety of standardized reading assessments (Fuchs & Fuchs, 1999; Muyskens et al., 2009; Shinn et al., 1992). Oral reading fluency is a fast and easy assessment of reading ability (Fuchs & Fuchs, 1999). It is also a good means of monitoring student progress. Oral reading fluency CBMs may be first given during the second semester of first grade (L. S. Fuchs et al., 2004).

Oral reading fluency has been studied for its predictive strength in first and second grade students' reading abilities (Speece & Ritchey, 2005). The authors concluded that growth in first grade oral reading fluency was the strongest single predictor of second grade performance and growth at the end of the school year. The January oral reading fluency rate was also a significant predictor of reading performance and growth in first grade students at the end of the year. Oral reading fluency was a good measure of reading proficiency in third and fifth grade students (Shinn et al.,1992). Oral reading fluency is found to be valid and reliable predictor of performance on statewide achievement tests (Buck & Torgesen, 2003; Buck, Torgesen, & Schatschneider, 2009; Carlisle, Schilling, Scott, & Zeng, 2004; Good et al., 2001; Hunley, Davies, & Miller, 2013; Riedel, 2007; Shapiro, Solari, & Petscher, 2008; Sibley, Biwer, & Hesch, 2001; Stage & Jacobsen, 2001; Stewart, 2009; Yeo, 2010). Oral reading fluency measurements given in English to fifth grade culturally and linguistically diverse students (CLDS) were found to be good predictors of later reading performance on statewide achievement tests (Muyskens et al., 2009).

Oral reading fluency is also a useful indicator of reading comprehension (Kim, Wagner, & Foster, 2011; Muyskens et al., 2009; Roehrig et al., 2008; Shinn et al., 1992; Spear-Swerling, 2006). Oral reading fluency was a better predictor of reading comprehension than phoneme segmentation by correctly classifying 80% of first grade students at the end of the year and 71% of the students at the end of their second grade in a study of 1,518 children (Riedel, 2007). Furthermore, adding other measures including a story retell task designed to measure comprehension did not significantly enhance the predictive power of oral reading fluency. Oral reading fluency measurements correlated highly with reading comprehension in fifth grade students (Shinn et al., 1992). Oral reading fluency assessments have high correlations to reading comprehension scores on high stake assessments (Buck & Torgesen, 2003; Shapiro et al., 2008).

Criticisms of CBMs

Generic CBMs (e.g., AIMSweb) are developed from data sources other than the school's curriculum (Deno, 2003). One of the major criticisms of generic CBMs is that the measures may not match the curriculum of the consumer school system (Fuchs & Deno, 1994; Powell-Smith & Bradley-Klug, 2001). An advantage of teacher created measurements is the face validity. If teachers create their own curriculum-based measures (CBMs), teachers may be more likely to use CBMs as part of their assessment system (Powell-Smith & Bradley-Klug, 2001).

Despite the face validity of CBMs, there are disadvantages to reading CBMs created directly from the curriculum (Powell-Smith & Bradley-Klug, 2001). First, there may be a larger measurement error because of the wide variety of variability in reading passages. Secondly, passages taken from the curriculum might be familiar to the student. A student might perform well on a familiar passage, thus making it unclear if the student applied knowledge to a novel reading experience. Finally, a student may falsely demonstrate mastery on reading passages with highly controlled vocabulary (Fuchs & Deno, 1994; Powell-Smith & Bradley-Klug, 2001). Passages derived from social studies or science texts are more difficult to control for vocabulary. The students may not perform as well on reading passages from content areas. Content areas generally include science, social studies, history, and math (Brummitt-Yale, 2014). In other words, content areas are most subjects except English literature. Fuchs and Deno (1994) concluded that curriculum sampling is not the most important aspect of instructionally useful measurement (Powell-Smith & Bradley-Klug, 2001). The essential assessment features are (a) opportunities for repeated measures, (b) the ability to determine outcomes of instruction, and (c) representations of student performance. Thanks to the volume of CBM research, it is possible to use generic and standardized CBMs, “uncoupled” from the local school curriculum (Deno, 2003). Generic CBMs may be useful in making instructional decisions. Another advantage is that generic CBMs may assess skill development in early literacy that may not be part of the curriculum (Kaminski & Good, 1996). Lastly, generic CBMs may come with computer based applications to facilitate record keeping (Fuchs, Fuchs, & Hamlett, 1993).

A Rational for a Possible Predictive Relationship between Low IRI Scores and Subsequent Diagnosis of SLD

Learning to read is one of the most important achievements of school aged children. Reading ability is a powerfully predictive factor in success in school and life. The “Matthew Effect” is the phenomena of the consequences of being “good” or “poor” in reading ability (Stanovich, 1986). Specifically, good readers experience success in reading and read more, thus propelling further academic and reading success (Cunningham & Stanovich, 2001; Taylor et al., 1990). Too often, poor readers do not experience success in reading, read less, and do not experience academic success. Poor readers are at-risk for school failure because it is difficult for them to advance their reading abilities enough to achieve academic success (Torgesen, 1998).

One reason why students experience difficultly with reading is because of SLD. It is critical that students with SLD be identified sooner, rather than later, so that they might benefit from specialized instruction. The identification of SLD is based on long-term and poor academic performance (Reschly, 1996). Three methods of identification of SLD have been in use over the years. The traditional approach or IQ-academic achievement discrepancy model was widely used until recently. A contemporary approach or RTI requires the implementation of scientific research-based interventions in reading over an eight to 12 week period. Students must demonstrate a resistance to the interventions by demonstrating a learning rate that is below expectations before a referral can be made for a special education evaluation. The traditional and contemporary approaches have supporters and detractors. Currently, RTI is not required but may be used under the reauthorization of the Individuals with Disabilities Education Improvement Act of 2004 (Berkeley et al., 2009; Bradley et al., 2007). The act allows states to use RTI, continue to use the traditional IQ-achievement discrepancy criterion, or use a blended approach. The blended approach was adopted in Idaho in 2010. The blended approach combines RTI with comprehensive assessments of psychological processes (e.g. memory, phonological skills, and processing speed) as a means to identify students with SLD (Baskette et al., 2006; Barth et al., 2008; Batsche et al., 2006; Bradley et al., 2007; Fletcher et al., 2004; Fuchs et al., 2010; Kavale & Spaulding, 2008; Keogh, 2005; McKenzie, 2009).

The IRI is the kindergarten through third grade assessment that the state of Idaho implemented to screen students for potential reading difficulties. The IRI is the basis for many important decisions that impact individual student lives. Performance on the IRI determines which students may receive additional reading interventions and school funding for professional development, technical assistance, curriculum, and other activities by the SDE (Idaho Comprehensive Literacy Act, 1999).

A few studies involving the IRI exist. All of the studies were commissioned by the SDE. Two studies by Nave and Burke (2007a, 2007b) compared two versions of the IRI in terms of proficiency ratios between groups of students. Stewart (2009) examined the how well student performance on the IRI predicted student performance on the ISATS. The IRI was able to explain 84% of the variance on student performance.

As noted earlier, the ability to identify students with SLD early in the school career is critical for students, educators, and schools. For students with significant reading issues, a diagnosis of SLD can give access to intense interventions not otherwise available. An unresolved question is the degree to which IRI scores can serve as an early indicator and predictor of subsequence diagnosis of SLD.

CHAPTER 3: METHOD

Introduction

This study included eight research questions that examine the correlation between Idaho Reading Indicator (IRI) scores and the subsequent identification of Specific Learning Disabilities (SLD). Three types of statistical analyses were used in this study. First, a logistic regression analysis was used to determine the relationship between IRI scores and the identification of SLD at any point between the years of 2009 to 2012. Secondly, a Pearson chi-square was used to further examine phenomenon that might exist in the data. A value of less than 0.05 would reflect a significant relationship between IRI scores and the subsequent identification of SLD. A value of greater than 0.05 would not reflect a significant relationship between IRI scores and the subsequent identification of SLD. The Cramer's V statistic was used to determine the magnitude of the relationship. The expected count and actual count chi-square statistic was used to measure the gap between the expected count (the number if there was no relationship between IRI scores and the subsequent identification of SLD) and the actual count for each cell in the 2 X 3 table. Finally, odds ratios were calculated to demonstrate the relationship between IRI scores and the subsequent identification of SLD.

Research Questions and Hypotheses

This study targeted eight questions. This study examined the relationship between the IRI and the subsequent identification of SLD students by the end of the first semester in fourth grade. The eight questions that this study sought to resolve were as follows:

Question 1: What is the correlation between Letter Sound Fluency scores by students on the first grade fall IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

H0: There will not be a statistically significant correlation between first grade fall Letter Sound Fluency scores and the subsequent identification of these students as SLD by the end of the first semester in fourth grade.

H1: There will be a statistically significant correlation between the first grade fall Letter Sound Fluency scores and the subsequent identification of these students as SLD by the end the first semester in fourth grade.

Question 2: What is the correlation between Letter Sound Fluency scores by students on the first grade spring IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

H0: There will not be a statistically significant correlation between first grade spring Letter Sound Fluency scores and the subsequent identification of these students as SLD by the end of the first semester in fourth grade.

H1: There will be a statistically significant correlation between the first grade spring Letter Sound Fluency scores and the subsequent identification of these students as SLD by the end of the first semester in fourth grade.

Question 3: What is the correlation between R-CBM scores by students on the first grade fall IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

H0: There will not be a statistically significant correlation between first grade fall R-CBM on the IRI and the subsequent identification of these students as SLD by the end of the first semester in fourth grade.

H1: There will be a correlation between first grade fall R-CBM on the IRI and the subsequent identification of these students as SLD by the end of the first semester in fourth grade.

Question 4: What is the correlation between R-CBM scores by students on the first grade spring IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

H0: There will not be a statistically significant correlation between first grade spring R-CBM on the IRI and the subsequent identification of these students as SLD by the end of the first semester in fourth grade.

H1: There will be a statistically significant correlation between first grade spring R-CBM on the IRI and the subsequent identification of these students as SLD by the end of the first semester in fourth grade.

Question 5: What is the correlation between R-CBM scores by students on the second grade fall IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

H0: There will not be a statistically significant correlation between second grade fall R-CBM on the IRI and the subsequent identification of these students as SLD by the end of the first semester in fourth grade.

H1: There will be a correlation between second grade fall R-CBM on the IRI and the subsequent identification of these students as SLD by the end of the first semester in fourth grade.

Question 6: What is the correlation between R-CBM scores by students on the second grade spring IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

H0: There will not be a statistically significant correlation between second grade spring R-CBM on the IRI and the subsequent identification of these students as SLD by the end of the first semester in fourth grade.

H1: There will be a statistically significant correlation between second grade spring R-CBM on the IRI and the subsequent identification of these students as SLD by the end of the first semester in fourth grade.

Question 7: What is the correlation between R-CBM scores by students on the third grade fall IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

H0: There will not be a statistically significant correlation between third grade fall R-CBM on the IRI and the subsequent identification of these students as SLD by the end of the first semester in fourth grade.

H1: There will be a statistically significant correlation between third grade fall R-CBM on the IRI and the subsequent identification of these students as SLD by the end of the first semester in fourth grade.

Question 8: What is the correlation between R-CBM scores by students on the third grade spring IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

H0: There will not be a statistically significant correlation between third grade spring R-CBM on the IRI and the subsequent identification of these students as SLD by the end of the first semester in fourth grade.

H1: There will be a statistically significant correlation between third grade spring R-CBM on the IRI and the subsequent identification of these students as SLD by the end of the first semester in fourth grade.

Design of Study

Participants

The study focused on the performance of students on the IRI and the subsequent identification of SLD within this population of students by fourth grade. Treatments were not administered to the students examined in this study. There was no need to recruit participants for this study, as all data used already existed. These data included (a) performance data on the IRI from fall and spring administrations in first through third grade and (b) the subsequent identification of students with or without SLD by the end of the first semester in fourth grade.

The study concluded at the first semester of fourth grade for several reasons. As reported earlier, 75% of students with SLD are identified by the third and fourth grades (Lyon, 1996). A majority of the students with SLD were likely to have been identified by the end of the first semester in fourth grade. The study followed the introduction of the current IRI in the fall of 2009 through the spring of 2012 and the first semester of fourth grade, a span of three years and one semester. The study ended at fourth grade because of the time restraints imposed by the author.

In 2009, the current AIMSweb (AIMSweb Pro Complete, 2010) version of the IRI was given to all Idaho students in grades kindergarten through third grade (Idaho State Department of Education, 2013). Fourth grade student data was selected because they were first grade students and among the first group to take the 2009 AIMSweb version of the IRI. In 2013, the group of first grade students entered the fourth grade. In the fall of 2013, there were 2078 fourth grade students enrolled in the BSD (S. Tyree, personal communication, October 2, 2014).

The IRI data from students who attended BSD public schools in first grade in 2009 and fourth grade in 2013 was compared to the subsequent identification of SLD within this population. Students who did not attend first through fourth grades in the BSD were eliminated from the study, with attrition statistics reported (more detailed information about the available scores is presented in Appendix A).

All data used came from the BSD, a data pool to which the author had access. The BSD was selected because of the large size of the data pool and the availability of the Infinite Campus student information system. Student identification numbers, performance on the IRI, and special education designation are maintained in the Infinite Campus system. The current 2009 AIMSweb version of the IRI was implemented in the BSD from the fall of 2009 to present (Idaho State Department of Education, 2014b).

Setting

The BSD is located in Boise, Idaho. Boise is the capital and the largest city in Idaho. Boise is in southwestern Idaho on the Boise River. The BSD is described as an urban school district (Boise School District, 2015). The most recent population data from 2012 estimates the population of Boise as 212,303 (United States Census Bureau, 2015). The BSD is the second largest school district in the state.

Demographics

The most recent BSD demographic data was from 2011. In 2011, the BSD had an enrollment of 25,430 students, including the special education preschool population and general education grades kindergarten through twelve (Boise School District, 2015). The number of students enrolled in the preschool program, and kindergarten through grade sixth was 13,720.

A description of the BSD demographics in 2011 (Boise School District Statistics, 2015) is presented on Table 3.1. The demographic information applies to BSD students from preschool and kindergarten through 12th grade, with the exception of the percentage of students identified as enrolled in the free and reduced lunch program. The percentage of students enrolled in the free and reduced lunch program in 2011 applies to students in first through sixth grade only. The number of students enrolled in the free and reduced lunch program was determined by the number of students receiving free and reduced lunches. The percentage of students in the BSD identified as having Limited English

Proficiency (LEP) was calculated by dividing the number of BSD students labeled as LEP (2087) by the total number of BSD students (25,430).

Table 3.1

Boise School District Student Characteristics in 2011

Characteristic	Percentage
Free/reduced lunch	47.2
White ethnicity	79.3
Hispanic ethnicity	10.3
Asian ethnicity	5.9
Black ethnicity	3.3
Native American ethnicity	0.9
Pacific Islander	0.8
LEP	9.0

Note. LEP = students with limited English proficiency.

BSD Student Population of 2009

Student demographic characteristics from 2009 are of interest, since this study was concerned with the performance of first grade students on the IRI during that year. In 2009, the total enrollment number was 25,097. The population of students enrolled in preschool through grade sixth was 13,539.

A description of the BSD demographics in 2009 (Boise School District Statistics, 2015) is presented on Table 3.2. All demographic information applies to BSD students from preschool and kindergarten through 12th grade, with the exception of the percentage of students identified as receiving free and reduced lunch. The percentage of students receiving free and reduced lunch in 2009 applies to the percentage of first through sixth grade students receiving free and reduced lunches. The percentage of students in the BSD identified as LEP was calculated by dividing the number BSD students labeled as LEP (2255) by the total number of BSD students (25,097).

Table 3.2

Boise School District Student Characteristics in 2009

Characteristic	Percentage
Free/reduced lunch	42.9
White ethnicity	82.0
Hispanic ethnicity	9.1
Asian ethnicity	4.1
Black ethnicity	3.2
Native American ethnicity	0.7
Pacific Islander	0.7
LEP	8.9

Note. LEP = students with limited English proficiency.

Recruitment Plan

There was no need to recruit participants for this study. The BSD district level administrators were contacted to obtain permission to conduct the study. The BSD uses Infinite Campus student information system. Each student is given an identification number as part of the state and district level student tracking. The identification numbers of first, second, third, and fourth grade students were used to match raw IRI scores to students identified with SLD. The identification numbers that did not pair with the IRI or SLD identification were discarded. Students were coded as “SLD” and “not SLD” by using district level information recorded in Infinite Campus. The presence of SLD was connected with the student identification numbers. Personal and identifying information (name, date of birth, school) that could potentially reveal the identity of any of the students was not be collected. The IRI scores from the years 2009 - 2013 and the BSD documentation of students identified as "SLD" on the December 1st count of 2013 was studied.

Variables to be Measured

In this study, the correlations between IRI scores and the subsequent identification of these students as SLD or not SLD by the fourth grade were compared. The IRI scores and the identification of students as SLD or not SLD were requested from the BSD. For purposes of this study, a BSD administrator removed official student identification numbers and reassigned pseudo-identification numbers prior to the data being released to the author.

IRI Scores

The performance of students on the first, second, and third grade IRI was the first variable in this study. The IRI is administered to students in Idaho, grades kindergarten through third. In the BSD, the IRI is administered three times a year, in the fall, winter and spring ("Assessment: What do assessments measure?", n.d.). In kindergarten, the IRI assesses Letter Naming Fluency and Letter Sound Fluency. In first grade, the IRI assesses Letter Sound Fluency and oral reading fluency (R-CBM). In second and third grades, the IRI contains R-CBM assessments (Idaho State Department of Education, 2013).

The first grade IRI was included in this study because it contains Letter Sound Fluency and R-CBM subtests. While there is a great deal of research available on R-CBMs, there is less research concerning Letter Sound Fluency (Speece et al., 2003). The kindergarten IRI was not included in this study. The kindergarten IRI utilizes Letter Name Fluency and Letter Sound Fluency. Although Letter Name Fluency is a useful early reading measure, it has a ceiling effect, once young students master the letter names. The study followed the first grade students because it allowed time for the potential identification of SLD. In 2012, this group of students entered the fourth grade.

A three point cut score system is used in the IRI. Cut scores are used to rank student performance. The points are used to classify student performance on the IRI as “benchmark,” “strategic,” or “intensive” (Idaho State Department of Education, 2013). On the IRI, a score of “3” is benchmark and defined as reading skills at or above grade level. A student score of “2” is called “strategic.” A student score of “1” is called “intensive.”

The cut scores for the fall and spring first grade IRI on Letter Sound Fluency and R-CBM were presented in relationship to the raw scores of each subtest. Letter Sound Fluency has more utility in the fall, because first grade students are not expected to be fluent readers as measured on the R-CBM. There is a smaller range of data points on the first grade fall R-CBM.

The fall and spring cut scores are presented by grade level in the next section. The winter cut scores were not available, therefore they were not included. It was required to administer the IRI three times a year in 2009 – 2010. The BSD student performance on the fall, winter, and spring IRIs were included within each grade level.

First Grade IRI Scores

Letter Sound Fluency on the first grade fall IRI was the significant measure and was used to determine reading proficiency levels. In the fall, the Letter Sound Fluency is used to determine proficiency. Both Letter Sound Fluency and R-CBM are used to measure student growth rates in proficiency, when the scores are compared to the spring scores. In the fall, a "benchmark" score is 31 or above on Letter Sound Fluency. A "strategic" score is 20-30 on Letter Sound Fluency. An "intensive" score is 0-19 on Letter Sound Fluency.

In the spring, the R-CBM becomes more important and is used to determine proficiency levels by the SDE. A "benchmark" score is 53 or above on R-CBM. A "strategic" score is 28 - 52 on the R-CBM. An "intensive" score is 0 – 27 on the R-CBM. Table 3.3 summarizes the cut scores in relation to raw scores on the first grade IRI.

Table 3.3

First Grade IRI Cut Scores and Raw Scores

IRI Cut Score Rating	Letter Sound Fluency	R-CBM
Fall		
Intensive	0 - 19 = 1	0 - 0 = 1
Strategic	20 - 30 = 2	1 - 1 = 2
Benchmark	31 - above = 3	2 - above = 3
Spring		
Intensive	0 - 51 = 1	0 - 27 = 1
Strategic	52 - 71 = 2	28 - 52 = 2
Benchmark	72 – above = 3	53 - above = 3

Note. Benchmark = 3; Strategic = 2; Intensive = 1.

A summary of the BSD 2009-2010 first grade performance is presented on Table 3.4 (Idaho State Department of Education, 2012). In 2009, the Idaho State Department of Education (SDE) required school districts to administer the IRI and report the results three times a year.

Table 3.4

Boise School District 2009-2010 First Grade IRI Scores

Score	Fall Percentage	Winter Percentage	Spring Percentage
Benchmark	68.97	79.09	79.16
Strategic	19.62	16.62	13.11
Intensive	11.41	4.28	7.16

Second Grade IRI Scores

The R-CBM is the only measure used to determine reading proficiency on the second grade IRI. The R-CBM data is collected and used to measure student growth rates in proficiency, when the fall R-CBM score is compared to the R-CBM in the spring.

In the fall, a “benchmark” is 54 or above on the R-CBM. A “strategic” score is 27 - 53 on the R-CBM. An “intensive” score is 0 - 26 on R-CBM. In the spring, the “benchmark” score is 92 or above. A “strategic” score is 68 - 91 on the R-CBM. An “intensive” score is 0 – 67 on R-CBM. Table 3.5 summarizes the cut scores in relation to raw scores on the second grade IRI.

Table 3.5

Second Grade IRI Cut Scores and Raw Scores

Fall Cut Scores	Spring Cut Scores
0 - 26 = 1	0 - 67 = 1
27 - 53= 2	68 - 91= 2
54 - above = 3	92 - above = 3

Note. Benchmark = 3; Strategic = 2; Intensive = 1.

A summary of the BSD 2010 - 2011 second grade proficiency scores is presented on Table 3.6 (Idaho State Department of Education, 2012). In 2010, the Idaho SDE was required to administer the IRI and report the results two times a year. The winter administration of the IRI was optional.

Table 3.6

Boise School District 2010-2011 Second Grade IRI Scores

Score	Fall Percentage	Winter Percentage	Spring Percentage
Benchmark	58.88	70.48	77.21
Strategic	22.9	15.46	11.68
Intensive	18.22	14.06	11.10

<u>Third Grade IRI Scores</u>

Like the second grade IRI, the R-CBM is the only measure used to determine reading proficiency on the third grade IRI. Student growth in reading proficiency in third grade is determined by comparing the fall R-CBM to the spring R-CBM on the IRI. In the fall, a “benchmark” score is 77 or above on the R-CBM. A “strategic” score is 49 –

76 on the R-CBM. An “intensive” score is 48 or less on R-CBM. In the spring, the “benchmark” score is 110 or above. The proficiency level is considered “strategic” with a score of 82 – 109 on the R-CBM. An “intensive” score is 81 or less on the R-CBM. Table 3.7 summarizes the cut scores in relation to raw scores on the third grade IRI.

Table 3.7

Third Grade IRI Cut Scores and Converted Scores

Fall Cut Scores	Spring Cut Scores
0 - 48 = 1	0 - 81 = 1
49 - 76 = 2	82 - 109 = 2
77 - above = 3	110 - above = 3

Note. Benchmark = 3; Strategic = 2; Intensive = 1.

A summary of the BSD 2011-2012 third grade proficiency rates is presented on Table 3.8 (Idaho State Department of Education, 2012). In 2010, the Idaho SDE required school districts to administer the IRI and report the results two times a year. The administration of the winter IRI was optional.

Table 3.8

Boise School District 2011-2012 Third Grade IRI Scores

Score	Fall Percentage	Winter Percentage	Spring Percentage
Benchmark	69.88	77.01	81.69
Strategic	18.20	12.23	9.50
Intensive	11.93	10.75	8.81

Subsequent Diagnosis of SLD

The second variable in the study was the subsequent identification of students as SLD by the end of the first semester in fourth grade. BSD documentation of students with SLD who were in first grade in 2009, second grade in 2010, third grade in 2011, and fourth grade in the fall of 2012 was requested. For the purposes of this study, a BSD administrator removed official student identification numbers and reassigned pseudo-identification numbers to match the pseudo-identification numbers of students identified as SLD. The pseudo-identification numbers were used to make the statistical comparisons. The pseudo-identification numbers were coded as SLD or not SLD. The fall and spring IRIs for the students who entered first grade in 2009 were collected over the course of three years, as these students attended first, second, and third grades.

The raw scores of Letter Sound Fluency on the fall and spring first grade IRI were correlated to the subsequent identification of SLD in these students as they progress through first, second, third, and fourth grades. The raw scores of the fall and spring R-CBM on the IRI for grades first through third were correlated to the subsequent identification of SLD in these students as they advanced through grades one through four. The winter IRI scores were not included in this study because the requirement to administer the IRI in the winter was eliminated by the SDE in 2010.

It is necessary to note that at the time of this study the approach to identifying students with SLD changed. Until August of 2010, the traditional IQ-achievement discrepancy model was the way in which students were identified as having SLD in Idaho. At the start of the 2010 school year, the new blended approach was implemented. The traditional IQ-achievement discrepancy approach is no longer accepted as a means to label students with SLD. The data from students who entered first grade in 2009 is the focus of this study. Although there is a slight possibility that a very few of the students in first grade are labeled as SLD under the discrepancy model, these numbers are likely insignificant. From the fall of 2010 through the present, the blended approach is the method used to identify students as SLD.

Techniques for Gathering and Managing Data

The BSD administration was asked for the first grade student identification numbers paired with their IRI scores for the fall of 2009 and spring of 2010, for second grade students in the fall of 2010 and the spring of 2011, and for third grade students in the fall of 2011 and the spring of 2012. In addition, the BSD was asked to provide the pseudo-identification numbers of students identified as SLD in first grade in 2009, second grade in 2010, third grade in 2011, and fourth grade in 2012.

A BSD administrator removed official student identification numbers and reassigned pseudo-identification numbers. The pseudo-identification numbers were used to make the statistical comparisons.

A coding system was created for the identification of with and without SLD for the student data used in this study. Students identified with SLD were coded as “one” and students without SLD were coded as “zero.” Student identification numbers, IRI scores, students identified as SLD (one) and not SLD (zero) were entered into the Statistical Package for the Social Sciences (SPSS) 20.0.0 Student Version spread sheet columns.

Techniques for Analyzing, Interpreting, and Presenting Data

A non-experimental, cross-sectional design was used in this study (Horn, Plazas, Coverdale, Louie, & Weiss Roberts, 2009). The purpose of this study was to determine the relationships between IRI scores at repeated points and subsequent identification of students as SLD. The strength of the relationships between IRI scores and subsequent identification of SLD was examined. Five types of statistical analysis were used to accomplish this task, logistic regression analysis, Pearson chi-square test, Cramer's V statistic, expected count and actual count, and odds ratios. The Statistical Package for the Social Sciences (SPSS) version 21 was the computer software system used in this study. The SPSS is able to conduct the logistic regression, Pearson chi-square test, Cramer's V statistic, expected count and actual count, and odds ratio data analysis.

A logistic regression analysis was used to determine the relationship between IRI raw scores and the identification of SLD at any point between the years of 2009 to 2012. A logistical regression model was used because a binary predicted variable of SLD or not SLD was paired with a continuous predictor variable, the raw IRI scores. It was expected that there would be a strong relationship between poor student performance on the IRI and the identification of SLD. It was expected that there would be a weak relationship between student performance of "proficient" on the IRI and the subsequent identification of SLD.

A Pearson chi-square test was used to further examine any relationships that might exist in the data. On the Pearson chi-square test, a value of less than 0.05 reflects a significant relationship between IRI scores and the subsequent identification of SLD. A value of greater than 0.05 does not reflect a significant relationship between IRI scores and the subsequent identification of SLD. The Cramer's V Statistic was used to determine the magnitude of the relationships between the variables. The expected count and actual count measured the gap between the expected count (the number if there was no relationship between IRI scores and the subsequent identification of SLD) and the actual count for each cell in the 2 X 3 table.

The last statistical analysis employed was an odds ratio to demonstrate the relationship between IRI scores and the subsequent identification of SLD. The odds ratios determined how likely IRI scores linked with the identification of SLD. Conversely, the odds ratios demonstrated how students with SLD performed on the IRI.

Limitations of the Study

Several limitations were associated with this study. The limitations were connected with sample size, the four-year time frame of this study, the change in Idaho State Department of Education requirements for the identification of students with SLD, and the BSD documentation of SLD in the database.

First, the sample size may have been too small to make generalizations. The sample may not have been representative of primary elementary students at the national level. The student population of Boise may not be diverse enough to mirror the national population. Furthermore, Boise is considered an urban school district, thus making generalizations to rural, small town, and metropolitan communities potentially questionable.

Another potential problem with the study was that it encompassed a four year and one semester time frame. A longer period might have allowed more time for the subsequent identification of SLD. That is, a student has the potential of being identified as SLD at any point during an elementary and secondary educational career. Although most students are typically identified as SLD around ages 8 or 9 (Lyon, 1996), it is possible that some of the students in this study were identified as SLD after the first semester of fourth grade, but would not have been identified in this analysis due to the imposition of an SLD diagnosis by the end of the first semester in fourth grade.

At the time that this study was conducted, a dramatic shift was made in the way students with SLD were labeled. Idaho eliminated the IQ discrepancy criteria for SLD and moved to a blended model. In the spring of 2010, the IQ discrepancy model was in use. In the fall of 2010, all Idaho school districts were required to use the new eligibility criteria to identify students as SLD. There was an overlap between the IQ discrepancy model and the blended approach utilization in Idaho within the time frame of this study. The impact of the different eligibility criteria is likely to be minimal, as first grade students are not typically identified as SLD under the IQ discrepancy model (Stage et al., 2003). Most of the students with SLD who were in first grade in the spring of 2010 were more likely to be identified with the new blended model of eligibility.

Finally, the way SLD was documented by the BSD limited the generalization of the findings of this study. Not all students identified as SLD have difficulty with reading achievement. A SLD might be present in eight different academic areas: (a) oral expression, (b) listening comprehension, (c) written expression, (d) basic reading skills, (e) reading fluency, (f) reading comprehension, (g) mathematic calculation, and (h) mathematics problem solving (Idaho State Department of Education, 2009). A student identified with SLD may display significant difficultly in any one or combination of these eight academic areas. The BSD records the students as SLD but does not identify the academic areas. The overwhelming majority of students labeled as SLD have reading difficulties (Lyon, 1996). However, it is possible to have SLD and be proficient in reading but not in another area, such as math.

CHAPTER 4: RESULTS

The purpose of this study was to determine if there is a relationship between Idaho Reading Indicator (IRI) scores and the subsequent identification of specific learning disabilities (SLD). Existing IRI scores, curriculum-based measurement scores, and special education records from the Boise School District (BSD) were used in this study. The IRI scores were from students who received the first AIMSweb version of the IRI in 2009. IRI data was collected twice a year from the same group of students through 2012. The IRI is a statewide reading rate assessment that is administered annually in Idaho to students in kindergarten through third grade.

In this chapter, a review of the types of data will be provided. The results of the study in relationship to each hypothesis will be reported. An analysis of the fourth grade AIMSweb reading rate (R-CBM) is also included. (The fourth grade R-CBM is not part of the state required IRI assessment system, but the information was provided by the BSD.) A summary of the mean raw scores for the fall and spring IRI at each grade level is reported. The mean raw scores of the fourth grade R-CBM are also included.

The Data

IRI raw scores for grades one through three were used in this study to determine the relationship between IRI performance and the subsequent identification of SLD. The fall and spring IRI scores for grades one, two, and three were correlated with the dichotomous data of student diagnosis of SLD (with SLD or not with SLD) over the four-year period from the fall of 2009 to the spring of 2013.

The IRI is scored with a three-point rating system: “1,” “2,” or “3” (Idaho State Department of Education, 2013). A student score of “3” indicates “mastery of the skills.” A score of “3” is considered “proficient.” A student score of “2” indicates “partial mastery of some or all skills.” A student score of “1” indicates “a lack of mastery of some or all skills.” Cut scores are set for each grade level. Cut scores are used to classify student performance (Zieky & Perie, 2006). Cut scores are selected points on a test scale. The correlation between students identified as SLD and not SLD with IRI was examined with the chi-square test of association.

The raw IRI scores from 2009 through 2012 were collected for the Boise School District (BSD). The scores were from students who were in first grade in 2009, in second grade in 2010, and in third grade in 2011. Additionally, the list of students identified with SLD from the years 2009 through 2013 was provided by the BSD (with all personal identification information removed).

Some missing data occurred across the grade levels. The percentage of IRI scores available for analysis range from approximately 97% for the first grade fall IRI to 75% for the third grade spring IRI (more detailed information about the available scores is presented in Appendix A). The data was analyzed using the Statistical Package for the Social Sciences (SPSS) version 21.

Hypotheses and the Data Analysis

Eight research questions were the focus of this study. Each hypothesis targeted the connection between IRI scores and the subsequent identification of SLD. This section will report the results in response to each hypothesis.

Statistical Analyses

Several statistical tests were used to examine the data in this study. The Pearson chi-square statistic was used to determine if there was a significant relationship between IRI scores and the subsequent diagnosis of SLD. The Cramer's V statistic was used to determine the magnitude of the relationship between the IRI scores and the subsequent identification of SLD. Odds ratios were calculated to provide an additional metric for the strength of the association between IRI scores and the likelihood of being identified as SLD. To calculate the odds ratios, IRI scores of "1" and "2" were combined because the scores are defined as "below proficient."

Question 1. What is the correlation between first grade fall Letter Sound Fluency IRI scores and subsequent identification of SLD by the end of the first semester in fourth grade?

Pearson chi-square statistic. The Pearson chi-square statistic was used to determine if there was a relationship between first grade fall Letter Sound Fluency IRI scores and subsequent identification of SLD. The Pearson chi-square statistic was statistically significant (Pearson chi-square= 79.3; df= 2; $p < .001$).

Cramer's V statistic. The Cramer's V statistic was used to determine the magnitude of the relationship between first grade fall Letter Sound Fluency IRI scores and the subsequent identification of SLD. The Cramer's V statistic was significant but weak to moderate in magnitude (Cramer's V value = .204; $p < .001$).

Expected count and actual count. The chi-square statistic was used to measure the difference between the expected count (the value if there was no relationship between IRI scores and the subsequent identification of SLD) and the actual count for each cell in the 2 X 3 table. In the first grade, the chi-square model predicted 6.2 students as level "1," 10.9 as level "2," and 11 as level "3" on the 2009 fall Letter Sound Fluency IRI. The actual number of students who scored "1," which indicated "a lack of mastery of some or all skills," was 23. The actual number of students with SLD who fell within level 1 was almost four times greater than what was expected. The model predicted that 10.9 students with SLD would score a "2" on the IRI. The actual count was 22, nearly double what was predicted. Finally, 38.9 students were expected to obtain a "3" on the Letter Sound Fluency IRI but the actual count was 11. The actual count of first grade students who were considered "proficient" was three times less than the expected count. Expected versus actual counts are displayed on Table 4.1.

Table 4.1

2009 First Grade Fall IRI Chi-Square Count and Expected Count Statistics

Total Students	SLD	IRI Score of 1	IRI Score of 2	IRI Score of 3
1856	56			
Count		23.0	22.0	11.0
Expected		6.2	10.9	38.9

The odds ratio. An odds ratio was used to determine the likelihood of students with SLD and without SLD scoring below proficient (“1” or “2”) on the IRI. IRI scores of “1” and “2” were combined to calculate the odds ratios because the scores are defined as “below proficient.” On the fall first grade Letter Sound Fluency IRI, students with SLD were 10 times more likely to score a “1” or “2” on the IRI when compared to students without SLD. The odds ratios for grades one through three on the fall and spring IRI are in Appendix E.

Summary of the first grade fall Letter Sound Fluency IRI data. A statistically significant relationship between first grade fall Letter Sound Fluency IRI scores and the subsequent identification of SLD was found, but the relationship between the two variables was found to be weak to moderate in magnitude. More students with SLD scored below proficient on the first grade fall Letter Sound Fluency IRI than expected. Students with SLD were 10 times more likely to score below proficient on the fall first grade Letter Sound Fluency IRI than students without SLD.

Question 2. What is the correlation between first grade spring Letter Sound Fluency IRI scores and subsequent identification of SLD by the end of the first semester in fourth grade?

Unfortunately, given that spring Letter Sound Fluency data was not available for analysis, the relationship between the spring Letter Sound Fluency and the subsequent identification of SLD could not be determined. In the spring, Letter Sound Fluency is not used to determine spring proficiency scores. Instead, the R-CBM score is used to determine spring proficiency levels.

Question 3. What is the correlation between first grade fall R-CBM IRI scores and subsequent identification of SLD by the end of the first semester in fourth grade?

As with the case of the spring Letter Sound Fluency, given that fall R-CBM data was not available for analysis, the relationship between the fall R-CBM and the subsequent identification of SLD could not be determined. In the fall, the R-CBM is administered only as a baseline measure to determine student growth on the spring R-CBM. Thus, in the fall, the score that was used to determine proficiency levels of first grade students is Letter Sound Fluency.

Question 4. What is the correlation between first grade spring R-CBM IRI scores and subsequent identification of SLD by the end of the first semester in fourth grade?

Pearson chi-square statistic. The Pearson chi-square statistic was used to determine if there was a relationship between first grade spring R-CBM IRI scores and subsequent identification of SLD. The Pearson chi-square statistic was statistically significant (Pearson chi-square= 222.1; df= 2; p< .001).

***Cramer's V statistic*.** The Cramer's V statistic was used to determine the magnitude of the relationship between first grade spring R-CBM IRI scores and the subsequent identification of SLD. The Cramer's V statistic was significant but weak to moderate in magnitude (Cramer's V value = .340; p< .001).

Expected count and actual count. The chi-square statistic was used to measure the difference between the expected count (the value if there was no relationship between IRI scores and the subsequent identification of SLD) and the actual count for each cell in the 2 X 3 table. On the spring 2010 R-CBM IRI for first grade, the chi-square test of association predicted that 4.7 students with SLD would score at level “1.” The actual count was 33, seven times more than predicted. It was expected that 7.5 students with SLD would score “2” on the IRI, however the actual number of students with SLD was nearly double at 18. It was expected that 46.7 students with SLD would score “3” on the IRI and the actual count was eight. The actual count of eight was six times lower than the predicted amount. Expected versus actual counts are displayed on Table 4.2.

Table 4.2

2010 First Grade Spring IRI R-CBM Chi-Square Count and Expected Count Statistics

Total Students	SLD	IRI Score of 1	IRI Score of 2	IRI Score of 3
1859	59			
Count		33.0	18.0	8.0
Expected		4.7	7.5	46.7

The odds ratio. An odds ratio was used to determine the likelihood of students with SLD and without SLD scoring below proficient (“1” or “2”) on the IRI. IRI scores of “1” and “2” were combined to calculate the odds ratios because the scores are defined as “below proficient.” The odds of first grade students with SLD scoring a “1” or “2” on the spring R-CBM IRI was nearly 28 times more likely than a student without SLD. The odds ratios for grades one through three on the fall and spring IRI are in Appendix E.

Summary of the first grade spring R-CBM IRI data. A statistically significant relationship between first grade spring R-CBM IRI scores and the subsequent identification of SLD was found, but the relationship between the two variables was found to be weak to moderate in magnitude. More students with SLD scored below proficient on the first grade spring R-CBM than expected. Students with SLD were almost 28 times more likely to score below proficient on the first grade spring R-CBM IRI than students without SLD.

Question 5. What is the correlation between second grade fall R-CBM IRI scores and subsequent identification of SLD by the end of the first semester in fourth grade?

Pearson chi-square statistic. The Pearson chi-square statistic was used to determine if there was a relationship between second grade fall R-CBM IRI scores and subsequent identification of SLD. The Pearson chi-square statistic was statistically significant between the second grade fall R-CBM IRI scores and the subsequent identification of SLD (Pearson chi-square = 145.8; df = 2; $p < .001$).

Cramer's V statistic. The Cramer's V statistic was used to determine the magnitude of the relationship between second grade fall R-CBM IRI scores and the subsequent identification of SLD. The Cramer's V statistic was significant but weak to moderate in magnitude (Cramer's V value = .286; $p < .001$).

Expected count and actual count. The chi-square statistic was used to measure the difference between the expected count (the value if there was no relationship between IRI scores and the subsequent identification of SLD) and the actual count for each cell in the 2 X 3 table. When the students were in second grade in the fall of 2010, the expected count for students with SLD who scored a "1" on the IRI was 8.5 but the actual count was 39, nearly five times more than predicted in the chi-square test of association. It was expected that 12.1 students with SLD would score a "2" on the IRI. The actual count was 14. Two students with SLD obtained a score "3" on the IRI but 34.5 was the expected count. The actual count of two was nearly 18 times less than the expected count predicted by the model. Expected versus actual counts are displayed on Table 4.3.

Table 4.3

2010 Second Grade Fall IRI R-CBM Chi-Square Count and Expected Count Statistics

Total Students	SLD	IRI Score of 1	IRI Score of 2	IRI Score of 3
1721	55			
Count		39.0	14.0	2.0
Expected		8.5	12.1	34.5

The odds ratio. An odds ratio was used to further explore the performance of students with SLD and without SLD on the second grade fall R-CBM IRI of 2010. IRI scores of “1” and “2” were combined to calculate the odds ratios because the scores are defined as “below proficient.” Students with SLD were 48.2 times more likely to obtain a below proficient score (1 or 2) on the second grade fall R-CBM IRI than students without SLD. The odds ratios for grades one through three on the fall and spring IRI are in Appendix E.

Summary of the second grade fall R-CBM IRI data. A statistically significant relationship between the second grade spring R-CBM IRI scores and the subsequent identification of SLD was found, but the relationship between the two variables was found to be weak to moderate in magnitude. More students with SLD scored below proficient on the second grade spring R-CBM IRI than expected. Students with SLD were 48.2 times more likely to score below proficient on the second grade fall R-CBM IRI than students without SLD.

Question 6. What is the correlation between second grade spring R-CBM IRI scores and subsequent identification of SLD by the end of the first semester in fourth grade?

Pearson chi-square statistic. The Pearson chi-square statistic was used to determine if there was a relationship between second grade spring R-CBM IRI scores and subsequent identification of SLD. The Pearson chi-square statistic was statistically significant (Pearson chi-square = 196.7; df = 2; $p < .001$).

Cramer's V statistic. The Cramer's V statistic was used to determine the magnitude of the relationship between second grade spring R-CBM IRI scores and the subsequent identification of SLD. The Cramer's V statistic was significant but weak to moderate in magnitude (Cramer's V value = .340; $p < .001$).

Expected count and actual count. The chi-square statistic was used to measure the difference between the expected count (the value if there was no relationship between IRI scores and the subsequent identification of SLD) and the actual count for each cell in the 2 X 3 table. In the spring of 2011, 4.3 of the second grade students with SLD were expected to score "1" on the IRI. The actual count was 29, about seven times greater than predicted. Five students with SLD scored a "2" on the IRI but the expected count was 16. The actual count was about three times more than expected. Eight students with SLD scored a "3" on the IRI, however the expected count was 43.3. The actual count was approximately five times less than the expected count. Expected versus actual counts are displayed on Table 4.4.

Table 4.4

2011 Second Grade Spring IRI R-CBM Chi-Square Count and Expected Count Statistics

Total Students	SLD	IRI Score of 1	IRI Score of 2	IRI Score of 3
1647	53			
Count		29.0	16.0	8.0
Expected		4.3	5.3	43.3

The odds ratio. An odds ratio was used to determine the likelihood of students with SLD and without SLD scoring below proficient ("1" or "2") on the IRI. IRI scores of "1" and "2" were combined to calculate the odds ratios because the scores are defined as "below proficient." Students with SLD were nearly 30 times more likely to score below proficient when compared to students without SLD. The odds ratios for grades one through three on the fall and spring IRI are in Appendix E.

Summary of the second grade spring R-CBM IRI data. A statistically significant relationship between second grade spring R-CBM IRI scores and the subsequent identification of SLD was found, but the relationship between the two variables was found to be weak to moderate in magnitude. More students with SLD scored below proficient on the second grade spring R-CBM IRI than expected. Students with SLD were nearly 30 times more likely to score below proficient on the second grade spring R-CBM IRI, when compared to students without SLD.

Question 7. What is the correlation between third grade fall R-CBM IRI scores and subsequent identification of SLD by the end of the first semester in fourth grade?

Pearson chi-square statistic. The Pearson chi-square statistic was used to determine if there was a relationship between third grade fall R-CBM IRI scores and subsequent identification of SLD. The Pearson chi-square statistic was statistically significant between the third grade fall R-CBM IRI scores and the subsequent identification of SLD (Pearson chi-square = 182.0; df = 2; p< .001).

Cramer's V statistic. The Cramer's V statistic was used to determine the magnitude of the relationship between third grade fall R-CBM IRI scores and the subsequent identification of SLD The Cramer's V statistic was significant but weak to moderate in magnitude (Cramer's V value = .338; p< .001).

Expected count and actual count. The chi-square statistic was used to measure the difference between the expected count (the value of no relationship between IRI scores and the subsequent identification of SLD) and the actual count for each cell in the 2 X 3 table. Twenty-eight students with SLD obtained a score of "1" on the IRI but 4.2 was the expected count. There were nearly seven times more students in the actual count than the expected count predicted by the model. Twenty students with SLD obtained a score of "2" but 9.6 was the expected number. The actual number of students with SLD who scored a "2" on the IRI was double the expected number predicted by the model. Five students with SLD obtained a score of "3" on the IRI but 39.2 were expected. There were nearly eight times as many students predicted to score a "3" on the IRI as the actual number of five. Expected versus actual counts are displayed on Table 4.5.

Table 4.5

2011 Third Grade Fall IRI R-CBM Chi-Square Count and Expected Count Statistics

Total Students	SLD	IRI Score of 1	IRI Score of 2	IRI Score of 3
1537	53			
Count		28.0	20.0	5.0
Expected		4.2	9.6	39.2

The odds ratio. An odds ratio was used to determine the likelihood of students with SLD and without SLD scoring below proficient ("1" or "2") on the IRI. IRI scores of "1" and "2" were combined to calculate the odds ratios because the scores are defined as "below proficient." Students with SLD were 31 times more likely to score below proficient than students without SLD on the third grade fall IRI in 2011. The odds ratios for grades one through three on the fall and spring IRI are in Appendix E.

Summary of the third grade fall R-CBM IRI data. A statistically significant relationship between the third grade fall R-CBM IRI scores and the subsequent identification of SLD was found, but the relationship between the two variables was found to be weak to moderate in magnitude. More students with SLD scored below proficient on the fall third grade R-CBM IRI than expected. Students with SLD were nearly 31 times more likely to score below proficient, when compared to students without SLD on the third grade fall R-CBM IRI.

Question 8. What is the correlation between R-CBM scores by students on the third grade spring R-CBM IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

Pearson chi-square statistic. The Pearson chi-square statistic was used to determine if there was a relationship between third grade spring R-CBM IRI scores and subsequent identification of SLD. The Pearson chi-square statistic was statistically significant between the third grade R-CBM IRI scores and the subsequent identification of SLD (Pearson chi-square = 198.9; df = 2; p< .001).

Cramer's V statistic. The Cramer's V statistic was used to determine the magnitude of the relationship between third grade R-CBM IRI scores and the subsequent identification of SLD. The Cramer's V statistic was significant but weak to moderate in magnitude (Cramer's V value = .360; p< .001).

Expected count and actual count. The chi-square statistic was used to measure the difference between the expected count (the value of no relationship between IRI scores and the subsequent identification of SLD) and the actual count for each cell in the 2 X 3 table. In the spring of 2012, 2.6 students with SLD were expected to score a "1" on the IRI. Twenty-one was the actual count, about eight times more than predicted. Eighteen students scored a "2" on the IRI but 4.9 was the number predicted. The actual count was nearly four times more than the expected count of the chi-square test of association. Ten students with SLD scored a "3" on the IRI but 41.6 were predicted. There were less than four times as many students predicted to score a "3" on the IRI than the actual number of students. Expected versus actual counts are displayed on Table 4.6.

Table 4.6

2012 Third Grade Spring IRI R-CBM Chi-Square Count and Expected Count Statistics

Total Students	SLD	IRI Score of 1	IRI Score of 2	IRI Score of 3
1484	49			
Count		21.0	18.0	10.0
Expected		2.6	4.9	41.6

The odds ratio. An odds ratio was used to determine the likelihood of students with SLD and without SLD scoring below proficient ("1" or "2") on the IRI. IRI scores of "1" and "2" were combined to calculate the odds ratios because the scores are defined as "below proficient." A student with SLD was 26 times more likely to score below proficient than students without SLD on the third grade spring R-CBM IRI. The odds ratios for grades one through three on the fall and spring IRI are in Appendix E.

Summary of the second grade fall R-CBM IRI data. A statistically significant relationship between the third grade spring R-CBM IRI scores and the subsequent identification of SLD was found, but the relationship between the two variables was found to be weak to moderate in magnitude. More students with SLD scored below proficient on the spring third grade R-CBM IRI than expected. Students with SLD were nearly 26 times more likely to score below proficient on the third grade spring R-CBM IRI, when compared to students without SLD.

2012 Fourth Grade R-CBM Scores

Although the fourth grade R-CBM was not a part of the IRI and was not the initial focus of the hypotheses, data from this source was also analyzed. The BSD has used R-

CBMs at all elementary grade levels for decades. Thus, knowing the relationship between fourth grade R-CBM and the IRI for grades one through three might give insight into the continuity of the assessments. It is also important to know if the relationship between fourth grade R-CBM scores and subsequent identification of SLD is consistent across the IRI and BSD R-CBM boundaries. The information may be of value to the BSD. The R-CBM data for the fourth grade students was included with the requested IRI data.

The BSD administers R-CBMs to students in grades four through six. The R-CBM is an AIMSweb product. The R-CBMs for grades four through six are not a state but a district requirement. The administration of the R-CBM is the same as the R-CBMs in the IRI. Three reading passages at grade level reading ability are given to students in grades four, five, and six. The students are asked to read the grade level reading passages in one minute. The students reading fluency is measured by the number of words read per minute, minus the erred words read. The median score of the three reading passages is recorded. The scoring system is identical to the IRI with students cut scores ranked as "1," "2," or "3" (B. Anderson, personal communication, November 28, 2014). Probes were selected and reserved as universal screens from the AIMSweb progress monitoring system that the BSD purchased in 2009. The norms used to determine the reading rates were obtained from the AIMSweb system.

A raw score of 0 to 74 is a "1" or "intensive" on the fall R-CBM for fourth grade students. A score of 75 to 100 is a "2" or "strategic" on the fall R-CBM. A score of 101 or more words per minute read is considered a "3" or "benchmark." In the spring, fourth grade students who score a zero to 101 are considered at level "1" or "intensive." A score of 102 to 128 is a “2” or “strategic” on the fall R-CBM. A score of 129 or more words per minute read is considered a “3” or “benchmark.”

Statistical Analyses

The same statistical tests that were used to examine IRI scores and the subsequent identification of SLD were applied to the fourth grade fall and spring R-CBM data. First, the Pearson chi-square statistic was used to determine if there was a significant relationship between fourth grade fall and spring R-CBM scores and subsequent diagnosis of SLD. The Cramer’s V statistic was used to test the magnitude of the relationship between the R-CBM scores and the subsequent identification of SLD. Odds ratios were calculated to further explore the association of fourth grade fall and spring R-CBM scores and the likelihood of being identified as SLD. To calculate the odds ratios, IRI scores of “1” and “2” were combined because the scores are defined as “below proficient.” The percentage of participants on the fourth grade fall and spring R-CBM was slightly less than 70%. Information about the percentage of scores used and missing from the data appear in Appendix A.

Pearson Chi-Square Statistic

The Pearson chi-square statistic was used to determine if there was a relationship between fourth grade fall and spring R-CBM scores and subsequent identification of SLD. In the fall, the Pearson chi-square statistic was statistically significant between the fourth grade fall R-CBM and the subsequent identification of SLE (Pearson chi-square = 166.8; df = 2; $p < .001$). In the spring, the Pearson chi-square statistic was also significant between the fourth grade spring R-CBM scores and the identification of SLD (Pearson chi-square = 163.5; df = 2; p< .001).

Cramer's V Statistic

The Cramer's V statistic was used to determine the magnitude of the relationship between the fourth grade fall and spring R-CBM and the subsequent identification of SLD. The relationship between the fall fourth grade R-CBM and the subsequent identification of SLD was significant but weak to moderate in magnitude (Cramer's V value = .344; p< .001). In the spring, the magnitude between the fourth grade R-CBM and subsequent identification of SLD was also significant but weak to moderate in power (Cramer's V value = .337; p< .001).

Expected Count and Actual Count

The chi-square statistic was used to explore the differences between expected counts (the frequency expected when there is no relationship between R-CBM scores and subsequent identification of SLD) and the actual counts for each cell in the 2 X 3 table. Table 4.13 shows the expected and actual counts.

In the fall, the chi-square measure of association predicted that there would be approximately four students with SLD who obtained a “1” on the R-CBM, if the relationship was the result of chance. The actual count was 27, about seven times more than expected. The expected count of fourth grade students with SLD who obtained a “2” was 8.6 but 13 students actually received a “2” on the R-CBM. Approximately 33 students were expected to obtain a “3” on the R-CBM. Only five students with SLD actually obtained a proficient score, nearly seven times less than the expected number.

In the spring, 3.4 students with SLD were expected to obtain a “1” on the R-CBM. The actual number was 24, about eight times more than predicted. It was expected that 8.2 students with SLD would receive a “2” on the R-CBM but 17 was the actual number. While 33.4 was the expected number of students with SLD to score “3” on the R-CBM, only four students actually obtained a score of proficient. The expected number of students scoring in the proficient range was eight times greater than the actual number. Expected versus actual counts are displayed on Table 4.7.

Table 4.7

2012 – 2013 Fourth Grade R-CBM Chi-Square Count and Expected Count Statistics

Total Students	SLD	IRI Score of 1	IRI Score of 2	IRI Score of 3
2012 Fall				
1368	45			
Count		27.0	13.0	5.0
Expected		3.6	8.6	32.5
2013 Spring				
1396	45			
Count		24.0	17.0	4.0
Expected		3.4	8.2	33.4

The Odds Ratio

An odds ratio was used to determine the likelihood of students with SLD and without SLD scoring below proficient (“1” or “2”) on the R-CBM. R-CBM scores of “1” and “2” were combined to calculate the odds ratios because the scores are defined as “below proficient.” Students with SLD were nearly 23 times more likely to score below proficient than students without SLD on the fourth grade fall R-CBM in 2012. In the spring, students with SLD were likely to obtain a score of below proficient at a rate 33 times greater than students without SLD on the fourth grade R-CBM. The odds ratios for the fourth grade fall and spring R-CBM are in Appendix E.

Summary of the Fourth Grade Fall and Spring R-CBM Data

Statistically significant relationships for both the fall and spring fourth grade R-CBM and the subsequent identification of SLD were found. The relationship between the fall fourth grade R-CBM and the subsequent identification of SLD was found to be weak to moderate in magnitude. The relationship between the spring fourth grade R-CBM and subsequent identification of SLD was also found to be weak to moderate in magnitude. More students with SLD scored below proficient on the fall and spring fourth grade R-CBM than expected as predicted. Students with SLD were nearly 23 times more likely to score below proficient, when compared to students without SLD on the fall fourth grade R-CBM. In the spring, students with SLD were 33 times more likely to score below proficient on the fourth grade R-CBM.

Additional Findings

In addition to the above, two additional questions emerged and were examined. The first question looked at each grade-level gap between the mean raw scores of students without SLD and with SLD. The second question examined the annual drop in the mean raw scores for all students (both SLD and not SLD) from the spring to the fall of the succeeding school year. This phenomenon has been characterized as summer reading loss (Cooper, Nye, Charlton, Lindsay, and Greathouse, 1996).

The Mean Raw Score Gaps Between Students with and Without SLD

Table 4.8 and Figure 4.1 display the mean raw scores of the fall and spring IRI and the fourth grade R-CBM for students without SLD and those with SLD from the fall of 2009 to the spring of 2013.

Table 4.8

IRI Mean Raw Scores for Fall and Spring from 2009 through 2013

	Students without SLD		Students with SLD	
Assessment	Fall	Spring	Fall	Spring
2009 First Grade IRI	24.80	92.43	2.44	31.17
2010 Second Grade IRI	79.22	134.18	20.68	64.22
2011 Third Grade IRI	113.16	156.04	49.63	85.88
2012 Fourth Grade R-CBM	127.45	157.93	67.34	96.71

Note. The first grade IRI measures Letter Sound Fluency in the fall. In the spring, the first grade IRI is a R-CBM.

A visual comparison between the fall and spring mean raw scores for students without and with SLD is displayed in Figure 4.1. The bottom of the vertical line is the mean score for the fall IRI and fourth grade R-CBM scores. The top of the vertical line is the mean score for the spring IRI and fourth grade R-CBM scores. The graph displays the mean raw scores for students without SLD and for students with SLD. Both groups demonstrated growth in their ability to read words per minute; however, the mean raw scores for students with SLD remain lower when compared to students without SLD for each year represented.

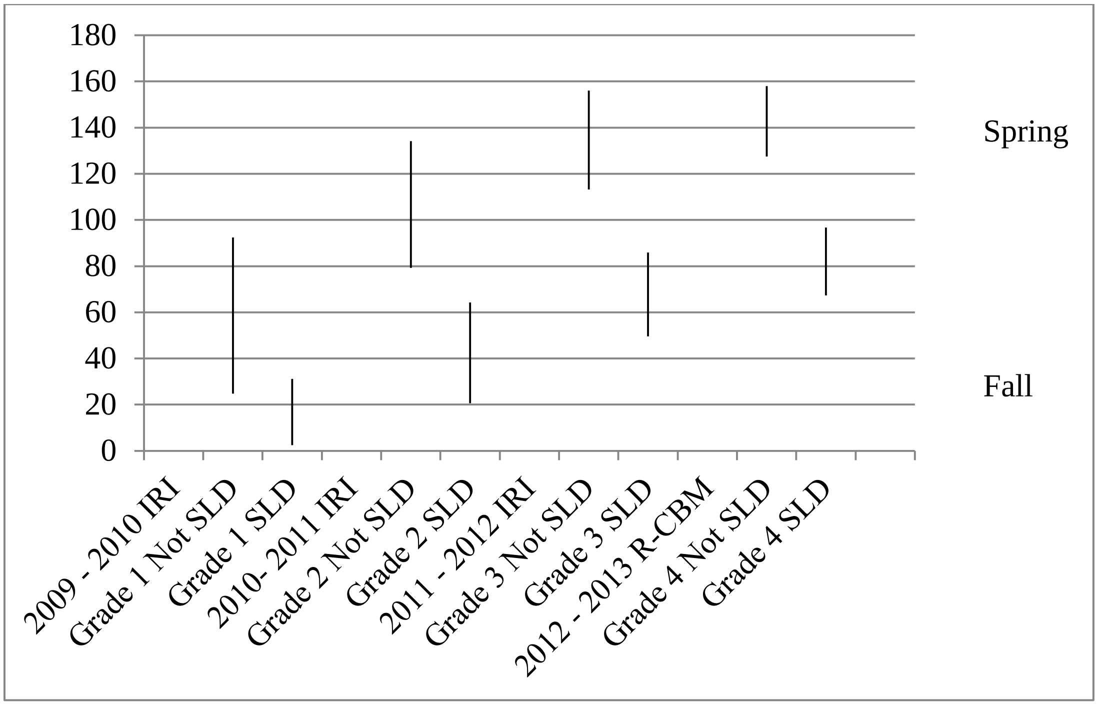

Figure 4.1. Fall and Spring Mean Raw Scores on the IRI and Fourth Grade R-CBM

Figure 4.1. Fall and spring mean raw scores of students with SLD and without SLD on the IRI and fourth grade R-CBM for 2009-2013. The IRI is designed for grades 1 through 3. In grade 4, a district R-CBM is given. The *bottom* of each vertical data line is the mean average raw score for the fall administration of the IRI. The *top* of each vertical data line is the mean average raw score for the spring administration of the IRI.

It should be noted that the fall first grade IRI was a Letter Sound Fluency measurement. The first grade spring IRI was a R-CBM measurement. In first grade, there is some overlap between the mean raw scores of the two groups of students (SLD and non-SLD) taking the IRI, with some students diagnosed with SLD obtaining scores in the *spring* administration that were higher than scores of some students without SLD in the *fall* administration of the instrument.

The second and third grade IRI consisted of R-CBMs. In second grade (and thereafter), no such overlap between the scores of the two groups existed. Instead, a gap between the fall mean raw scores on the IRI for students without SLD and subsequent spring mean raw scores for students with SLD emerges. This means that students without SLD are reading better (as measured by the IRI) in the fall of second grade than their peers subsequently diagnosed with SLD are reading by the *end* of second grade.

This gap continues after second grade. While students with SLD do demonstrate consistent annual growth in IRI scores, their mean raw scores remain consistently and substantially behind those of their grade peers without SLD. The gap widens for these students in second grade, with this pattern maintained through fourth grade.

In first grade, the fall mean score for students with SLD was 2.44 and the spring mean score was 31.17 (Table 4.8). For first grade students without SLD, the fall mean score was 24.8 and the spring score was 92.43. In second grade, the fall mean score was 20.68 and the spring mean score was 64.22 for students with SLD. The second grade fall mean score was 79.22 and the spring mean score was 134.18 for students without SLD.

This pattern continued in third grade. In third grade, the fall mean score was 49.63 and the spring mean score was 85.88 for students with SLD. In the third grade, the fall mean score was 113.16 and the spring mean score was 156.04 for students without SLD. In fourth grade, students in the BSD were given the R-CBM, which is similar to the IRI but not required by the state. The spring fourth grade R-CBM mean score was 67.34 and the spring mean score was 96.71 for students with SLD. For students without SLD, the spring fourth grade R-CBM mean score was 127.45 and the spring mean score was 157.93.

The Drop in Mean Raw Scores over Summer Vacation

One interesting observation is the consistent drop in mean raw scores from the spring to the subsequent fall (over summer vacation) in both students without SLD and students with SLD. A substantial drop in mean raw scores was observed annually for both students with and without SLD. (Note: The reading probes change for each grade level and are written at the readability level for that grade level. For example, the second grade IRI reading passages administered in the spring of second grade were written at a second grade readability level. The third grade IRI passages subsequently administered in the fall of third grade were written at a third grade readability level.)

The mean score for students without SLD on the spring first grade R-CBM IRI was 92.43. In the fall, the mean score for the same students without SLD on the second grade R-CBM IRI declined to 79.22. The 13-point difference represented a 14% drop in the mean score for students without SLD. This pattern of regression is also consistent for students with SLD. The mean score for students with SLD on the spring first grade IRI is 31.17. In the fall, the mean score for the same students with SLD on the fall IRI in second grade declined to 20.68. The 10-point gap represented a drop of nearly 34% in the mean score for students with SLD.

In the second grade, the mean score for students without SLD was 134.18 on the spring IRI. In the fall, the mean score for the same group of students without SLD declined to 113.16 on the third grade IRI. The 21-point difference represented a drop of about 16%. In the second grade, the mean score for students with SLD was 64.22 on the spring IRI. In the fall, the mean score for the same group of students with SLD declined to 49.63 on the third grade IRI. The approximate 15-point difference represented a drop of about 23% in mean raw scores for students with SLD.

In third grade, the spring mean score for students without SLD was 156.04 on the IRI. In the fall, the mean score for the same group of students without SLD declined to127.45 on the fourth grade R-CBM. This represented a difference of nearly 29 points and an 18% drop in mean raw scores for student without SLD. In third grade, the spring mean score was 85.88 on the IRI for students with SLD. In the fall, the mean score for the same group of students with SLD declined to 67.34 on the fourth grade R-CBM. This represented a difference of nearly 19 points and a 22% drop in mean raw scores for student with SLD. A summary of the drop in mean raw scores over summer vacation is presented on Table 4.9.

Table 4.9

The Drop in Mean Raw Scores between the Fall and Spring over Summer Vacation

Summer Loss Period	Designation	Spring Mean Score	Fall Mean Score	Difference	Percent Drop
1st Spring IRI to 2nd Fall IRI	Without SLD	92.43	79.22	13.21	14.3
	With SLD	31.17	20.68	10.49	33.7
2nd Spring IRI to 3rd Fall IRI	Without SLD	134.18	113.16	21.03	15.7
	With SLD	64.22	49.63	14.59	22.7
3rd Spring IRI to 4th Fall R-CBM	Without SLD	156.04	127.45	28.59	18.3
	With SLD	85.88	67.34	18.54	21.5

CHAPTER 5: DISCUSSION

Introduction

The primary purpose of this study was to examine the relationship between the Idaho Reading Indicator (IRI) scores and the subsequent identification of students with specific learning disabilities (SLD). More specifically, the study was designed to respond to the following eight questions:

1. What is the correlation between Letter Sound Fluency scores by students on the first grade fall IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

2. What is the correlation between Letter Sound Fluency scores by students on the first grade spring IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

3. What is the correlation between first grade fall R-CBM IRI scores and subsequent identification of SLD by the end of the first semester in fourth grade?

4. What is the correlation between first grade spring R-CBM IRI scores and subsequent identification of SLD by the end of the first semester in fourth grade?

5. What is the correlation between second grade fall R-CBM IRI scores and subsequent identification of SLD by the end of the first semester in fourth grade?

6. What is the correlation between second grade spring R-CBM IRI scores and

subsequent identification of SLD by the end of the first semester in fourth grade?

7. What is the correlation between third grade fall R-CBM IRI scores and subsequent identification of SLD by the end of the first semester in fourth grade?

8. What is the correlation between R-CBM scores by students on the third grade spring R-CBM IRI and subsequent identification of SLD by the end of the first semester in fourth grade?

The variables used in this study were the IRI scores and the identification of students with SLD and those without SLD in the BSD from the years 2009 to 2013.

Summary of the Study

Existing IRI scores from students enrolled in the Boise School District (BSD) from the fall of 2009 through the spring of 2012 were used in this study. This group of students was enrolled in first grade in 2009. Their fall and spring IRI scores from the first, second, and third grades were used in the analysis. These six sets of IRI scores were correlated with subsequent identification of these students as either having or not having SLD by the end of the first semester in fourth grade. In addition, the two fourth grade Curriculum-Based Measurement (R-CBM) evaluations of reading (fall and spring) were included in the data provided by the BSD. This fourth grade data was similarly statistically analyzed.

Statistical Analyses and Results Summary

To best understand any potential relationships between (a) IRI scores and (b) the binary categories of students with SLD and students without SLD, statistical analyses included (a) the Pearson chi-square, (b) Cramer's V statistic, (c) expected outcomes and actual outcomes, and (d) odds ratios.

The same statistical analyses were conducted on the fourth grade R-CBM scores for the fall and spring of 2012-2013. The Pearson chi-square, Cramer's V statistic, expected outcomes and actual outcomes, and odds ratios were used to determine the relationship between fourth grade R-CBM scores and the subsequent identification of SLD.

Support for the Hypotheses

Weak to moderate correlations were found between all IRI scores (fall and spring in first, second, and third grades) and the subsequent identification of SLD. The statistical analyses were carried out as follows.

First, the Pearson chi-square statistic was used to determine if a relationship existed between IRI scores and the subsequent identification of SLD. A statistically significant relationship was found between each spring and fall IRI for grades one, two, and three and the subsequent identification of SLD. A significant relationship was also found between the fourth grade spring and fall R-CBM scores and the subsequent identification of SLD. While the Pearson chi-square statistic was able to demonstrate that a statistically significant relationship existed between the variables of IRI scores and the subsequent identification of SLD, it was unable to tell the magnitude of the relationship.

The Cramer's V statistic was used to find the magnitude of the relationship between IRI scores and the subsequent identification of SLD. The results of the Cramer's V statistic revealed a weak to moderate magnitude of the relationship between the IRI and the subsequent identification of SLD. Each fall and spring IRI for grades one, two, and three had a weak to moderate correlation with the subsequent identification of SLD. This pattern was also true for the fourth grade R-CBM and subsequent identification of SLD.

The expected count and actual count of the chi-square statistic further illustrated that the relationship between the IRI scores and the subsequent identification of SLD was not the result of chance. The actual count for students with SLD was higher than the expected count for each fall and spring IRI for grades one, two, and three, and for the fourth grade R-CBM. For example, in the third grade, the actual count of students with SLD who obtained a score of “2” on the IRI was 18; nearly four times the expected count of 5.

Odds ratios were used to demonstrate how much more likely a student with SLD would obtain a “below proficient” score on the IRI when compared to students without SLD. The lowest odds ratio was on the fall first grade IRI, indicating that students with SLD were 10 times more likely to score “below proficient” than students without SLD. The highest odds ratio was on the second grade fall IRI, revealing that students with SLD were 48 times more likely to score “below proficient” than students without SLD. The average likelihood of students with SLD obtaining a score of “below proficient” on the IRI was 29 (see Appendix E for complete odds ratio list).

<u>Consistency of the Gaps in IRI Scores Between Students Identified as SLD or Not SLD</u>

The comparison of the mean raw scores of the fall and winter IRI and fourth grade R-CBM scores yielded a finding not directly associated with the hypotheses but nevertheless interesting in its own right. Students with SLD consistently obtained lower mean raw scores on the IRI and R-CBM than did students without SLD. This result was found in fall and spring mean IRI scores in first, second, and third grades (that is, for every IRI data point) (Table 4.9, Figure 4.1).

Furthermore, although all students (both SLD and not SLD) showed growth in their reading ability throughout the years, their rankings relative to each other remained essentially unchanged. That is, once a student scores low on the IRI, it is likely that he/she will maintain that low ranking relative to grade peers. The converse is also true, as students who obtain initially high scores on the IRI maintain that high ranking through at least fourth grade.

The gap between the scores between students with SLD and students without SLD begins in first grade (Table 4.9, Figure 4.1). Interestingly, there was some overlap between the range of first grade mean raw scores for students with SLD and students without SLD in the two IRI administrations that year, with some students with SLD scoring higher in the spring administration of the test (after several months of reading instruction) than did their non-SLD peers earlier that academic year, in the fall administration. In considering this, it should be noted that the first grade IRI is a combination of Letter Sound Fluency and R-CBM. In the fall, Letter Sound Fluency is used to measure proficiency. In the spring the R-CBM is used to measure proficiency. The fall score that was reported in the BSD data was the Letter Sound Fluency measurement. The spring first grade IRI score that was recorded in the BSD data was the R-CBM. The second and third grade IRI assessments are R-CBM. Thus, the two assessments in first grade potentially were measuring slightly different literacy skills.

After first grade, any overlaps between IRI scores of students with and without SLD disappear. No overlap of mean raw scores existed between second, third, and fourth grades between students without SLD and with SLD. Instead, the gap widened for these students in second grade, with that pattern maintained through fourth grade (Table 4.9, Figure 4.1).

Summer Vacation Drop in Mean Raw Scores

Another interesting finding was the consistent drop in mean raw scores from the spring of one year to the fall in the next year. This phenomenon was noted on both the IRI and the fourth grade R-CBM (Table 4.9, Figure 4.1), and was observed for both students with and without SLD. In the first grade, for students without SLD, the percentage drop on IRI scores was 14% from the first grade spring to the second grade fall IRI. For first grade students with SLD, the magnitude of the drop over summer vacation was larger. These students recorded an IRI score drop of nearly 34% from the first grade spring administration of the IRI to the second grade fall administration.

For students without SLD, there was a percentage drop of approximately 16% on IRI scores from the second grade spring administration to the third grade fall administration. For students with SLD, there was a percentage drop of approximately 23% on the second grade spring to the third grade fall IRI.

For students without SLD, there was a percentage drop of approximately 18% on IRI scores from the third grade spring administration to the fourth grade fall administration of the R-CBM. For students with SLD, there was a percentage drop of approximately 22% on the third grade spring to the fourth grade fall R-CBM.

Conclusions and Implications

The Hypotheses

Statistically significant relationships were consistently found between IRI scores and the subsequent identification of SLD. However, the magnitude of these relationships was weak to moderate, and thus may be of questionable practical utility. There are a number of possible reasons why the relationships between IRI scores and subsequent diagnoses of specific learning disabilities were not stronger.

One possible explanation for the results of this study emerges from the fundamental purpose and intent in the design of the IRI and R-CBM. The IRI was created as a statewide assessment to identify students below grade level in reading (Idaho Comprehensive Literacy Act, 1999). The IRI and other R-CBMs typically are used as universal screens to identify students who are at risk for reading failure (Hughes & Dexter, n.d. b; Jenkins, Hudson, & Johnson, 2007; Jenkins & Johnson, 2008). The IRI (Idaho State Department of Education, 2013) and R-CBMs (Fuchs & Fuchs, 2011) were not designed, and cannot reasonably be expected to function, as comprehensive reading diagnostic tools. More specifically to the purposes of the present study, the IRI was not created to identify students with SLD. Instead, it was designed more generically as a universal screen to identify students at-risk for reading failure.

There are many reasons that students may have difficulties in reading. In addition to the presence of SLD, other possible causal factors include (a) cultural and linguistic differences, (b) low intelligence, and (c) poor reading instruction (Committee on the Prevention of Reading Difficulties in Young Children, 1998; Fletcher et al., 1994; Foorman et al., 1998; Fuchs et al., 2002). Since the IRI is a universal screen designed to identify *all* students with potential reading difficulties, and was *not* developed to diagnose the *reason* for the reading problem, a stronger magnitude in the Cramer's V statistic between the IRI and the subsequent identification of SLD would be unexpected. The IRI's purpose is to cast a wide enough net to capture *all* students at-risk for reading failure, not just those students whose issues in reading are directly and specifically attributable to the presence of SLD.

In this study, consideration was given to the potential statistical impact on IRI scores of cultural and linguistic differences in the test takers (categorized as English Language Learners (ELL) in the data provided by the BSD). Of the over 2000 students whose data was analyzed in this study, there were only 104 ELL students. The number of such students was not significant compared to the whole group. It thus seems unlikely that the inclusion or exclusion of the ELL students would make any significant difference in the observed outcomes of the statistical analyses.

Similar consideration was given to the potential impact of low socio-economic status (SES) on IRI scores. Unfortunately, only incomplete data on SES of the students was provided by the BSD. There is the potential that if students with low SES were removed from the analysis, correlations might be somewhat stronger. However, given (a) the substantial overlap between learning disabilities and poverty (e.g., Coutinho, Oswald, & Best, 2002) and (b) the fact that the magnitude of the correlations were found to be only weak to moderate, it seems unlikely that the exclusion of low SES students (even if they had been able to be identified) would have generated significant practical implications.

A number of potential implications emerge from the results of this study. First, the IRI may be an excellent universal screen for reading issues in general. However, because there is a limited amount of available psychometric information specific to the development of the IRI, it is difficult to make a definitive determination as to the IRI's efficacy in this task (Santi & Francis, 2012).

As noted previously, the efficacy of the IRI cannot be fully determined because of the limited amount of psychometric information available to the public (Santi & Francis, 2012). The effectiveness of a universal screen requires the examination of the following elements: (a) sensitivity, (b) specificity, (c) practicality, and (d) consequential validity (Hughes & Dexter, n.d. b; Jenkins et al., 2007, Jenkins & Johnson, 2008).

Sensitivity refers to the degree to which a universal screen can accurately predict which students will perform *poorly* on a future measure (Hughes & Dexter, n.d. b; Jenkins et al., 2007, Jenkins & Johnson, 2008; Johnson, Jenkins, Petscher, & Catts, 2009). While the initial findings as outlined in this study are encouraging as to the sensitivity of the IRI for this purpose, more public information about the sensitivity of the IRI will be necessary to draw more definitive conclusions. Unfortunately, information about the sensitivity of the IRI was not made available for this study.

Specificity is the degree to which a universal screen can accurately identify students who perform *satisfactorily* at present and will later perform satisfactorily on a future measure. No universal screen is perfect. Thus, cut scores must be set at an optimal predictive level (Hughs & Dexter, n.d. b; Jenkins et al., 2007, Jenkins & Johnson, 2008; Johnson et al., 2009). In general, effective universal screens seek to identify 90% of the students who will perform poorly on later reading achievement measures (Jenkins & Johnson, 2008).

For the IRI, more public information allowing for more accurate conclusions about its level of specificity is necessary. When specificity rates are too high, there is an increase in the number of false positives. A large number of false positives, or students who did not make the cut scores but will perform satisfactorily in a later assessment, may cost school district resources that might be better used on students with true reading difficulties. Unfortunately, the information required to make more informed judgments about the specificity of the IRI is not available.

Only one study could be identified that sought to examine the ability of the third grade 2007 AIMSweb IRI to predict the performance of students on the high stakes academic achievement assessment: the Idaho Statewide Achievement Test (ISATS; Stewart, 2009). The ISATS assesses reading, language usage, and mathematics. Stewart (2009) found that the AIMSweb IRI correctly predicted student performance for all but 15.8% of the students. The strongest correlation between the IRI and the ISATS was found among white students who were not enrolled in special education or attending Title I schools. The R-CBM probes of the 2007 AIMSweb IRI were discontinued in 2009 and replaced with the current probes, which are now the intellectual property of the Idaho SDE.

Practicality refers to the logistical issues of administering and scoring a test. The IRI is quick and easy to administer, so the criteria of practicality is met.

Finally, *consequential validity* refers to the likelihood that the outcomes of any evaluative measure will incorrectly result in harm to students. Consequential validity of a universal screen is violated when one or more of the following occurs: (a) students are misidentified, (b) there is inequitable treatment, or (c) interventions based on the evaluation data are ineffective. At this point in time, it cannot be determined if the IRI meets the criteria of consequential validity because detailed psychometric information is not available to the public for examination. Information about the sensitivity and specificity of the IRI would be helpful to determine consequential validity.

It may be that the predictive functions of the IRI could be enhanced through the incorporation of multiple measures in a screening battery. Multiple screening measures have been found to be more accurate in classification than a single measure (Clemens et al., 2011; Jenkins & Johnson, 2008; Jenkins & O'Connor, 2002; Johnson et al., 2009). Predictive accuracy was also improved when teacher's ratings of student attention and behavior were added to the screening battery (Davis et al., 2007).

In terms of the original questions of this study, it may be that with the addition of select multiple measures (e.g., cognitive processing efficiency, intellectual ability), the IRI might be able to more accurately identify early on those students likely to subsequently be diagnosed with specific learning disabilities. However, to the degree that an assessment instrument is made more complex with the incorporation of additional measures, it loses some of the qualities that one looks for in an effective screening device (e.g., practicality, or ease of administration and scoring).

Other Findings

In addition to the major hypotheses studied in this work, a number of additional findings of interest emerged. These include (a) the consistency of the grade level gaps in IRI scores between students with and without SLD, (b) the possible Matthew Effect, (c) the magnitude of the drop in mean IRI scores over summer vacations, and (d) the implications of the use of the IRI as a teacher accountability measure.

Consistency of the Grade Level Gaps in IRI Scores

Both students with and without SLD demonstrated growth in reading rates from first through fourth grades. However, the relative rankings of these two groups of students remained largely unchanged. Students with SLD obtain lower scores than students without SLD on each administration of the IRI as well as on the fourth grade R-CBM.

Essentially, what can be concluded from this is that both groups of students (those diagnosed and not diagnosed with SLD) evidence growth in reading skills over time. However, despite their ongoing improvement, students identified as SLD never “catch up” to their non-SLD grade peers in reading skills. The challenge for the schools, then, is to assist such students, not to simply make progress in acquisition of reading skills year after year, but to make enough progress to catch up with their grade peers.

The Possible Matthew Effect

The consistent and growing difference in mean raw scores of students with SLD and students without SLD on the IRI and fourth grade R-CBM likely is an illustration of the so-called *Matthew Effect* (Stanovich, 1986). In the Matthew effect, in essence, the “poor get poorer” and the “rich get richer” in reading ability. That is, if a student starts out as a poor reader, he/she will likely remain a poor reader. However, if a student begins as a skilled and successful reader, he/she will likely remain so.

This is likely attributable to the fact those good readers are more likely to enjoy reading, and experience fewer reading frustrations, than are their poorer reading counterparts. Thus, good readers are likely to read more and more, in the process of becoming even better readers. Similarly, because of their difficulties, poor readers are less likely to enjoy reading, and thus are less likely to read as much. With less reading experience and success, they are then more likely to continue to struggle with reading (Cunningham & Stanovich, 2001; Taylor et al., 1990). It is difficult for poor readers to make enough gains in reading to become proficient in reading (Torgesen, 1998).

As noted earlier, Response to Intervention (RTI) has emerged as an efficacious intervention subsequent to the early identification of school issues, including reading difficulties. This study evaluated gains in IRI scores in a group of students in the Boise Public Schools from 2009 to 2013. In Idaho, RTI was not required until 2010, and likely was not widely implemented until sometime after that. The students examined in this study were enrolled in the BSD at a time when RTI was just beginning to be implemented in Idaho schools. The quality and fidelity of the early RTI interventions in the BSD were not examined in this study, but its implementation was likely to have been relatively nonsystematic and unsophisticated early on. It may well be that the early and comprehensive implementation of RTI for struggling readers (as is more likely to be the case at present) may help to minimize the Matthew Effect for these learners in the future.

Another consideration related to the Matthew Effect may be the inadvertent effect that the implementation and interpretation of the IRI had on the type of reading instruction subsequently provided to students. The IRI started out as a universal screen made to identify potential students at-risk for reading failure. However, in 2001, the IRI additionally became a tool for teacher accountability (Idaho Comprehensive Literacy Act, Idaho Code 33-1616). Accountability tests may pressure educators to shape curriculum and instruction to mirror what is being measured (Anagnostopoulos, 2003; Gordon Commission, 2012; McNeil, 2002; Paris, 2005; Pearson, 2006; Resnick & Berger, 2010; Samuels, 2007; Sparks, 2012).

It might be argued that the type of reading skill measured by the IRI and other related CBM systems is not in fact reading fluency, but instead may be more accurately described as a measure of *reading rate and error* (Pearson, 2006). The testing focus on increasing reading rate and words read correctly may have resulted in narrowed reading curriculum. That is, an overreliance on oral reading fluency may place more emphasis on "word calling" skills at the expense of developing higher levels of reading comprehension (Johnson et al., 2009; Paris, 2005; Pearson, 2006; Riedel, 2007; Samuels, 2007).

If indeed the implementation of the IRI has resulted in instruction that de-emphasizes reading comprehension, then for many readers comprehension issues may make reading less enjoyable and thus less likely to happen. Such a theoretical outcome may explain in part the probable Matthew Effect observed here.

<u>The Magnitude of the Drop in Mean IRI Raw Scores over Summer Vacations</u>

Another interesting and somewhat unexpected finding was the *magnitude* of the consistent drops in the mean raw scores between the spring and fall IRI administrations over summer vacation for all students, both with and without SLD. This observed outcome is consistent with the body of research targeting academic skill regression during summer vacation (Kim, 2006).

In a review of 39 studies concerned with achievement tests and summer vacation, Cooper et al. (1996) found that achievement test scores consistently declined over the summer. A meta-analysis of 13 of these studies found a loss of about one month on a grade-level equivalent scale, or one tenth of a standard deviation relative to spring test scores. Gaps in reading achievement are impacted by the socio-economic status (SES) of students, with reading gaps experienced by poorer students growing more over the summer than during the school year (Alexander, Entwisle, & Olson, 2001; Cooper et al., 1996; Downey, von Hippel, & Broh, 2004; McCoach, O'Connell, Reis, & Levitt, 2006). An equivalent relationship between drops in reading achievement and SES has been reported in the BSD. This is not insignificant in the present study, given the aforementioned strong relationship between low SES and SLD.

Another potential factor in the summer drop in achievement scores that should be noted is that the types and complexities of the reading passages incorporated in the IRI changes between grade levels. Each academic year, the passages are written at the readability level for that grade level. For example, students in the spring of the second grade taking the IRI that term read passages at a *second* grade readability level. When these students begin third grade in the following fall, the IRI passages they are to read were at a *third* grade readability level. If students had little or no opportunity to read over the summer (as is likely the case for many, especially for those students with less reading success), encountering new and more advanced grade level passages in the fall may explain some of the observed regression.

While some slight regression in reading skills might reasonably be expected over a period of three months characterized by little systematic reading instruction, the size of these losses (from 14% to over 30%), especially for those students with SLD and thus already at high risk for reading issues, presents a largely unacknowledged challenge for the schools. An interesting avenue for further research would be an investigation of year round schools to see if similar regressions in reading occur for students (both with and without SLD) in these schools without elongated periods of non-instruction.

Implications of the Use of the IRI as a Teacher Accountability Measure

The use of the IRI as a teacher accountability measure has the potential to compromise test administration. Such a use may tempt teachers to "teach to the test," which in turn is likely to defeat the purpose of a universal screen (Idaho State Board of Education, 2014).

To this end, the Idaho Literacy Technical Advisory Committee recently recommended that the IRI be used *only* as a universal screen and *not* as a teacher accountability measure. Further, the Idaho Literacy Technical Advisory Committee recommended that (a) the IRI be reviewed to address concerns about its technical adequacy, and (b) alternative measures be identified and explored (Idaho State Board of Education, 2014). The recommendations will be given to the Idaho Legislature for consideration during the 2015 term.

Future Research

Several recommendations for future research and practice emerge from the findings identified and discussed in this study. First, more comprehensive information about the psychometric properties of the IRI should be made available (Santi & Francis, 2012). (Such a recommendation actually applies to any assessment adopted as a statewide achievement test or universal screen.) In its present state, the IRI does seem to have some moderate power for the early identification of reading difficulties in general. Its ability to predict subsequent diagnoses of Specific Learning Disabilities is considerably weaker. The use of multiple measures incorporated into universal screens such as the IRI might be considered. However, as noted earlier, the incorporation of additional complexities in the IRI may hamper its utility as a practical and efficient screening device.

Second, it was interesting to note that, despite a persistent gap in reading skills between students with and without SLD, both groups of students demonstrated significant annual growth in this area. Thus, students with SLD can indeed evidence growth in reading. The question, then, is one of the eternal questions in special education. *How can schools help students with Specific Learning Disabilities catch up to their age and grade peers?*

One possible answer may lie in another additional finding, that in which the regression in reading skills over summers was examined. In each case (first to second grade, second to third grade, and third to fourth grade), all students (both with and without SLD) showed regression in reading skills. However, in each case, students *with* SLD showed more regression.

This phenomenon was especially significant in the summer between first and second grade. Students without SLD showed a regression of 14.3%. However, students subsequently identified as SLD showed a regression of 33.7%. Thus, one potentially promising response would be for schools to offer targeted special summer literacy programs for students who had (a) completed first grade and (b) recorded low scores on the first grade fall and spring administrations of the IRI. Based on the data analyzed in this study, if such a program could minimize the size of the reading losses experienced by students subsequently diagnosed with SLD over the summer, they would then enter second grade closer to their grade peers in reading skills.

Second, as recommended by the Idaho Literacy Technical Advisory Committee to the Idaho State Board of Education (2014), perhaps it is time to focus more on intervention than assessments; more on teaching than testing. Such an outcome may be politically unlikely at a time when greater and greater calls for higher and more teacher accountability continue. However, most practitioners are likely to agree that such a restructuring of educational priorities is long overdue.

The present study has tentatively suggested that the implementation and impact of the IRI may have resulted in an emphasis on word fluency at the cost of reading comprehension. This unanticipated outcome is likely to make reading less enjoyable for less skillful readers, ultimately generating the Matthew Effect for these individuals. This adverse result might be counterbalanced through an instructional de-emphasis on reading fluency accompanied by an increase in reading comprehension skill instruction. At a minimum, reading comprehension assessments and interventions should receive greater attention.

Third, the IRI should be absolutely disconnected from any involvement with teacher accountability. Its present dual function makes it extraordinarily challenging to carry out either professionally and ethically.

One final question emerges. Are the results found in this study of students who were in first grade in 2009 still valid after the subsequent widespread adoption of RTI practices throughout the state of Idaho? This study focused on students who took the first version of the AIMSweb IRI in 2009 prior to the implementation of RTI. Given the philosophical and educational goals of RTI, including its emphasis on very early intervention of school difficulties, one might optimistically anticipate that students today are receiving reading interventions of appropriate intensities at an earlier point in their school careers, before their literacy deficits have grown to such challenging proportions.

REFERENCES

Aaron, P. G. (1991). Can reading disabilities be diagnosed without using intelligence tests? *Journal of Learning Disabilities, 24*(3), 178–186, 191.

Adams, M. J. (1994). *Beginning to read: Thinking and learning about print*. Cambridge, MA: A Bradford Book.

AIMSweb Pro Complete. (2010). *AIMSweb*. Pearson Inc. Retrieved from www.AIMSweb.com

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2001). Schools, achievement, and inequality: A seasonal perspective. *Educational Evaluation and Policy Analysis, 23*(2), 171–191.

American Psychiatric Association. (2012). *Diagnostic and statistical manual of mental disorders*. Retrieved from http://www.dsm5.org/ProposedRevision/Pages/proposedrevision.aspx?rid=429

Anagnostopoulos, D. (2003). The new accountability, student failure, and teachers' work in urban high schools. *Educational Policy*, *17*(3), 291-316.

Apel, K. (2011). What is orthographic knowledge*? Language, Speech, and Hearing Services in Schools*, *42*(4), 592–603.

Artiles, A. J., & Trent, S. C. (1994). Overrepresentation of minority students in special education: A continuing debate. *Journal of Special Education*, *27*(4), 410-437.

Assessment: What do assessments measure? (n.d.) Retrieved January 31, 2015, from http://www.boiseschools.org/assessment/measure.html

Barr, R. D., & Flachbart, M. (2003). Idaho Reading Initiative status report: 1999-2003. Boise, ID: Center for School Improvement and Policy Studies, College of Education, Boise State University.

Barr, R. D., Flachbart, M., & Stewart, R. (2002). Idaho Reading Initiative status report: 1999–2002. Boise, ID: Center for School Improvement and Policy Studies, College of Education, Boise State University.

Barr, R. D., Joyner, S., Parrett, W., & Willison, S. (2004). *Idaho Reading Initiative: Focus on minority and low socio-economic students.* Center for School Improvement and Policy Studies, Boise, ID: Boise State University. Retrieve from http://csi.boisestate.edu/wp-content/uploads/2011/08/IdahoReadingInitiativeStatusReport2004.pdf

Barth, A. E., Stuebing, K. K., Anthony, J. L., Denton, C. A., Mathes, P. G., Fletcher, J. M., & Francis, D. J. (2008). Agreement among response to intervention criteria for identifying responder status. *Learning and Individual Differences, 18*(3), 296–307.

Baskette, M. R., Ulmer, L., & Bender, W. N. (2006). The emperor has no clothes! Unanswered questions and concerns on the response to intervention procedure. *Journal of the American Academy of Special Education Professionals,* Fall, 5–25. Retrieved from http://www.naset.org/fileadmin/user_upload/JAASEP/JAASEP_Fall_2006.pdf#page=5

Batsche, G. M., Kavale, K. A., & Kovaleski, J. F. (2006). Competing views: A dialogue on response to intervention. *Assessment for Effective Intervention*, *32*(1), 6–19.

Bear, D. R., Invernizzi, M., Templeton, S., & Johnston, F. (1996). Words their way: Word study *for phonics, vocabulary, and spelling instruction*. Upper Saddle River, NJ: Prentice-Hall, Inc.

Beaver, J. (1997). *Developmental reading assessment*. Parsippany, NJ: Celebration Press.

Berkeley, S., Bender, W. N., Gregg Peaster, L., & Saunders, L. (2009). Implementation of response to intervention: A snapshot of progress. *Journal of Learning Disabilities*, *42*(1), 85–95.

Berninger, V. W., Abbott, R. D., Zook, D., Ogler, S., Lemos-Britton, Z., & Brooksher, R. (1999). Early intervention for reading disabilities: Teaching the alphabet principle in a correctionist framework. *Journal of Learning Disabilities*, *32*(6), 491–503.

Birch, K. G. (2004). Phonological processing, automaticity, auditory processing, and memory in slow learners and children with reading disabilities (Doctoral dissertation). Retrieved from ProQuest UMI Dissertation Publishing (UMI Microform 3116264).

Blackorby, J. & Wagner, M. (1996). Longitudinal postschool outcomes of youth with disabilities: Findings from the National Longitudinal Transition Study. *Exceptional Children, 62*(5), 399–413.

Boise School District Statistics. (2015). *Boise School District statistics: Enrollment figures*. Retrieved from http://www.boiseschools.org/modules/cms/pages.phtml?pageid=272094

Boise School District. (2015). *Human Resources: Our community*. Retrieved from http://school.boiseschools.org/modules/cms/pages.phtml?pageid=252345&sessionid=75a8b92418ca2df987c2bb63c79dbcb9&sessionid=74a8b92418ca2df987c2bb63c79dbcb9

Bormuth, J. R. (1973 - 74). Reading literacy: Its definition and assessment. *Reading Research Quarterly*, *9*(1), 7–66.

Boscardin, C. K., Muthen, B., Baker, E. L., & Francis, D. J. (2008). Early identification of reading difficulties using heterogeneous developmental trajectories. *Journal of Educational Psychology, 100*(1), 192–208.

Bradley, J. M, & Ames, W. S. (1977). Readability parameters of basal readers. *Journal of Literacy Research*, *9*(2), 175–183.

Bradley, R., Danielson, L., & Doolittle, J. (2005). Response to intervention. *Journal of Learning Disabilities, 38*(6), 485–486.

Bradley, R., Danielson, L., & Doolittle, J. (2007). Responsiveness to intervention: 1997 to 2007. *Teaching Exceptional Children*, *39*(5), 8–12.

Bradley, R., Danielson, L. C., & Hallahan, D. P. (2002). *Identification of learning disabilities: Research to practice.* Mahwah, NJ: L. Erlbaum Publishers.

Brummitt-Yale, J. (2014). What is content area reading*? K12 Reader: Reading Instruction Resources*. Retrieved from http://www.k12reader.com/what-is-content-area-reading/

Buck, J., & Torgesen, J. (2003). The relationship between performance on a measure of oral reading fluency and performance on the Florida Comprehensive Assessment Test. Retrieved from http://www.fcrr.org/technicalreports/TechnicalReport1.pdf.

Buck, J., Torgesen, J., & Schatschneider, C. (2009). *Predicting FCAT- SSS scores using prior performance on the FCAT-SSS, FCAT-NRT, and SAT9.* Technical Report # 4, Florida Center for Reading Research. Tallahassee, FL. Retrieved from http://www.fcrr.org/TechnicalReports/TechnicalReport1.pdf

Buffum, A., Mattos, M., & Weber, C. (2009). *Pyramid response to intervention*. Bloomington, IN: Solution Tree.

Burke, M. D., Hagan-Burke, S., Zou, Y., & Kwok, O. (2009). A structural equation model using fluency-based early literacy measures to predict emerging reading ability in kindergarten. *Remedial and Special Education*, 31(5), 385–399.

Bus, A. B., van IJzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. Review of Educational Research, *65*(1), 1–21.

Carlisle, J. F., Schilling, S. G., Scott, S. E., & Zeng, J. (2004). *Do fluency measures predict reading achievement? Results from the 2002-2003 school year in Michigan's Reading First schools.* Technical Report #1, Evaluation of Reading First in Michigan. Ann Arbor: University of Michigan.

Carnevale, A.P. (1991). American and the new economy: How new competitive standards are radically changing *American workplaces.* San Francisco, CA: Jossey-Bass.

Castillo, J. M. (2005). *The predictive validity of four reading fluency measures on a state's "High Stakes."* (Doctoral dissertation). Retrieved from Graduate School Thesis and Dissertations, http://shttp://scholarcommons.usf.edu/etd/2815

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading, 3*(4), 331–361.

Chall, J. S. (1967). *Learning to read: The great debate*. New York: McGraw Hill.

Chall, J. S. (1983). *Stages of reading development*. New York: McGraw Hill.

Christ, J. J., & Silberglitt, B. (2007). Estimates of the standard error of measurement for curriculum-based measures of oral reading fluency. *School Psychology Review, 36*(1), 130–146.

Cipielewski, J. & Stanovich, K. E. (1992). Predicting growth in reading ability from children's exposure to print. *Journal of Experimental Child Psychology*, *54*(1), 74–89.

Clemens, N. H., Shapiro, E. S., & Thoemmes, F. (2011). Improving the efficacy of first grade reading screening: An investigation of word identification fluency with other early literacy indicators. *School Psychology Quarterly*, *26*(3), 231–244.

Committee on the Prevention of Reading Difficulties in Young Children. Snow, C.E., Burns, M.S., & Griffin, P. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Consortium on Reading Excellence. (2008). Assessing reading: multiple measures for kindergarten through twelfth grade. Novato, CA: Arena Press.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research, 66*(3), 227–268.

Cortiella, C. (2011). *The State of Learning Disabilities.* New York, NY: National Center for Learning Disabilities.

Coutinho, M. (1995). Who will be learning disabled after the reauthorization of IDEA? Two very distinct perspectives. *Journal of Learning Disabilities*, *28*(10), 664–668.

Coutinho, M. J., Oswald, D. P., & Best, A. M. (2002). The influence of sociodemographics and gender on the disproportionate identification of minority students as having learning disabilities. *Remedial and Special Education*, *23*(1), 49–59.

Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology, 82*(4), 733–740.

Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology*, 83(2), 264–274.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*(6), 934–945.

Cunningham, A. E., & Stanovich, K. E. (1998). What reading does for the mind. *American Educator*, *22*(1), 8–15.

Cunningham, A. E., & Stanovich, K. E. (2001). What reading does for the mind. *Journal of Direct Instruction*, *1*(2), 137–149.

Cunningham, A., & Stanovich, K. (2003). Reading can make you smarter. *Principal*, *83*(2), 34–39.

Davis, G. N., Lindo, E. J., & Compton, D. L. (2007). Children at risk for reading failure: Constructing an early screening measure. *Teaching Exceptional Children*, *39*(5), 32–37.

Davidson, E., Reback, R., Rockoff, J., & Schwartz, H. (2013). Fifty ways to leave a child behind: Idiosyncrasies and discrepancies in States' implementation of NCLB.

Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.columbia.edu/~ekd2110/Fifty_Ways_4_5_2013.pdf

Deno, S. (2003). Curriculum-based measures: Development and perspectives. *Assessment for Effective Intervention, 28*(3-4), 2–12.

Deno, S .L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*(3), 219–232.

Deno, S. L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review, 15*(3), 358–374.

Deno, S. L. (1989). Curriculum-based measurement and alternative special education services: A fundamental and direct relationship. In M.R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 1-17). New York, NY: Guilford Press.

Deno, S. L., Marston, D., & Tindal, G. (1985). Direct and frequent curriculum-based measurement: An alternative for educational decision making. *Special Services in the Schools, 2*(2), 5–27.

Dion, E., Dube, I., Roux, C., Landry, D., & Bergeron, L. (2012). How curriculum-based measures help us detect word recognition problems in first graders. In C. A. Espin, K. L. McMaster, S. Rose, & M. M. Wayman, (Eds.), *A measure of success: The influence of curriculum-based measurement on education* (pp.101–111). Minneapolis, MN: University of Minnesota Press.

Downey, D. B., von Hippel, P. T., Broh, B. A. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. *American Sociological Review, 69*(5), 613–635.

Ehri, L. C. (1998). Grapheme–phoneme knowledge is essential for learning to read words by sight in English. In J. L. Metsala & L. C. Ehri (Eds), Word recognition in beginning literacy (pp. 3-40). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9*(2), 167–188.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36*(3), 250–287.

Elley, W. B. (1991). Acquiring literacy in a second language: The effect of book-based programs. *Language Learning, 41*(3), 375–411.

Elliot, J., Lee, S. W., & Tollefson, N. (2001). A reliability and validity study of the Dynamic Indicators of Basic Early Literacy Skills–Modified. *School Psychology Review, 30*(1), 33–49.

Elementary and Secondary Education Act of 1965, P.L. 89-750, 20 U.S.C. § 6301 et seq.

Feitelson, D., Goldstein, Z., Iraqi, J., & Share, D. L. (1993). Effects of listening to story reading on aspects of literacy acquisition in a diglossic situation. *Reading Research Quarterly, 28*(1), 70–79.

Finlan, T. G. (1992). Do state methods of quantifying a severe discrepancy result in fewer students with learning disabilities? *Learning Disability Quarterly, 15*(2), 129–134.

Fisk, J. (n.d.). Idaho Reading Indicator: Teacher information guide. *Idaho State Department of Education*. Retrieved from http://www.google.com/search?client=safari&ris=en&q=Idaho+Reading+Indicator+Teacher+guide+fisk&ie=UTF-8

Fletcher, J. M, Coulter, W. A., Reschly, D. J., & Vaughn, S. (2004). Alternative approaches to the definition and identification of learning disabilities: Some questions and answers. *Annals of Dyslexia, 54*(2), 304–331.

Fletcher, J. M., Denton, C., & Francis, D. J. (2005). Validity of alternative approaches for the identification of learning disabilities: Operationalizing unexpected underachievement. *Journal of Learning Disabilities*, *38*(6), 545–552.

Fletcher, J. M., Foorman, B. R., Boudousquie, A., Barnes, M., Schatschneider, C., & Francis, D. (2002). Assessment of reading and learning disabilities: A research-based intervention-oriented approach. *Journal of School Psychology, 40*(1), 27–63.

Fletcher, J. M., Morris, R. D., & Lyon, G. R. (2003). Classification and definition of learning disabilities: An integrative perspective. In H. L. Swanson, K. R. Harris, and S. Graham (Eds.), *Handbook of learning disabilities* (pp. 30-56). New York, NY: Guilford.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L. Liberman, I. Y., Stuebing, K. K.,...Shaywitz, B. A. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definition. *Journal of Educational Psychology*, *86*(1), 6–23.

Foorman, B. R. (2007). Primary prevention in classroom reading instruction. *Teaching Exceptional Children, 39*(5), 24–30.

Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*(1), 97–55.

Fuchs, D., Compton, D. L., Fuchs, L. S., Bryant, J. & Davis, G. N. (2007). Making “secondary interventions” work in a three-tier responsiveness-to-intervention model: findings from the first-grade longitudinal reading study of the National Research Center on Learning Disabilities. *Reading and Writing*, *21*(4), 413–436.

Fuchs, D., & Deshler, D. D. (2007). What we need to know about Responsiveness to Intervention (and shouldn’t be afraid to ask). *Learning Disabilities Research & Practice, 22*(2), 129–136.

Fuchs, D., Deshler, D., & Reschly, D. (2004). National research center on learning disabilities: Multimethod studies of identification and classification issues. *Learning Research Quarterly*, *27*(4), 189–194.

Fuchs, D., Fuchs, L. S., & Compton, D. L. (2004). Identifying reading disabilities by responsiveness-to-instruction: Specifying measures and criteria. *Learning Disability Quarterly*, *27*(4), 216-227.

Fuchs, D., Fuchs, L. S., Mathes, P. G., Lipsey, M. W., & Roberts, P.H. (2002). Is "learning disabilities" just a fancy term for low achievement? A meta-analysis of reading differences between low achievers without the label. In R. Bradley, L. Danielson, & D.P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp.737-763). Mahwah, NJ: Erlbaum.

Fuchs, D., Fuchs, L., & Stecker, P. M. (2010). The "blurring" of special education in a new continuum of general education placements and services. *Exceptional Children*, *76*(3), 301–323.

Fuchs, D., & Young, C. L. (2006). On the irrelevance of intelligence in predicting responsiveness to reading instruction. *Exceptional Children*, *73*(1), 8–30.

Fuchs, L., & Fuchs, D. (1999). Monitoring student progress towards the development of reading competence: A review of three forms of classroom-based assessment. *School Psychology Review*, *28*(4), 659–671.

Fuchs, L. S., & Deno, S. L. (1994). Must instructionally useful performance assessment be based in the curriculum? *Exceptional Children, 61*(1), 15–24.

Fuchs, L. S., & Fuchs, D. (2007). A model for implementing responsiveness to intervention. *Teaching Exceptional Children*, *39*(5), 14–23.

Fuchs, L. S., & Fuchs, D. (2011). Using CBM for progress monitoring in reading. *National Center on Student Progress Monitoring.* Retrieved from http://eric.ed.gov/?id=ED519252

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2004). Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children*, *71*(1), 7–21.

Fuchs, L. S., Fuchs, D., Hosp, M., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*(3), 239–256.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1993). Technological advances linking the assessment of students' academic proficiency to instructional planning. *Journal of Special Education Technology, 12*(1), 49–62.

Garland, D. & Strosnider, R. (2005). Responsiveness to Intervention and learning disabilities. *Learning Disability Quarterly*, *28(*4), 249–260.

Gerber, M. (2005) Teachers are still the test: Limitations of response to instruction strategies for identifying children with learning disabilities. *Journal of Learning Disabilities*, *28*(6), 516–524.

Gersten, R. & Dimino, J. A. (2006). RTI (Response to Intervention): Rethinking special education for students with reading difficulties (yet again). *Reading Research Quarterly, 41*(1), 99–108.

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-states outcomes. *Scientific Studies of Reading*, *5*(3), 257–288.

Goodman, G., & Webb, M. A. (2006). Reading disability referrals: Teacher bias and other factors that impact response to intervention. *Learning Disabilities: A Contemporary Journal*, *4*(2), 59–70.

Gordon Commission. (2012, April). Shifting paradigms: Beyond the abstract. *Assessment, Teaching, and Learning, 2*(2). Retrieved from http://www.gordoncommission.org/rsc/pdfs/vol_2_no_2_18704.pdf

Hallahan, D. P., Keller, C. E., Martinez, E. Z., Byrd, E. S., Gelman, J. A., & Xitao, F. (2007). How variable are interstate prevalence rates of learning disabilities and other special education categories? A longitudinal comparison. *Exceptional Children, 73*(2), 136–146.

Henderson, R., & Hopper, G. (2010, October). *Specific Learning Disability criteria and resources.* Paper presented at the Council for Exceptional Children State Conference, Sun Valley, Idaho. Retrieved from http://idahotc.com/Portals/23/Docs/training/CEC%20state%20Conference.pdf

Hieronymus, A. N., Hoover, H. D., & Lindquist, E. F. (1982). *Iowa Tests of Basic Skills.* Chicago, IL: Riverside.

Hintze, R. M., Ryan, A. L., Stoner, G. (2003). Concurrent validity and diagnostic accuracy of the Dynamic Indicators of Basic Early Literacy Skills and the Comprehensive Test of Phonological Processing. *School Psychology Review, 34*(4), 541–556.

Horn, C., Plazas, S. B., Coverdale, J. H., Louie, A. K., & Weiss Roberts, L. (2009). Educational research questions and study design. Academic Psychiatry, *33*, 261–267.

Hosp, M. K., Hosp, J. L., & Howell, K. W. (2007). *The ABCs of CBM: A practical guide to curriculum-based measurement.* New York: Guilford Press.

Hudson, R. F., High, L., & Otaiba, S. A. (2007). Dyslexia and the brain: What does current research tell us? *The Reading Teacher*, *60*(6), 506–515.

Hughes, C., & Dexter, D. D. (n.d. a). Selecting a scientifically based core curriculum for Tier 1. *RTI Action Network*. Retrieved from http://www.rtinetwork.org/learn/research/selectingcorecurriculum-tier1

Hughes, C., & Dexter, D. D. (n.d. b). Universal Screening within a response-to-intervention model. Retrieved from http://www.rtinetwork.org/learn/research/universal-screening-within-a-rti-model

Hunley, S. A., Davies, S. C., & Miller, C. R. (2013). The relationship between curriculum-based measures in oral reading fluency and high-stakes tests for seventh grade students. *RMLE Online, 36*(5), 1–8. Retrieved from http://search.proquest.com/docview/1426958974?accountid=9649

Hurford, D. P., & Johnson, M. (1994). Early identification and remediation of phonological-processing deficits in first-grade children at risk for reading disabilities. *Journal of Learning Disabilities*, *27*(10), 647–660.

Hyatt, J. (2010, January). *The nuts and bolts of the new specific learning disability (SLD) criteria.* Idaho SDE Special Education Division. Idaho Live Webinar, Boise, Idaho. Retrieved from https://docs.google.com/viewer?a=v&q=cache:D0T9yfYu44EJ:idahotc.com/Portals/23/Docs/Archived%2520training/SLDNutsBolts_01-21-

10.ppt+specific+learning+disabilities+idaho+2010&hl=en&gl=us&pid=bl&srcid=ADGEESiXu3Un3gbFvbSRGys2nhmbkBBVzdHVsPopy2dIIDbINp3PkAS3rlP5djxvsqhCncJf395EJB_X_Y-2sUwD33KsH1vEDcO-mLdCEcrKOcVnJjPVVcon-CyJgmBi3WxIyiRanhu&sig=AHIEtbQs81KXRPM_9J_-dXeVwVNFyx4FSQ

Idaho Comprehensive Literacy Act, Idaho Code §§§33-1614, 33-1615, 33-1207A (1999).

Idaho Reading Indicator (IRI): Parent Brochure, n.d. Retrieved from https://www.sde.idaho.gov/site/reading_indicator/docs/parents/IRI%20Parent%20Brochure.pdf

Idaho Standards Achievement Tests (ISATs). (n.d.). Idaho State Department of Education. Retrieved from http://www.sde.idaho.gov/site/assessment/isat/

Idaho State Board of Education. (2014). *Idaho Literacy Task Force report*. Retrieved from http://www.sde.idaho.gov/site/literacyTech/

Idaho State Department of Education. (2004). *Idaho Reading Indicator: Instructional support guide.* Retrieved from http://www.sde.idaho.gov/readingindicator/docs/Teachers/SupportGuideFinal7202004.pdf

Idaho State Department of Education. (2007, June 13–14). *State Department of Education.* Meeting Minutes. Retrieved from http://www.boardofed.idaho.gov/meetings/board/archive/2007/06_13_07/sde.pdf

Idaho State Department of Education. (2009). *Idaho special education manual 2007, revised 2009.* Boise, Idaho: Idaho State Dept. of Education, Division of Student Achievement and School Accountability.

Idaho State Department of Education. (2012). *Idaho performance data IRI comparative reports.* Retrieved from http://www.sde.idaho.gov/ipd/iri/IriAnalysis.asp

Idaho State Department of Education. (2013) *Idaho Reading Indicator (IRI): Testing quick guide, 2012–2013 edition.* Retrieved from http://www.sde.idaho.gov/site/reading_indicator/docs/admin/IRITestingGuide2012-2013.pdf

Idaho State Department of Education. (2014a). Department announces Idaho Literacy Technical Advisory Committee. *Idaho Education News: The Official Blog of the Idaho State Department of Education.* Retrieved from http://educationidaho.blogspot.com/2014/06/department-announces-idaho-literacy.html

Idaho State Department of Education. Office of the State Department of Education. (2014b). *Public school information, Legislative report*; *Idaho Reading Initiative (IRI).* Retrieved from https://www.sde.idaho.gov/site/publications/communications_docs/2014/2014%20IRI%20Legislative%20Report_FINAL.pdf

Idaho State Department of Education. (2015). Idaho's ESEA waiver. *Elementary & Secondary Education Act (ESEA).* Retrieved from _indicator/ docs/admin/2ndGrd_IRIProctorinfoTemplate.pdf

Individuals with Disabilities Education Improvement Act [IDEA], P.L. §108-466. (2004).

Jenkins, J. R., & Fuchs, L. S. (2012). Curriculum-based measurement: The paradigm, history, and legacy. In C. A. Espin, K. L. McMaster, S. Rose, & M. M. Wayman, (Eds.), *A measure of success: The influence of curriculum-based measurement on education* (pp.7–23). Minneapolis, MN: University of Minnesota Press.

Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for at-risk readers in a response to intervention framework. *School Psychology Review, 36*(4), 582–600.

Jenkins, J. R., & Johnson, E. (2008). Universal screening for reading problems: Why and how should we do this? *RTI Action Network*. Retrieved from http://www.rtinetwork.org/_essential/assessment/screening/readingproblems

Jenkins, J. R., & O'Connor, R. E. (2002). Early identification and intervention for young children with reading/learning disabilities. In R. Bradley, L. Danielson, & D. Hallahan (Eds.), *Identification of learning disabilities: Research to Practice* (pp.99–149). Mahwah, NJ: Erlbaum

Johnson, E., Mellard, D. F., & Byrd, S. E. (2005). Alternative models of learning disabilities identification: Considerations and initial conclusions. *Journal of Learning Disabilities*, *38*(6), 569–572.

Johnson, E., Mellard, D. F., Fuchs, D., & McKnight, M. A. (2006). *Responsiveness to intervention (RTI): How to do it.* Lawrence, KS: National Research Center on Learning Disabilities. Retrieved from www.rtinetwork.org/essential/assessment/screening/screening-for-reading-problems-in-an-rti-framework

Johnson, E. S., Humphrey, M., Mellard, D. F., Woods, K., & Swanson, H.L. (2010). Cognitive processing deficits and students with specific learning disabilities: A selective meta-analysis of the literature. *Learning Disability Quarterly, 33*(1), 3–18.

Johnson, E. S., Jenkins, J. R., Petscher, Y., & Catts, H. W. (2009). How can we improve the accuracy of screening instruments? *Learning Disabilities Research & Practice*, *24*(4), 174–185.

Johnson, E. S., Pool, J., & Carter, D. R. (n.d.). Screening for reading problems in grades 1 through 3: An overview of select measures. *RTI Action Network*. Retrieved from http://www.rtinetwork.org/essential/assessment/screening/screening-for-reading-problems-in-grades-1-through-3

Jorgensen, M. A., & Hoffmann, J. (2003). *History of No Child Left Behind Act of 2001 (NCLB)*. (Assessment Report). Retrieved from Pearson Education, Inc. website: http://www.pearsonassessments.com/NR/rdonlyres/D8E33AAE-BED1-4743-98A1-BDF4D49D7274/0/HistoryofNCLB.pdf

Kaminski, R. H., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review*, *25*(2), 215.

Karlsen, B., Madden, R., & Gardner, E. F. (1984). *Stanford Diagnostic Reading Test* (33rd Ed.). San Antonio, TX: The Psychological Corporation, Harcourt Brace Jovanovich.

Kavale, K. A. (2002). Discrepancy models in the identification of learning disability. In R. Bradley, L. Danielson. & D. P. Hallahan (Eds), *Identification of learning disabilities: Research to practice* (pp. 369–426). Mahwah, NJ: Erlbaum.

Kavale, K. A. (2005). Identifying specific learning disabilities: Is responsiveness to intervention the answer? *Journal of Learning Disabilities*, *38*(6), 553–562.

Kavale, K. A., & Spaulding, L. S. (2008).*Is response-to-intervention good policy for specific learning disability?* Faculty Publications and Presentations. Paper 119. Retrieved from http://digitalcommons.liberty.edu/educ_fac_pubs/119

Keogh, B .K. (2005). Revisiting classification and identification*. Learning Disabilities Quarterly*, *28*(2), 100–102.

Kim, J. S. (2006). Effects of a voluntary summer reading intervention on reading achievement: Results from a randomized field trial. *Educational Evaluation and Policy Analysis*, *28*(4), 335–355.

Kim, Y. S., Wagner, R. K., Foster, E. (2011). Relations among oral reading fluency, silent reading fluency, and reading comprehension: A latent variable study of forst-grade readers. *Scientific Studies of Reading*, *15*(4), 338–362.

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, *95*(1), 23–31.

Lau, M. Y., Sieler, J. D., Muyskens, P., Canter, A., Vankeuren, B., & Marston, D. (2006). Perspectives on the use of the problem-solving model from the viewpoint of a school psychologist, administrator, and teacher from a large Midwest urban school district. *Psychology in the Schools*, *43*(1), 117–127.

Leipzig, D. H. (January, 2001). What is reading? *WETA*. Retrieved from http://www.readingrockets.org/article/what-reading

Linn, R. L., Baker, E. L., Betebenner, D.W. (2002). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. *Educational Researcher, 31*(6), 3–16.

Lyon, G. R. (1996). Learning Disabilities. *The Future of Children*, *6*(1), 54–76.

MacMillan, D. L., Gresham, F. M., Siperstein, G. N., & Bocian, K. M. (1996). The labyrinth of I.D.E.A.: School decisions on referred students with subaverage general intelligence. *American Journal on Mental Retardation*, *101*(2), 161–174.

MacMillan, D. L., & Speece, D. L. (1999). Utility of current diagnostic categories for research and practice. In R. Gallimore, L. P. Bernheimer, D. L. MacMillan, D. L. Speece, & S. Vaugh (Eds.), *Developmental perspectives on children with high-incidence disabilities* (pp.111–133). Mahwah, NJ: Erlbaum.

Marston, D. (2005). Tiers of intervention in responsiveness to intervention: Prevention outcomes and learning disabilities. *Journal of Learning Disabilities*, *8*(6), 539–544.

Marston, D. (2012). School-and district-wide implementation of curriculum-based measurement in the Minneapolis Public Schools. In C. A. Espin, K. L. McMaster, S. Rose, & M. M. Wayman, (Eds.), *A measure of success: The influence of curriculum-based measurement on education* (pp.59–78). Minneapolis, MN: University of Minnesota Press.

Mastropieri, M., & Scruggs, T. (2005). Feasibility and consequence of response to intervention: Examination of the issues and scientific evidence as a model for the identification of individuals and learning disabilities. *Journal of Learning Disabilities*, *28*(6), 525–531.

Mather, N., & Gregg, N. (2006). Specific learning disabilities: Clarifying, not eliminating, a construct. *Professional Psychology: Research and Practice, 37*(1), 99–106.

McCoach, D. B., O’Connell, A. A., Reis, S. M., Levitt, H. A. (2006). Growing readers: A hierarchical linear model of children’s reading growth during the first 2 years of school. *Journal of Educational Psychology*, *98*(1), 14–28.

McKenzie, R. G. (2009). Obscuring vital distinctions: The oversimplification of learning disabilities within RTI. *Learning Disability Quarterly, 32*(4), 203–215.

McNeil, L. (2002). *Contradictions of school reform: Educational costs of educational testing*. New York, NY: Routiedge.

Mellard, D. F., & Johnson, E. (2008). *RTI: A practitioner's guide to implementing response to intervention.* Thousand Oaks, CA: Crowin Press.

Mercer, C. D., Jordan, L., Allsoppp, D. H., & Mercer, A. R. (1996). Learning disabilities definitions and criteria used by state education departments. *Learning Disability Quarterly*, *19*(4), 217–232.

Mercer, C. D., King-Sears, P., & Mercer, A. R. (1990). Learning disabilities definitions and criteria used by state education departments. *Learning Disability Quarterly, 13*(32), 141–152.

Moats, L. C. & Lyon, G. R. (1993). Learning Disabilities in the United States: Advocacy, science, and the future of the field. *Journal of Learning Disabilities, 26*(5), 282–294.

Montgomery, J. K., & Moore-Brown, B. (2006). *Response to Intervention: An alternative to special education.* Rockville, MD: American Speech, Language, and Hearing Association.

Morgan, P., Fuchs, D., Compton, D., Cordray, D., & Fuchs, L. (2008). Does early reading failure decrease children's reading motivation? *Journal of Learning Disabilities, 41*(5), 387–404.

Muyskens, P., Betts, J., Lau, M. Y., & Marston, D. (2009). Predictive Validity of Curriculum-Based Measures in the Reading Assessment of students who are English language learners. *California Association of School Psychologists*, *14*, 11–21.

National Association of State Directors of Special Education [NASDE] and Council of Administrators of Special Education [CASE]. (2006). *Response to Intervention: White Paper on RtI*. Retrieved from http://www.nasdse.org/documents/RtIAnAdministratorsPerspective1-06.pdf

National Center for Family Literacy (U.S.). (2008). *Developing early literacy Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy.

National Center on Student Progress Monitoring [NICHHD]. (n.d.) Frequently asked questions about curriculum-based measurement. *Ideas that Work.* US. Office of Special Education Programs. Retrieved from http://www.studentprogress.org/faq.asp#_Toc89594736

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* (NIH Pub.NO. 00-4769). Washington, DC: U.S. Government Printing Office.

Nave, G., & Burke, A. (2007a). *Idaho Reading Indicator study: Technical Report, Fall 2006*. Portland, OR: Northwest Regional Education Laboratory.

Nave, G., & Burke, A. (2007b). *Idaho Reading Indicator study: Technical Report, Winter 2007*. Portland, OR: Northwest Regional Education Laboratory.

Neuman, S. B., Copple, C., & Brendekamp, S. (1998). Learning to read and write: Developmentally appropriate practices for young children. Young Children, *53*(4), 30–46.

Ofiesh, N. (2006). Response to intervention and the identification of specific learning disabilities: Why we need comprehensive evaluations as part of the process. *Psychology in the Schools, 43*(8), 883–888.

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40*(2), 184–202.

Pearson, P. D. (2006). Foreword. In K. S. Goodman (Ed.), The truth about DIBELS: What it is, what it does. Portsmouth, NH: Heinemann.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research* (3rd Ed.). Fort Worth, TX: Harcourt Brace.

Pikulski, J. J., & Chard, D. J. (2005). Fluency: Bridge between decoding and reading comprehension. *Reading Teacher*, *58*(6), 510–519.

Pinnell, G. S., Pikulski, J. J., Wilxson, K. K., Campbell, J. R., Gough, P. B., & Betty, A. S. (1995). *Listening to children read aloud.* Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.

Powell-Smith, K. A. & Bradley-Klug, K. L. (2001). Another look at the "C" in CBM: Does it really matter if curriculum-based measurement reading probes are curriculum-based? *Psychology in the Schools, 338*(4), 299–312.

Pretorius, E. J. (2000). "What they can't read will hurt them": Reading and academic achievement. *Innovation, 21*, 33–41. Retrieved from http://www.innovation.ukzn.ac.za/InnovationPdfs/No21pp33-41Pretorius.pdf

Rasinski, T. V. (2004). Assessing reading fluency. *Pacific Resources for Education and Learning*. Retrieved from http://www.prel.org/media/137837/assessing2011.pdf

Rasinski, T. V., & Hoffman, J. V. (2003). Theory and research into practice: Oral reading in the school literacy curriculum. *Reading Research Quarterly, 38*(4), 510 –522.

Reid, D., Hresko, W., & Hammill, D. (2001). *Test of early reading abilitiy-3rd edition.* Austin, TX: PRO-ED.

Reschly, D. J. (1996). Identification and assessment of children with disabilities. *The Future of Children: Special Education for Children with Disabilities, 6*(1), 40–53.

Resnick, L. B., & Berger, L. (2010). An American examination system. Retrieved from http://www.k12center.org/rsc/pdf/ResnickBergerSystemModel.pdf

Riedel, B. W. (2007). The relation between DIBELS, reading comprehension, and vocabulary in urban first-grade students. *Reading Research Quarterly, 42*(4), 546–562.

Ritchey, K. D. (2008). Assessing letter sound knowledge: A comparison of Letter Sound Fluency and nonsense word fluency. *Exceptional Children*, 74(4), 487–508. Retrieved from http://search.proquest.com/docview/201203384?accountid=9649

Roehrig, A.D., Petscher, Y, Nettles, S.M, Hudson, R.F, & Torgesen, J.K. (2008). Accuracy of the DIBELS oral reading fluency measure for predicting third grade reading comprehension outcomes. *Journal of School Psychology*, *46*(3), 343–366.

Rouse, H. L., & Fantuzzo, J. W. (2006). Validity of the Dynamic Indicators for Basic Early Literacy Skills as an indicator of early literacy for urban kindergarten children. *School Psychology Review, 35*(3), 341–355.

Samuels, S. J. (2002). Reading fluency: Its development and assessment. In A. E. Farstrup & S. J. Samuels (Eds.), What research has to say about reading instruction (3rd ed., pp. 166–183). Newark, DE: International Reading Association.

Samuels, S. J. (2007). The DIBELS test: Is speed of baking at print what we mean by reading fluency? *Reading Research Quarterly, 42*(4), 543–566.

Santi, K., & Francis, D. (2012). *Idaho k-3 reading assessment analysis*. Boise, Idaho: Idaho State Department of Education.

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In B.K. Shapiro, P. J. Accardo, & A. J. Capute (Eds.), *Specific reading disabilities* (pp.75–119). Timonium, MD: York Press, Inc.

Scarborough, H. S. (2009). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In F. Fletcher-Campbell, J. Soler, & G. Reid (Eds.), *Approaching difficulties in literacy development: Assessment, pedagogy and programmes* (pp. 23–38). Thousand Oaks, CA: SAGE Publications Inc.

Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology*, *96*(2), 265–282.

Schilling, S. G., Carlisle, J. F., Scott, S. E., & Zeng, J. (2007). Are fluency measures accurate predictors of reading achievement? *The Elementary School Journal, 107*(5), 429–428.

Scott, B. (n.d.). Free test readability consensus calculator. Retrieved from http://www.readabilityformulas.com/free-readability-formula-tests.php

Scruggs, E., and Mastropieri, M. A. (2002). On babies and bathwater: Addressing the problems of identification of learning disabilities. *Learning Disability Quarterly*, *25*(3), 155–169.

Semrud-Clikeman, M. (2005). Neuropsychological aspects for evaluating learning disabilities. *Journal of Learning Disabilities*, *38*(6), 563–568.

Semrud-Clikeman, M., Biederman, J., Sprich-Buckminster, S., Lehman, B. K., Faraone, S. V., & Normon, D. (1992). Comorbidity between ADHD and learning disability: A review and report in clinically referred sample. *Journal of the American Academy of Child & Adolescent Psychiatry*. *31*(3), 439–448.

Shapiro, E. S. (n.d.). Tiered instruction and intervention in a Response-to-Intervention model. *RTI Action Network*. Retrieved from http://www.rtinetwork.org/essential/tieredinstruction/tiered-instruction-and-intervention-rti-model

Shapiro, E. S., Solari, E., & Petscher, Y. (2008). Use of a measure of reading comprehension to enhance prediction on the state high stakes assessment. *Learning and Individual Differences*, *18*(3), 316–328.

Shaw, S. F., Cullen, J. P., McGuire, J. M., & Brincherhoff, L. C. (1995). Operationalizing a definition of learning disabilities. *Journal of Learning Disabilities*, *28*(9), 586–597.

Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P....Gore, J. C. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. *Biological Psychiatry*, *55*(9), 926–933.

Sherman, L. (2001) Birth of a standard: Idaho spells out statewide expectations for reading. *NW Education, 7*(1), 12–21.

Shinn, M. R. (1989). Identifying and defining academic problems: CBM screening and eligibility procedures (pp. 90–129). In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children.* New York, NY: Guilford Press.

Shinn, M. R (Ed.) (1998). *Advanced applications of curriculum-based measurement.* New York, NY: The Guilford Press.

Shinn, M. R, Good, R., Knutson, N., Tilly, W. D., & Collins, V. L. (1992). Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review*, *21*(3), 459–479.

Sibley, D., Biwer, D., & Hesch, A. (2001, April). *Establishing curriculum-based measurement oral reading fluency performance standards to predict success on local and state tests of reading achievement*. Paper presented at the Annual Meeting of the National Association of School Psychologists, Washington, DC.

Siegel, L. S. (1989). IQ is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities*, *22*(8), 469–486.

Siegel, L. S. (1999a). Issues in the definition and diagnosis of learning disabilities: A perspective on Guckenberger v. Boston University. *Journal of Learning Disabilities*, *32*(4), 304–319.

Siegel, L. S. (1999b). Learning disabilities: The roads we have traveled and the path to the future. In R. J. Sternberg & L. Spear-Swerling (Eds.), *Perspectives on learning disabilities.* (pp. 159–176). Boulder, CO: Westview Press.

Snowling, M .J. (2012). Early identification and interventions for dyslexia: A contemporary view. *Journal of Research in Special Education Needs, 13*(1), 7–14.

Sparks, D.S. (2012, November). Today’s tests seen as bar to better assessment. *Ed Week*, *32*(12), 8–9.

Spear-Swerling, L. (2006). Children's reading comprehension and oral reading fluency in easy text. *Reading and Writing*, *19*(2), 199–220.

Speece, D., Case, L., & Molloy, D. (2003). Responsiveness to general education instruction as the first gate to learning disabilities identification. L*earning Disabilities Research & Practice*, *18*(3), 147–156.

Speece, D. L., & Case, L. (2001). Classification in context: An alternative approach to identifying early reading disability. *Journal of Educational Psychology, 93*(4), 735–749.

Speece, D. L., Mills, C., Ritchey, K. D., & Hillman, E. (2003). Initial evidence that letter fluency tasks are valid indicators of early reading skill. *The Journal of Special Education*, *36*(4), 223–233.

Speece, D. L., & Ritchey, K. (2005). A longitudinal study of the development of oral reading fluency in young children at risk for reading failure. *Journal of Learning Disabilities, 38*(5), 387–399.

Stage, S. A, Abbott, R. D., Jenkins, J. R., & Berninger, V. W. (2003). Predicting responses to early reading intervention from verbal IQ, reading-related language abilities, attention ratings, and verbal IQ-word reading discrepancy: Failure to validate discrepancy method. *Journal of Learning Disabilities, 36*(1), 24–33.

Stage, S. A., & Jacobsen, M. D. (2001). Predicting student success on a state-mandated performance-based assessment using oral reading fluency. *School Psychology Review, 30*(3), 407–419.

Stage, S. A., Sheppard, J., Davidson, M. M., & Browning, M. M. (2001). Prediction of first-graders' growth in oral reading fluency using kindergarten letter fluency. *Journal of School Psychology*, *39*(3), 225–273.

Stanovich, K. E. (1986). Matthew Effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360–407.

Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the garden-variety poor reader: The phonological-core variable-difference model. *Journal of Reading Disabilities*, *21*(10), 590–604.

Stanovich, K. E. (1991a). Discrepancy definitions of reading disabilities: Has intelligence led us astray? *Reading Research Quarterly*, *26*(1), 7–29.

Stanovich, K. E. (1991b). Word recognition: Changing perspectives. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), Handbook of reading research (Vol. 2, pp. 418–452). New York, NY: Longman.

Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognition*, *20*(1), 51–68.

Stanovich, K. E., & Cunningham, A. E. (1993). Where does knowledge come from? *Journal of Educational Psychology*, *85*(2), 11–29.

Stewart, R. A. (2009). *Idaho Reading Indicator (IRI) R-CBM scores correlated with Idaho* Standards Achievement Test (ISATS) reading scores: 3rd grade, school year 2008-2009. Boise, ID: Boise State University.

Stewart, R. A. (2010). *Read to me program evaluation report 2009-2010: Case studies of Idaho Public Library (ICFL) supported programs and analyses of Idaho Reading Indicator scores*. Retrieved from http://libraries.idaho.gov/files/Case-Studies2009-2010.pdf

Stuebing, K, Fletcher, J., & LeDeux, J. (2002). Validity of IQ-discrepancy classification of learning disabilities: A meta-analysis. *American Educational Research Journal*, *39*(2), 469.

Swanson, E., & Vaughn, S. (2011). Implementing a response to intervention model to improve reading outcomes for all students. In S.J. Samuels and A. E. Farstrup (Eds.), *What research has to say about reading instruction* (pp. 266–285). (4th ed.). Newark, DE: International Reading Association.

Taylor, B. M., Frye, B. J., & Maruyama, G. M. (1990). Time spent reading and reading growth. *American Educational Research Journal*, *27*(2), 351–362.

TerraNova™. (1996). Monteray, CA: CTB/McGraw-Hill.

Thurlow, R., & van den Broek, P. (1997). Automaticity and inference generation. *Reading and Writing Quarterly*, *13*(2), 165–184.

Torgesen, J. K. (1998). Catch them before they fall: Identification and assessment to prevent reading failure in young children. *American Educator*, *22*(1), 32–39.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of word reading efficiency*. Austin, TX: Pro-Ed.

Traub, N., & Bloom, F. (1990). Recipe for reading. Cambridge, MA: Educators Publishing Services Inc.

University of Oregon. (n.d.). *Big ideas in beginning reading: Center on Teaching and Learning.* Retrieved from http://reading.uoregon.edu/big_ideas/flu/flu_what.php

United States Census Bureau. (2015, February, 15). State & county quickfacts. *U.S. Department of Commerce*. Retrieved from http://quickfacts.census.gov/qfd/states/16000.html

U.S. Department of Education. (2000). *Twenty-second annual report to Congress on the implementation of the Individuals with Disabilities Education Act*. Washington, DC: U.S. Department of Education. Retrieved from http://www2.ed.gov/about/reports/annual/osep/2000/preface.pdf

U.S. Department of Education, Office of the Secretary, Office of Public Affairs. (2003). *No Child Left Behind: A parent's guide*. Washington, DC: U.S. Department of Education. Retrieved from http://www2.ed.gov/parents/academic/involve/nclbguideparentsguide.pdf

U.S. Department of Education (2012a, October 17). Obama administration approves Idaho's request for NCLB flexibility. *Archived Information.* Retrieved from http://www.ed.gov/news/press-releases/obama-administration-approves-idahos-request-nclb-flexibility

U.S. Department of Education. (2012b, February 29). 26 more states and D.C. seek flexibility from NCLB to drive education reforms in second round of requests.

ED.gov. Retrieved from http://www.ed.gov/news/press-releases/26-more-states-and-dc-seek-flexibility-nclb-drive-education-reforms-second-round-requests

Vadasy, P. F., Jenkins, J. R., & Pool, K. (2000). Effects of tutoring in phonological and early reading skills on students at risk for reading disabilities. *Journal of Learning Disabilities*, *33*(6), 579–590.

Valas, H. (1999). Students with learning disabilities and low-achieving students: Peer acceptance, loneliness, self-esteem, and depression. *Social Psychology of Education*, *3*(3), 173–192.

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.

Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice, 18*(3), 173–146.

Vaughn, S., Linan-Thompson, S., & Hickman, P. (2003). Response to instruction as a means of identifying students with reading/learning disabilities. *Exceptional Children*, *69*(4), 391–409.

Vellutino, F. R., Scanlon, D. M., & Lyon, G. R. (2000). Differentiating between difficult-to-remediate and readily remediated poor readers: More evidence against the IQ-achievement discrepancy definition of reading disability. *Journal of Learning Disabilities*, *33*(3), 223–238.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficult to remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognition and experimental deficits as basic cause of specific learning disability. *Journal of Educational Psychology*, *88*(4), 601–638.

Vellutino, F. R., Scanlon, D., Small, S., & Fanuele, D. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities. *Journal of Learning Disabilities, 39*(2), 157–169.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing*. Austin, TX: PRO:ED.

Walcutt, C. C. (1967). Reading-a professional definition. *The Elementary School Journal*, *67*(7), 363–365.

Wilhelm, J. (2015). Understanding reading comprehension. *Scholastic.* Retrieved from http://www.scholastic.com/teachers/article/understanding-reading-comprehension

Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., Wolf, M., Kuhn, M.,...Schwanenflugel, P. (2010). The relationship between different measures of oral reading fluency and reading comprehension in second-grade students who evidence different oral reading fluency difficulties. *Language, Speech, and Hearing Services in Schools*, *41*(3), 340–348.

Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading*, 5(3), 211–239.

Wong, B.Y.L., Graham, L., Hoskyn, M., & Berman, J. (2011). *The ABCs of learning disabilities.* Amsterdam: Elsevier/Academic Press.

Woodcock, R. W., & Johnson, M. B. (2001). *Woodcock-Johnson psycho-educational battery-revised*. Allen, TX: DLM Teaching Resources.

Yeo, S. (2009). Predicting performance on state achievement tests using curriculum-based measurements in reading: A multilevel meta-analysis. *Remedial and Special Education*, *34*(5), 259–268.

Ysseldyke, J. (2005). Assessment and decision making for students with learning disabilities: What if this is as good as it gets? *Learning Disability Quarterly*, *28*(2), 125–128.

Ysseldyke, J. E., Thurlow, M., Graden, J., Wesson, C., Algozzine, B., Deno, S. (1983). Generalizations from five years of research on assessment and decision making: The University of Minnesota Institute. *Exceptional Education Quarterly*, *4*(1), 75 –93.

Zieky, M. J., & Perie, M. (2006). A primer on setting cut scores on tests of educational achievement. *ETS: Policy & Research Reports.* Retrieved from http://www.ets.org/Media/Research/pdf/Cut_Scoes_Primer.pdf

APPENDEX A

Number of Student IRI and R-CBM Scores Included and Missing in the Study

Table A.1

Number of Student IRI and R-CBM Scores Included and Missing in the Study

Number Included	Percent	Number Missing	Percent	Total
2009 First Grade IRI Fall				
1912	97%	173	8%	2085
2010 First Grade IRI Spring				
1918	92%	167	8%	2085
2010 Second Grade IRI Fall				
1776	85%	309	15%	2085
2011 Second Grade IRI Spring				
1700	81%	385	19%	2085
2011 Third Grade IRI Fall				
1590	76%	495	24%	2085
2012 Third Grade IRI Spring				
1533	73%	552	27%	2085
2012 Fourth Grade R-CBM Fall				
1413	68%	672	32%	2085
2013 Fourth Grade R-CBM Spring				
1441	69%	644	31%	2085

Note. Only IRI and R-CBM scores of students who participated in the assessments through the years 2009 – 2013 were included in the study. Missing data refers to the excluded test scores because one or more test scores from the IRI or R-CBM.

APPENDEX B

Summary of the Pearson Chi-Square Analysis

Table B.1

Summary of the Pearson Chi-Square Analysis

Assessment	Value	Degree of Freedom	Asymptotic Significance (2-sided)
2009 First Grade IRI Fall			
Pearson Chi-Square	79.34	2	.000
2010 First Grade IRI Spring		2	.000
Pearson Chi-Square	222.07	2	.000
2010 Second Grade IRI Fall		2	.000
Pearson Chi-Square	145.77	2	.000
2011 Second Grade IRI Spring		2	.000
Pearson Chi-Square	196.70	2	.000
2011 Third Grade IRI Fall		2	.000
Pearson Chi-Square	182.04	2	.000
2012 Third Grade IRI Spring		2	.000
Pearson Chi-Square	198.91	2	.000
2012 Fourth Grade R-CBM Fall		2	.000
Pearson Chi-Square	166.77	2	.000
2013 Fourth Grade R-CBM Spring		2	.000
Pearson Chi-Square	163.45	2	.000

APPENDIX C

Summary of Symmetric Measures: Cramer's V Analysis

Table C.1

Summary of Symmetric Measures: Cramer's V Analysis

Assessment	Value	Approximate Significance
2009 First Grade IRI Fall		
Cramer's V	.204	.000
2010 First Grade IRI Spring		
Cramer's V	.340	.000
2010 Second Grade IRI Fall		
Cramer's V	.286	.000
2011 Second Grade IRI Spring		
Cramer's V	.340	.000
2011 Third Grade IRI Fall		
Cramer's V	.338	.000
2012 Third Grade IRI Spring		
Cramer's V	.360	.000
2012 Fourth Grade R-CBM Fall		
Cramer's V	.344	.000
2013 Fourth Grade R-CBM Spring		
Cramer's V	.337	.000

APPENDIX D

2009 – 2012 IRI Chi-Square Count and Expected Count Statistics

Table D.1

2009 – 2012 IRI Chi-Square Count and Expected Count Statistic

Total Students	SLD	IRI Score of 1	IRI Score of 2	IRI Score of 3
2009 Fall First Grade				
1856	56			
Count		23.0	22.0	11.0
Expected		6.2	10.9	38.9
2009 Spring First Grade				
1859	59			
Count		33.0	18.0	8.0
Expected		4.7	7.5	46.7
2010 Fall Second Grade				
1721	55			
Count		39.0	14.0	2.0
Expected		8.5	12.1	34.5
2011 Spring Second Grade				
1647	53			
Count		29.0	16.0	8.0
Expected		4.3	5.3	43.3
2011 Fall Third Grade				
1537	53			
Count		28.0	20.0	5.0
Expected		4.2	9.6	39.2
2012 Spring Third Grade				
1484	49			
Count		21.0	18.0	10.0
Expected		2.6	4.9	41.6

APPENDIX E

Odds Ratios

Table E.1

Odds Ratios

Variable	OR for SLD	OR for not SLD	Total OR
2009 First Grade IRI Fall	45/11 = 04.1	539/1317 = .41	4.1/.41 = 10.0
2010 First Grade IRI Spring	51/08 = 06.4	348/1511 = .23	6.4/.23 = 27.8
2010 Second Grade IRI Fall	53/ 2 = 26.5	610/1111 = .55	26.5/.55 = 48.2
2011 Second Grade IRI Spring	45/ 8 = 05.6	265/1382 = .19	5.6/.19 = 29.5
2011 Third Grade IRI Fall	48/ 5 = 09.6	366/1171 = .31	9.6/.31 = 31.0
2012 Third Grade IRI Spring	39/10 = 03.9	193/1291 = .15	3.9/.15 = 26.0
2012 Fourth Grade CBM Fall	40/ 5 = 08.0	354/1014 = .35	8.0/.35 = 22.9
2013 Fourth Grade CBM Spring	41/ 4 = 10.3	332/1064 = .31	10.3/.31 = 33.2