Summer 2018

An exploratory study of leadership, literacy assessments and acceleration for closing the reading achievement gap by third grade

Laura Frances Estes
College of William and Mary - School of Education, l1estes@cox.net

Follow this and additional works at: https://scholarworks.wm.edu/etd

Part of the Educational Leadership Commons

Recommended Citation
http://dx.doi.org/10.25774/w4-be4h-5289

This Dissertation is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Dissertations, Theses, and Masters Projects by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.
AN EXPLORATORY STUDY OF LEADERSHIP, LITERACY ASSESSMENTS, AND ACCELERATION FOR CLOSING THE READING ACHIEVEMENT GAP BY THIRD GRADE

By

Laura Frances Estes

Approved March 2018 by

Megan Tschannen-Moran, Ph.D.
Chairperson of the Doctoral Committee
The College of William and Mary

Thomas Ward, Ph.D.
The College of William and Mary

Denise Johnson, Ed.D.
The College of William and Mary
Dedication

I dedicate this achievement to my family and friends. There is no truer adage than the ‘it takes a village’ saying! If the miles of pavement and thousands of footsteps taken alongside my run club peeps could talk, we would all be in trouble—or just rich—from a really entertaining novel! So here goes—Run, ride, cheers to the posse—Mary Jo & Bruce and Denise & Travis who are the original run clubbers who started the whole thing; Amy & Scott, Jim, Johnny Mac, and Dave M, who I fondly call ‘the core’ that endured the epic whirlwind winter of 2017 with me—you know what I’m talking about; Debbie, Diana & David, Rachel, Jen R., Jen & Peter, Amanda & David, and Patti—all of whom rounded the gang out for an amazing group of people I am proud to call friends at this stage of life. I hope we keep the miles going forever! Of course, run club also brought Kristie into my life, and while our paths didn’t cross for a terribly long time, that journey brought me to Jim, and for that, I am forever grateful! I could never have predicted a sweeter outcome than the joy of Jim and Annalise in my life. While the dream of earning my doctorate is now realized, a whole new journey with family and friends still awaits! Dilly, dilly!
Table of Contents

Acknowledgements...........................................................................................................vii

List of Tables ..................................................................................................................... ix

List of Figures .................................................................................................................... x

Chapter 1: Introduction ......................................................................................................2

Statement of the Problem...............................................................................................6

Theoretical Framework ..................................................................................................9

   Leadership for Literacy Learning .............................................................................12

   Reading Measures ......................................................................................................15

   Standards Driven Usage Versus Personalized Usage of IRI Data .........................19

   Exploring Leadership and Literacy Assessment Practice ....................................20

Research Questions ........................................................................................................22

Significance of the Study ...............................................................................................24

Definition of Terms .......................................................................................................27

Chapter 2: Review of Literature .......................................................................................31

School Leaders and Data Informed Decision Making ................................................32

Assessment in Reading Instruction ..............................................................................37

   Standards Based Curriculum Assessment ..............................................................37

   Historical Overview of Reading Assessment ..........................................................40

   Systems of Reading Assessments ............................................................................42

   Standards of Learning ...............................................................................................43

   Informal Reading Inventories ..................................................................................46

Grade Three Reading Conundrum .................................................................................54
Acknowledgments

Just as I believe in the power and magic of a book to transform the reader, I believe there are those whose paths have crossed mine for valuable life lessons. Truly, I am forever grateful for each innumerable interaction that has shaped my path, but I do wish to acknowledge a few special, very notable, souls for which my thanks is unending.

Our parents are our first teachers who start from birth, follow us to the ends of the earth and through all our days, so I begin with Mom and Dad, of course, whose hard work and lifetimes of enduring love provided the framework for all my success. While Dad is in heaven now, I know he is celebrating with us! You said I could do anything I set my mind to, Dad, and I always believed you, but I am forever thankful for Mom’s everlasting support to help me remember that along the way! It amazes me still, how I reach and fall, but you always help me dust off and get back up! Thank you, Mom!

My siblings also amaze me! Billy, Shelley, and Randy—each with their own fascinating and successful journeys in life, have taught me many lessons. I might always be the baby of the family, but I’ve benefitted from your lives in so many ways. Thank you!

There are always those special teachers that reach us in a distinctive way, so I must thank Gwen Bennett. While she has proceeded us in this life, her mark is with us always. She was one of the most impactful educators in my life who taught me to be true to myself, a lesson not found in a textbook or curriculum. Thank you Gwen!

I swear just about everything I know about literacy came from Sue Duncan, and thus, Marie Clay! Reading Recovery started my journey in literacy teaching and learning, and I’ll be forever thankful for that impact!
Then there are those who teach us informally along the way—Ann Colorado, Micah Smith, and those in the cohort, all of whom were instrumental in helping me make it to the end! Thank you! Of course, my committee members Dr. Tschannen-Moran, Dr. Ward and Dr. Johnson, who shared their experience and expertise with me so generously, deserve a lot of credit and an award for patience! I am so thankful you were each on my team! Each of your specialty areas contributed so richly to the final product. I am forever grateful for the academic discourse and dissonance that challenged my thinking and helped me grow in literacy, research, and leadership. Again, thank you all!
List of Tables

Table 1. Commonly Used Informal Reading Inventories ........................................ 48
Table 2. Matrix of Extant Correlational Literature .............................................. 75
Table 3. Summary of Basic Procedures for Administering a Rigby IRI ................. 90
Table 4. Rigby Informal Reading Inventory Levels and Grade Level Equivalences .... 91
Table 5. Survey Protocol Design Matrix ........................................................... 96
Table 6. Research Questions, Theoretical Constructs, Data Sources, and Statistical Analyses ................................................................. 101
Table 7. Descriptive Statistics on Rigby IRI and SOL Scores for Three Variables .... 106
Table 8. Rigby Quarter 1, Quarter 4 and Third Grade Reading SOL Pearson r Correlations ................................................................. 107
Table 9. Cross Tabulation of SOL categorization and Rigby IRI Reading Level .... 110
Table 10. Averages on School Leader Use of Reading Measures Survey ............. 115
Table 11. Results, Findings, and Implications ................................................... 144
List of Figures

Figure 1. Third Grade Reading Conundrum.........................................................7

Figure 2. Theoretical Background for the Study..............................................11

Figure 3. Instructional Level Expectations for Reading..................................59
Abstract

The third grade year is a seminal moment for children moving into a fluent reader stage that continues to evolve well across their school career. Research indicates that not learning to read well by the third grade sets some children on a path of overall diminished school and life success, and thus, school leaders are faced with the challenge of altering the trajectory for students behind their peers in reading development. School leaders and teachers have a limited number of tools to assess literacy progress of beginning readers; therefore, it is important educators understand the connections between two of the most commonly used assessments. This exploratory study investigated the correlations between an Informal Reading Inventory (IRI), specifically, the Rigby, and the Virginia third grade Standards of Learning (SOL) assessment as well as a survey to capture principals’ self-ratings around the use of IRI quantitative and qualitative information within a data-informed instructional decision-making model. Findings indicated a significant relationship between the Rigby IRI and the Virginia third grade reading SOL as well as a significant relationship between where students scored on the quarter 1 Rigby IRI and where they end on both the quarter 4 Rigby IRI and the SOL. Moreover, principals’ perceptions of their leadership skills and processes indicated a stronger knowledge base and use of quantitative data from IRIs within a data monitoring system and a benchmarking process rather than the qualitative personalized instructional use of data from an IRI. If acceleration in reading progress is to be achieved in order to close reading gaps, IRI qualitative data needs to be utilized for a more dynamic instructional approach. Recommendations for practice and future studies are offered.
AN EXPLORATORY STUDY OF LEADERSHIP, LITERACY ASSESSMENTS, AND ACCELERATION FOR CLOSING THE READING ACHIEVEMENT GAP BY THIRD GRADE
Chapter 1: Introduction

Some researchers have identified principals’ use of data to lead schools as the most salient means for substantial impact on student outcomes (Fullan, Hill, & Crevola, 2006; Kowalski, Lasley & Mahoney, 2008; Lai & Hsiao, 2014; Schildkamp, Karbautzki, & Vanhoof, 2013; Slavin, Cheung, Holmes, Madden, & Chamberlain, 2013). Leadership models, over the years, have varied in their conceptual frameworks for emphasizing the role and activities of school leaders, but currently school leadership as an influence on student learning is being heralded by some researchers as second only to classroom teaching (Leithwood, Harris, & Hopkins, 2008). Experts in the field of leadership seek to merge data use from an isolated component within a leadership model toward a model whereby data use systems underpin all realms of leadership practice (Leithwood & Jantzi, 2006; Louis, Leithwood, Wahlstrom, & Anderson, 2010; Sun, 2014; Sun, Johnson, & Przybylski, 2016). One strand within the data-driven school leadership (DDSL) model specifically features the school leader’s work towards developing teachers’ decision-making capacity by encouraging teachers to provide precise instruction to each student based on student data (Farley-Ripple & Buttram, 2014; Jimerson, 2013; Lachat & Smith, 2005; Sun et al., 2016; Wayman & Cho, 2008). The use of reading data and the modeling of data use by school leaders, specifically, comes into particular play within this realm of data-driven school leadership for a variety of critical reasons.
Reading is arguably the most crucial skill that students learn in school. In this information age, reading is an important vehicle through which people access the information needed to navigate their daily lives. If children do not master the art and skill of reading in their early years of schooling, when brain development favors language acquisition, they are likely to struggle throughout their school years and through their lives. School leaders and teachers have a limited number of tools to assess the reading skills of beginning readers. It is important that educators understand the purposes, capabilities and connections between two of the most commonly used assessments, an Informal Reading Inventory and the Virginia Standards of Learning Grade 3 reading assessment.

Third grade reading data can be so predictive of future life outcomes that data from the third grade is used well beyond school administrators and teachers in the elementary years and has proven to be predictive of other outcomes in life overall. This mantra is repeated by the works of many in the field of education who advocate for urgent changes to the work schools do with literacy (Askew, 2011; Booker, Invernizzi, & McCormick, 2007; Christie & Rose, 2012; Education Commission of the States, 2002; Hernandez, 2011; Joshi et al., 2009; National Center for Education Statistics, 2011; Ravitch, 2009; Siegrest & Van Patten, 2007). To underscore the implications of the third grade reading outcomes shown in the data and the lifelong impact on individuals in society, Shippen, Houchins, Crites, Derzis, and Patterson (2010) found standardized reading measure outcomes for a wide variety of demographic profiles indicated a difference of one to two standard deviations below the mean in reading performance for
prison inmates as compared to their non-incarcerated peers, painting a grim outcome for some who enter into adulthood as poor readers. According to the U.S. Department of Education (2016), the National Assessment of Education Progress (NAEP) data for fourth grade reading shows little change over the years in reading for the nation with scores holding at 36% at or above proficient. The data clearly show the persistent urgency for school leaders and educators to continue to work towards solutions for interrupting the trajectory of lagging progress in early literacy development for children behind their normally progressing peers in the primary grades. This holds true for Virginia as well. While they fared somewhat better than the national average with the last reported data from 2015 as 47% at or above the proficient level, it is unacceptable for less than half of students to achieve proficiency in reading in the primary grades.

Students not reading on or above grade level with proficiency by third grade will likely continue to struggle, and according to Morris’s (2004) as well as Cramer’s (2010) research, will most likely not pass the Virginia Reading SOL. Additional statistics show students with various demographic backgrounds, such as low socioeconomic factors, disability factors, and limited English proficiency factors when added to low reading acquisition skills, sets some students on a path of failure, dropping out of school, and an overall lower lifetime of earnings than those students who complete high school or go onto college with proficient reading skills (Booker et al., 2007; Christie & Rose, 2012; Education Commission of the States, 2002; Hernandez, 2011; Joshi et al., 2009; Ravitch, 2009; Siegrest & Van Patten, 2007; National Center for Education Statistics, 2011).
Despite numerous reform efforts, consider, that “from 1970 to 2011, the average reading scores for nine year-olds remained relatively flat, increasing by just 12 points on a 500-point scale” as shown on NAEP results (Christie & Rose, 2012, p. 4). The National Reading Panel Report in 2000, and the implementation of the reauthorization of the Elementary and Secondary Education Act as the No Child Left Behind Act in 2001, were both seminal events in U.S. reading reform; however, the NAEP data indicate few gains overall (Education Commission of the States, 2002). The data, as cited by Hernandez (2011), show one third of U.S. students still reading below basic level as indicated by the NAEP, and another one third reaching only a basic level, with proficiency eluding two thirds of all rising fourth grade students. Part of the seminal outreach of NCLB legislation included the introduction of policies that emphasized assessment and accountability, which evolved into mandated testing in grades three through eight and once in high school, along with a large focus on the core areas of reading and math. Furthermore, imposed sanctions emerged for schools missing external benchmarks set under NCLB for not making what was deemed as adequate yearly progress (Virginia Department of Education SOL Innovation Committee Meeting, 2016). However, the NCLB legislation not only produced new policies on mandatory assessments for reading, but also produced new research in the form of The National Reading Panel Report (Shanahan, 2006). The final report, submitted on April 13, 2000, to the U.S. Department of Education, served as the framework of President George W. Bush’s plan for improving education in the NCLB Act of 2011. The report indicated that research supported five essential elements of early literacy instruction: phonemic
awareness, phonics, and oral reading fluency, vocabulary, and comprehension strategies. These elements became the cornerstone of current instructional practices and assessments in the educational setting. The main assessments used in the state of Virginia pertain directly or indirectly to the five essential elements in various ways. Informal Reading Inventories (IRIs) can assess fluency, use of vocabulary, application of phonics and phonemic awareness, as well as comprehension. The SOL test can assess comprehension and overall reading achievement. Score reporting categories include word analysis and reference materials, and comprehension of fiction and non-fiction texts.

**Statement of the Problem**

While the field of school leadership has evolved to show the importance and impact of leadership within the organization, as well as shown data systems are beneficial to the work school leaders do each day, school leaders need to know more about the potential impact of the use of the two predominantly used forms of reading assessments in Virginia to better inform their instructional decision making.

A combined picture of the school leader’s use of data to inform decision-making along with assessment practices over time in reading instruction frame the component parts of the reading conundrum we face in education when considering the critical timeframe of early literacy acquisition from kindergarten to grade three. Figure 1 depicts a visual representation of the overlapping nature of the three areas of school leadership, standards based assessments and informal reading assessments, which are typically treated as separate entities, but for the purpose of this study were explored for their interconnected properties to inform practice in the field of leadership.
Figure 1. This figure shows the overlap of research and practice for the conceptual components of the study.
This study explored the nature of the correlation between one type of commercially produced informal reading inventory (IRI) and the state standardized reading test along with elementary school principals’ practices in organizational structures for data usage and dynamic instructional practices to frame instructional reading goals. By studying the correlational nature of the Rigby IRI to obtain students’ estimated reading levels equated with the pass/fail categorizations on the SOL, new information could shed light on trends that will help school leaders retool and reshape the information they already collect and use in more effective ways. Further, surveying the school leaders regarding specific practices that fall under the qualitative informational side in the use of what IRIs have to offer help paint a clearer picture about potential data usage practices to promote instructional practices that accelerate students’ reading achievement. This study provides the reading level ranges from one commercially produced IRI, the Rigby, within particular testing windows on multiple cohorts of third grade readers in a mid-sized suburban district in Virginia. Patterns revealed in relation to the SOL assessment as suggested by the Joint Legislative Audit and Review Commission report (2011) could raise awareness of potential misalignment in using only the developmentally assigned benchmarking levels from IRIs compared with the reading levels students actually tend to fall in by pass/fail categorizations on the VA SOL. This is critical information needed to confirm or reshape the trajectory of expectations in reading levels, if school leaders are to implement structures for using data to truly accelerate learning and reduce the number of students failing to learn to read successfully by third grade.
Our goal is to reduce the number of students failing to read on grade level by the end of third grade, which is considered the watershed year of the elementary grades (Booker et al., 2007; Christie & Rose, 2012; Education Commission of the States, 2002; Hernandez, 2011; Joshi et al., 2009; Ravitch, 2009; Siegrest & Van Patten, 2007; National Center for Education Statistics, 2011). As much as educators have tried to raise the number of children reading on grade level by third grade in the U.S., over 40 years of NAEP results show few gains in overall reading achievement (Christie & Rose, 2012; U.S. Department of Education, 2016).

**Theoretical Framework**

Figure 2 provides a visual representation of the theoretical background for the study. Part of what principals do to set high expectations is through their understanding of teacher practice via observations, either formal or informal walk-throughs, and subsequent conversations, coaching through data monitoring meetings or systems, and asking questions to guide teacher growth and practice (Coburn & Turner, 2011; Deike, 2009; Farley-Ripple & Buttram, 2014; Jimerson, 2013; Lachat & Smith, 2005; Simpson, 2011; Wayman & Cho, 2008). The types of questions and structures for dialogue around literacy instruction can foster either a dynamic or fixed instructional practice mindset with teachers. Understanding why school leaders may or may not delve further into questioning or structures that get closer to dynamic instruction offers insight into why we see the continuation of stagnating reading data overall. As shown in Figure 2, the instructional leader can set the tone for practices that bridge the gap between the stated
goal for accelerated reading outcomes and actualized outcomes for more students reaching grade level expectations.
Figure 2. This figure shows school leaders and the possible paths for data-driven decision making for third grade reading outcomes through the use of differing leadership practices as explored in this study.
Leadership for literacy learning. Principals in schools today are expected to be instructional leaders, not just managers of schools (Darling-Hammond, LaPointe, Meyerson, Orr, & Cohen, 2007). Part of the role as an instructional leader encompasses setting high expectations within a school setting. One way an elementary school principal can assume an instructional leadership role is through the use of data monitoring systems and coaching/feedback loops that are part of a school-wide literacy model which stresses dynamic instructional practices for literacy instruction. Dynamic instructional practices include differentiated or personalized learning experiences and regrouping or flexible groups centered on student literacy profiles.

Flexible groups in early literacy instruction means that students are assessed for their developmental reading level proficiency and then placed in small groups of three to five students where they are instructed through literacy practices that align with the coordinating stage of reading development for that level reader. Students respond to this type of personalized instruction in varying rates, and therefore, make varying rates of progress in reading levels. Once student levels in a group differ enough to cause misalignment in goals from the initial grouping, the instruction will be less personalized. New groups need to be formed to realign the stage of readers for appropriate instructional goals. This is considered flexible grouping where students in the group do not just stay the same for long periods of time. Teachers may find they need to change students in a group due to other reading stage characteristics, such as fluency, comprehension, problem-solving in text, and so on. It is considered a more personalized approach to literacy instruction and creates the conditions for accelerating learning progress as
students are moved frequently based on growth, and where student literacy profiles drive instructional planning.

A fixed instructional practice includes an emphasis on static grouping in literacy instruction at the elementary grades and instruction is centered on standards for learning or curricular pacing and are not necessarily tied to student literacy profiles. A fixed approach includes grouping students based on an initial assessment of reading level, and generally leaves students in the group for an undetermined amount of time, if not the entire year. It also centers on a standards-based learning environment where the driving force for instructional goals stems from the curricular standards for the grade level. The curriculum may emphasize some subsets of reading skills that are important to developing readers in the elementary grades, but it is incumbent for the teacher to align the instructional practices for the proper stage of developing readers along with teaching the curricular standards for the grade level. As one literacy expert put it, the curricular goals and a personalized literacy method need not be at odds with one another, but rather need to be interconnected. This notion is best summed up as stated by Routman (2014): “Although standards and curriculum guide our instruction, what and how we teach must be interconnected to the responsive practices that lead to high student engagement, achievement, and independence as learners” (p. 38).

The pace in a fixed instructional practice tends to follow a curricular pacing guide, and benchmarking goals for reading levels that are approximate levels of movement across a grade level. While many students may follow a traditional trajectory of reading level progress across a grade year, others will move at widely varying rates or
have needs that differ greatly from the curricular pace. The fixed instructional practices
tend to group students by levels, monitor for students making the set benchmarked
reading levels at intervals across the year, and leaves students within those original
groups largely unchanged across a year. This leads to a rigid approach emphasizing a set
of subskills instead of a reading process. In contrast, an approach that focuses on the
learner first and the content second makes it more likely students will reach reading
goals. Accelerating the pace of learning with purpose and urgency is needed to close
reading achievement gaps (Routman, 2014).

School leaders modeling the use of data and providing individual support for data
use in a one-to-one teacher setting can create the emphasis towards either a dynamic or
fixed instructional mindset for teachers in the school setting within their literacy
framework. This is achieved through the data monitoring systems that principals
institute, facilitate, and utilize to guide instruction (Fullan et al., 2006; Kowalski et al.,
2008; Lai & Hsiao, 2014; Schildkamp et al., 2013; Slavin et al., 2013), and
coaching/feedback loops that stress dynamic instructional practices for literacy
instruction (Coburn & Turner, 2011; Wayman, Spring, Lemke & Lehr, 2012). With a
limited set of reading assessment tools for the acquisition phase of literacy learning in the
early grades, the use of two predominantly used measures can add to major implications
for school leaders making data-informed decisions for school goals, professional
development, teacher coaching, interventions, staffing resources, and material purchases
(Crum, Sherman, & Myran, 2009; Deike, 2009; Kennedy & Datnow, 2011; Mandinach,
Reading measures. There are two predominantly used measures of student progress in reading at the elementary level in the state of Virginia. These include the Virginia State Standards of Learning (SOL) Reading assessment for grades three through five and some form of an Informal Reading Inventory (IRI) to monitor reading development kindergarten through fifth grade. The VDOE considers the SOL results as a method for informing parents about not only their individual child’s achievement, but also collective results of students within communities as to the degree of progress in meeting the commonwealth’s expectations in each content area tested. The state Board of Education uses the information to identify schools that are in need of assistance and support as well as using the assessments as an “objective means for measuring achievement gaps between student subgroups and for determining the progress of schools, districts, and the state toward closing these gaps” (Virginia Department of Education, 2015).

Part of the assessment structures schools use includes an IRI, but in addition to the close monitoring of graded reading passage progress through the IRI process, statewide focus on passing the outcomes-based reading SOL exams adds to the pressing nature for appropriate reading growth over time. The charge for school leaders to use reading data as effectively as possible between kindergarten and third grade, if schools are to truly reduce reading failure rates, is pressing. The Virginia State Department of Education monitors student outcomes through criterion referenced content and skills
assessments administered yearly, and these tests are considered high stakes tests for
districts and schools across the state. Research indicates the use of IRI information does
provide teachers with the appropriate range of instructional reading level needed for
optimal gains in reading acquisition in the early grades (Allington, McCuiston, & Billen,
2015).

Informal reading inventories usually include oral reading accuracy (Clay, 1993)
and some form of miscue analysis (Goodman & Burke, 1972). Additionally, IRIs contain
a form of graded word lists to help place children within graded reading passages,
followed up by a series of comprehension questions or retelling rubrics. It contains both
quantitative and qualitative information about reading behaviors of children. Even
though IRIs have multiple benefits, Paris and Carpenter (2003) reported that educators
appear to have institutionalized a narrow interpretation, or fixed instructional practices,
for the use of IRIs—to obtain quarterly benchmarks across the year. They also note there
are several other more dynamic intended uses to include determining reading levels and
placing students in reading materials, informing grouping practices, and monitoring
progress over time. One of the most powerful intentions of the assessment tool is the
least used and that is the intended use for determining specific reading skill strengths and
needs. By looking deeply at the reading behaviors observed and analyzing them for
instructional teaching points, not just finding reading levels, teaching goals can become
more powerful and specific to meet student reading needs. Schools over-emphasizing the
quantitative information obtained in an IRI over the vast qualitative information IRIs can
provide risk a more fixed approach.
The fixed practices and systems that use an IRI assessment as a benchmarking tool that is framed around a standards-based curriculum potentially hinders the concept of accelerated movement intended to close critical reading gaps by third grade and falls under the fixed instructional method umbrella. Schools often use an IRI to obtain a quantitative reading level, frequently without the use of the additional qualitative information available on an IRI. This additional qualitative information, that is less frequently utilized, can be used to accelerate learning through differentiated or personalized instruction and frequent regrouping practices, which is more dynamic in nature. Even though studies have shown IRIs can have modest to significant predictive value, schools could improve the instructional use of their added value, besides just benchmarking, by using the qualitative information to personalize feedback to students, and frequently regroup students for accelerated reading progress. Benchmarking in the primary years involves the use of a widely accepted leveling system that indicates the reading levels students should be working on across the quarters of the school year within the early grades.

As students obtain a certain level of proficiency by the third grade, they switch from learning how to read to more of a reading to learn stance. Promising instructional methods employing differentiated instruction using informal reading records and the use of frequent regrouping to meet the instructional needs of all students has shown student reading levels can be significantly accelerated, closing early reading gaps starting as early as the kindergarten year (Duncan, 2016; Duncan & Johnson, n.d.a; Phillips, McNaughton, & MacDonald, 1997).
A study that focused on strategies to promote third grade reading performance in Virginia by the Commonwealth of Virginia Joint Legislative Audit and Review Commission Report (Joint Legislative Audit and Review Commission, 2011) claimed that pass and pass advanced levels for the SOL “appear to reasonably approximate grade level performance” (p. 25), and that students who pass the third grade reading SOL test should be “where they need to be” to enter fourth grade in reading (p. 25).

The Commonwealth of Virginia JLARC (2011) study also acknowledged that while districts across the state use IRIs as the major determination for student reading acquisition progress, there is no standardized method for collecting and tracking IRI progress for schools across the state. However, a noteworthy indication from the state report includes that the IRI is the “preferred method for determining reading grade level” over a standardized question and answer format, such as the SOL, and the report further suggested that aggregated data from IRIs could provide the number of students in VA reading on grade level, if done within a particular testing window (Joint Legislative Audit and Review Commission Report, 2011, p. 25).

School leaders need a better understanding of the correlations between the use of an IRI and the state’s third grade reading test as well as the uses of an IRI for instructional decision-making. Although school leaders believe the use of an IRI is an appropriate reading measure tool, the organizational structure in the implementation and use of them may hinder the goal for increasing the number of students successfully reaching third grade reading level on-time, especially if leadership leans towards fixed instructional practices. There are potential conflicts between the practices school leaders
employ through organizational structures with using IRI data and the stated goals toward closing reading gaps in the early years. The overemphasis on obtaining just the numerical levels and not utilizing the qualitative information contributes to the carryover of an inflexible curricular program that moves students lock-step through materials based on grade level curricula (Paris & Carpenter, 2003; Routman, 2014). Informal reading level texts provide a gradient of reading difficulty moving up a ladder of leveling that allows teachers to pinpoint an exact level of instruction for the most proficient outcomes in teaching. Prior to the use of a leveled set of readers for assessment in determining the place where instruction best suited a young emerging reader, a basal series was used that provided a fixed set of learning skills paced across a school year focused primarily around a code emphasis approach, such as controlled vocabulary and phonics patterns, or a meaning emphasis approach. Even though leveled readers in IRIs are more appropriately graded than in the years of basal readers, the way the results are implemented may prevent students from being moved through instruction at an accelerated pace over a fixed curricular pacing method. Furthermore, basals did not allow for accelerated instruction since students were locked into the curriculum for a full year whereas, an IRI is an assessment that allows students to be placed in an appropriate level of instruction and dynamic instruction keeps movement in the curriculum flexible.

**Standards driven usage versus personalized usage of IRI data.** A rigid approach to benchmarking students at a quarterly level and only considering the numerical or alphabetical leveling system is a threat to accelerated learning goals that
schools desire for closing reading gaps among early grade students, which is deemed essential for lifelong success in and out of school.

Personalized instruction through the use of observations obtained on informal reading inventories would stem from qualitative information that is then available for use in planning specifically tailored instruction in small groups of three to five students for reading in the early grades. This type of information includes a student’s strengths and needs in their understanding of how print works in the very early grades, use of print cues in the form of meaning from text, structural cues in language, and visual cues within printed text. Further, information can be garnered regarding phrasing, fluency, and comprehension. These types of information notated in a qualitative form provide more than just a reading level; they provide clues as to what is cutting edge learning for each student. This allows instruction to be dynamic in meeting student needs to propel them forward while continuously working up a ladder of reading levels in the direction of grade level proficiency or higher. Using this information creates an opportunity to group and regroup students in alignment with continuously moving growth and needs in the primary years. These characteristics form the bedrock of an accelerated model of teaching and learning that is central to the goal of increasing the number of children reaching grade level expectations by third grade.

**Exploring leadership and literacy assessment practices.** Many factors go into the work school leaders do, but that work can and does contribute to improved school gains and the use of reading data in the early grades is part and parcel to that work (American Educational Research Association, American Psychological Association &
National Council on Measurement in Education, 2014; Bernhardt, 2003; Hallinger & Heck, 1998; Halverson, Grigg, Prichett, & Thomas et al., 2007; Lashaway, 2002; Marzano, 2002; Marzano, McNulty, & Waters, 2005; National Policy Board for Educational Administration, 2015; Reeves, 2002; Taylor, 2010). Experts in the field suggest “the capacity of school leaders and teachers to transform traditional schools into organizations able to respond to the feedback of standardized testing represents a significant step in our understanding of the next generation of school leadership practice” (Halverson et al., 2007, p. 5). Surveying school leaders on the level of implementation of practices that foster systems for using reading data, personalized instruction and frequent regrouping practices provides valuable information regarding leadership practices that support acceleration and reducing reading gaps.

As part of the impetus for change in American schools with the historical landscape of reform efforts, political rhetoric, and policy changes for increasing achievement scores, accountability systems have evolved into part of the daily disciplines of school leaders in helping to inform their decision-making about daily schooling in ways unseen previously (Halverson et al., 2007; Reeves, 2002). Even with schools taking steps to create assessment systems and practices, commonly in the form of purchased commercial testing packages, software databases, test banks, test prep materials, program evaluations, and such, it is only with significant organizational change, buy-in, and true implementation practices that deep structures will root themselves, and this is, generally, only achieved through effective leadership at the helm. Thus, it is not a matter of implementing accountability systems with data use, but rather,
retooling the kinds of data schools have and how they can use the data they have already, such as the IRI information (Halverson et al., 2007; Hoy & Miskel, 2008).

Clearly, it is unwise and unacceptable to continue with assessment practices that, while well-meaning, do not quite shift student learning to actually close reading gaps. By taking an in-depth look into the data presented in the student outcomes on the IRI and the SOL, along with a survey of school leader practices for fostering data usage at the qualitative level, beyond the quantitative level, revelations about effective school leader decision-making, organizational structures around data usage, and ultimately, teacher practice, may shed additional light on the third grade reading conundrum of little to no shifts in overall achievement by the end of third grade.

**Research Questions**

This study proposed the following research questions from a sample of third grade cohorts across several years in a mid-sized suburban district in the state of Virginia.

1. How are two main types of reading measures used in the state of Virginia correlated?
   a. To what extent are the Rigby Informal Reading Inventory (IRI) fourth quarter scores and state Standards of Learning (SOL) Reading scores correlated for a sample of third grade students in a midsized suburban district?
   b. What is the range of Rigby IRI scores of students who passed the SOL at the proficient level? At the advanced level? At the fail level?

2. To what degree do outcomes on the Virginia third grade reading SOL show evidence of student acceleration towards closing a reading achievement gap?
a. What are the proportions of change in categorizations of student reading by grade level (below, on, above) as measured by the Rigby IRI from beginning to end of the third grade year for the five years of cohorts sampled?

b. To what degree do the number of reading levels that students move, as assessed by the Rigby IRI, differ among students classified as fail, pass, or pass advanced on end-of-year VA SOL Reading tests across the school year for a sample of third grade students in a mid-sized suburban district?

3. Do school leader data-driven practices reflect structures within a literacy framework that foster dynamic instructional practices?

   a. To what degree do school leaders identify practices within a data monitoring system in the district as fostering the use of the quantitative and/or qualitative information provided in an IRI?

   b. To what degree do school leaders identify their leadership structures as fostering the use of the fixed instructional practices through quantitative leveling information (benchmarking) provided by an IRI?

   c. To what degree do school leaders identify their leadership structures as fostering the use of dynamic instructional practices through personalization of reading instruction using qualitative information from an IRI?

   d. To what degree do school leaders identify their leadership structures as fostering the use of dynamic instructional practices through regrouping or flexible grouping of students for reading instruction?
Significance of the Study

A few emerging studies have been conducted on the correlational nature of high stakes testing outcomes against informal reading inventories to include three forms of some of the more predominantly used ones in the field. These include the Scholastic Reading Inventory, the Developmental Reading Inventory, and the Qualitative Reading Inventory. This proposed study sought to replicate the correlational research design from other forms of an IRI through the correlational research regarding the Rigby (Smith, Nelley, & Croft, 2008) published version of an IRI. It then sought to go beyond those previously studied to establish a framework for the full scope of categorization on the SOL for fail, pass, or advanced against the performance of five years of third grade cohort’s reported Rigby IRI levels for a more complete picture of the patterns in the data. Additionally, it provided a closer look on how an IRI shows what, if any, movement exists in students being categorized as on, below, or above grade level status by IRI level across the third grade school year. It examined trends across five years of cohorts, unlike other studies completed on a single year cohort, and lastly, the proposed study differs because it surveyed principals’ practices on organizational structures in data usage using the information provided by the IRI. Looking deeper into not only the use of the fixed instructional practices using quantitative leveling and benchmarking information obtained on IRIs, it considered important questions regarding the practices that may foster dynamic qualitative information available on IRIs that has more value in shaping personalized instruction and grouping practices towards a more accelerative model. This adds valuable information that goes beyond the predictive values, and helps to further
establish a scope of all the categorizations on an IRI, such as the Rigby, against the
categorizations on the SOL, as well as provided a synopsis of the level movement during
the third grade year and how that translates to outcomes on the SOL comprehension test.

The goal was to use theoretical constructs and extant literature from three major
areas and overlay them in an exploratory study to investigate the potential to better
inform school leaders with using reading data in a more impactful manner. While these
major areas exist in school practice, and a great deal can be found in the literature and
research about them in isolation, this study sought to overlay the constructs to examine
each area as an interlocking set, and explored the potential outcomes from viewing them
in this interconnected manner. Some emergent research has begun to shed initial light
onto correlations, and information exists for the areas—leadership, standards based
assessments, and informal reading inventories—separately. While they are each
inextricable components of a school’s literacy system, research has not fully expanded,
yet, to include the nature to which is presented in this research study. Data on the
overlapping nature of informal reading inventories and standardized assessments suggests
an opportunity exists to reconsider the type of early reading data we have, how we use it,
and how it can inform instructional decision making. The ways in which school leaders
guide the implementation and use of reading assessments can be extremely impactful on
student outcomes, depending on school leaders’ application of the types and uses of
reading assessments. These concepts will be further explored in Chapter 2 and addressed
through the research questions in this study.
Given that the leadership models of the 21st Century are predicated upon the use of data acquisition for analysis, and ultimately, program alignment and design, the question evolves into whether school leaders in Virginia are adequately informed about the potential relationships of these assessments administered in reading at the elementary level. I posit it is prudent for school leaders to understand more about the correlational implications of the assessments to one another and as a whole. The research on school leadership and data informed instructional systems suggest a positive impact is possible using the identified leadership strategies and data processes, but little research exists about the probable correlations of the actual data we collect in schools regularly. Even with the use of an IRI assessment system in the state of Virginia, school systems still struggle to close the gap for some students failing to reach expected reading levels by the end of third grade.

This study adds to the correlational studies previously conducted in the form of the Rigby published IRI—one that has not yet been included in a correlational research design on reading assessments in literacy. It also provides an in-depth look into the actual reading level equivalency trends by student performance as categorized on the Virginia State SOL assessments for pass, fail, and pass advanced. A fundamental component of the study includes a close look into Virginia school leaders’ use of two predominantly used reading data measures.
Definition of Terms

*Balanced Literacy Model* – an instructional approach in the early elementary grades that encompasses a balance between instructional components for phonics, text reading, oral reading, and writing.

*Basal Readers* – a compilation of bound stories with vocabulary controlled texts for use in early literacy instruction characterized as less authentic than little books more commonly used today.

*Commercial Reading Inventories* – examples include the following thorough, but not exhaustive listing: Analytical Reading Inventory, Bader Reading and Language Inventory, Basic Reading Inventory, Burns and Roe Informal Inventory, Classroom Reading Inventory, Comprehensive Reading Inventory, Critical Reading Inventory: Assessing Students’ Reading and Thinking, Dr. Fry’s Informal Reading Assessments Grades K-8, Ekwall/Shanker Reading Inventory, Flynt-Cooter Reading Inventory for the Classroom, Informal Reading - Thinking Inventory, Qualitative Reading Inventory, Stieglitz Informal Reading Inventory, Texas Primary Reading Inventory, and 3-Minute Reading Assessments: Word Recognition, Fluency & Comprehension.

*Data informed decision making* – a part of leadership methodology that utilizes multiple data sources of student achievement as well as other school archival data to evaluate, analyze, and determine curricular needs, school programs, interventions, professional development, resources, and other functions of daily school operation (Halverson et al., 2007).
**Personalized Instructional Framework** – a model of instruction that de-emphasizes the use of whole class instruction with a pre-paced delivery of curriculum for a particular grade level. The focus is shifted toward identified individual strengths and needs of students and tailoring instruction in small group form to meet student instructional needs.

**Informal Reading Inventory** - “is an informal testing instrument which consists of graded reading passages which are used to determine a student’s reading level. Each passage is to be read orally or silently by the student who attempts to answer accompanying comprehension questions” (Russell, 2013, p. 5).

**NAEP Basic** – One of three NAEP achievement levels denoting partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade assessed (National Center for Education Statistics, 2011).

**NAEP Proficient** – One of three NAEP achievement levels representing solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter (National Center for Education Statistics, 2011).

**NAEP Advanced** – One of the three NAEP achievement levels denoting superior performance (National Center for Education Statistics, 2011).

**Phonics Based Instruction** – an approach to reading instruction emphasizing a systematic curriculum for the decoding process by understanding how letters are linked to
sounds and spelling patterns. This approach typically uses a series of controlled phonics readers to introduce letter sound correspondence in a systematic form.

*Standards Based Instruction* – an approach to education typically used in the United States as a system of instruction for setting minimum competencies and expectations of grade levels or content areas for student learning. Instruction centers around short and long term curriculum mapping for course or grade level instruction to ensure students are exposed to and learn the content and skills for a given grade level or content area.

*Standards of Learning* (SOL) – establish minimum expectations in Virginia public schools for what students should know and be able to do at the end of each grade or course as well as “measure the success of students in meeting the Board of Education’s expectations for learning and achievement” (Virginia Department of Education, 2014).

*Standards of Learning (SOL) Fail/basic* – achieving a scaled score of 399 or below

*Standards of Learning (SOL) Pass/Proficient* – achieving a scaled score of 400 to 499

*Standards of Learning (SOL) Pass/Advanced* – achieving a scaled score 500 or above

*Stanford Achievement Test, Tenth Edition (Stanford 10)* – An instrument designed to measure student achievement in multiple areas. Although the Stanford 10 is designed to measure listening, science, social science, and mathematics, this study utilizes the portions pertaining to reading achievement, which include general reading, spelling and language (Carney, 2005).
Whole Language – an approach to reading instruction that emphasizes children learn to read by recognizing words as whole pieces of language and de-emphasizes the decoding of language and words into letters or letter combinations.
Chapter 2: Literature Review

As highlighted in Chapter 1, educators of children in the primary grades between kindergarten and grade three are faced with the task of attempting to accelerate the pace of learning during the reading acquisition phase for those who arrive in school performing behind their average peers in literacy knowledge and experience. In order to better understand the nature of this third grade reading dilemma school leaders face in day-to-day decision-making and the predominant reading measures used in schools in Virginia, this Chapter provides a road map through the major constructs reviewed in the literature for the purpose of this study. Major areas examined include: a) the role of school leaders in instructional decision making, b) the role of assessment in reading instruction, and c) the reading conundrum school leaders face attempting to reduce the overall failure rates and increase the number of students reaching proficiency. These major constructs, lead to the final segment exploring how educators might better understand the measures we use in reading to inform instructional decision-making.
School Leaders and Data Informed Instructional Decision-Making

Planning for and implementing instruction is a complex and involved process. What we know about schools from systems-theory, is that no one part of the system stands alone, and an instructional leader can play a significant role in the instructional processes of a school (Senge et al., 2000). Leaders help set the focus, articulate the vision and mission of a school, and create academic press. Through this process, leaders have a great deal of impact on the types of instructional strategies utilized in classrooms. A growing body of evidence over the last several decades has solidified that principals matter in schools (Reeves, 2009; Sahlberg, 2013). Their work contributes to student learning, overall school improvement goals, and toward achieving instructional excellence. Increasingly, research studies have shown the strong influence school leaders can have on teacher retention and recruitment (Darling-Hammond et al., 2007), and significant, while indirect, effects on student achievement outcomes (Louis et al., 2010; Witziers, Bosker, & Kruger, 2003). Additionally, school leaders have the ability to influence policy implementation (Bryk, Sebring & Allensworth., 2010; Preston, Goldring, Guthrie & Ramsey, 2012). These are significant ways in which school leaders can and do contribute to outcomes in schools.

The national standards for educational leaders, developed by the National Policy Board for Educational Administration, require that principals have skills and knowledge of information sources, data collection, and data analysis (National Policy Board for Educational Administration, 2015). The standards state that school leaders are accountable for using data and for student achievement. Mounting political pressure for
improving American schools, the former legislature for accountability in No Child Left Behind (Education Commission of the States, 2002), and the standards for educational leaders created a catalyst for principals to become proficient in the use of student performance data. However, the Every Student Succeeds Act (ESSA), signed into law on December 10, 2015, by President Obama, reauthorized the Elementary and Secondary Education Act of 1965 and replaced the No Child Left Behind Act of 2001. ESSA regulations empower state and local decision-makers regarding their own systems for school improvement, while holding all students to high standards and preparing them for a level of college and career readiness that the workforce demands (United States of America Department of Education, 2015). Overall, ESSA seeks to protect students of low-income families and students of color from being taught at disproportionate rates by out of field, inexperienced or ineffective teachers. Some of the reported unintended consequences of NCLB created overly stressed students, families, and staff in schools due to an emphasis on high-stakes testing outcomes (Virginia Department of Education SOL Innovation Committee Meeting, 2016). Teachers felt pressure to do test preparations that took away from regular instructional goals and time, in order to ensure higher pass rates and to avoid sanctions placed on schools from state and federal regulations for testing scores. These outcomes deviated from the original intentions of the Elementary and Secondary Act of 1965, which focused on creating equality in schools and declared a “war on poverty” by providing federal funding to supplement local funds for students of need. Thus, ESSA appears to circle back toward the original intentions of the law toward equity and attempts to buffer the overly high-stakes testing
environment created under NCLB. Through these transitions in the federal law, research, theory, and practice of school leaders remains a fundamental ingredient for school success and outcomes.

In conjunction with the accountability era, educational leadership theory has expanded in the last decades to include complex theoretical constructs that move school principals beyond just managing schools to leading schools (Darling-Hammond et al, 2007). Central to this movement in accountability is the expansion of leadership theory to include the critical nature of instructional leaders who use school data effectively (Bernhardt, 2003; Hallinger & Heck, 1998; Lashaway, 2002; Marzano, 2002; Marzano et al., 2005; Reeves, 2002; Taylor, 2010). Combining the standards set for school leaders with the movement in accountability, and the current leadership requirements for the use of school data, raises the bar for school leaders.

In order to increase student achievement results on state standardized tests, research suggests effective principals employ specific leadership actions and daily disciplines in using student performance data to improve effectiveness of instructional practices and raise achievement scores (Bernhardt, 2003; Hallinger & Heck, 1998; Lashaway, 2002; Marzano, 2002; Marzano et al., 2005; Reeves, 2002; Taylor, 2010). Further, the National Policy Board for Educational Administration (2015) standards indicate the use of student performance data by principals increases effectiveness in schools and test results can be used for multiple purposes to impact student outcomes (American Educational Research Association et al., 2014). Therefore, principals’ effective use of the data they have depends on the types and purposes for those data.
sources. Marzano, Waters, and McNulty’s (2005) meta-analysis of 69 studies spans 35 years of research on school leadership by principals. They identify 21 responsibilities of principals, three of which focus on gathering and using data: input; involvement in curriculum, instruction, and assessment; and monitoring/evaluating.

A Data Driven Instructional System (DDIS; Halverson et al., 2007) was chartered in research during a five year study designed to identify how leaders go about creating systems for the use of data by teachers in instruction. The research served to not only confirm previous practices in research, but also to capture effective practices in use by efficacious schools identified for study. The DDIS framework documented six mainstay practices for effective leaders’ use. The framework includes data acquisition, program alignment, program design, formative feedback, and test preparation. Data acquisition includes any process designed to collect and prepare information for use in guiding teaching and learning and includes many various types of information from traditional forms, such as discipline, attendance, grades, demographics, budgets, and master schedules, to newer pieces of information, such as reading measures, standardized assessments, and progress monitoring data. Program alignment processes balance relevant content and performance standards as hallmarks for planning and program evaluation. Designing programs not only encompasses curricular, pedagogical, and instructional strategies and resources, but also, folds in policies and financial commitments to maintain program designs in schools.

Formative feedback is a regular phrase used in schools now, and is a natural part of the DDIS process because it gets to the heart of a “learner-focused” (Halverson et al.,
2007, p. 15) evaluative cycle for leaders and teachers to improve student learning and the quality of school programming, in general. This level of personalization in a dynamic learner-focused cycle is what can drive accelerative learning. Lastly, test preparation can have negative connotations to classroom teachers, even though it is a recognized component of an effective DDIS process that school leaders can and should use. They use it for informing decisions and improving performance, and as part of the feedback loop for students, as a motivational tool for self-reflection and goal setting. Together, the component parts of the DDIS system have merit and value in leading schools in the age of accountability. Specifically, they have value in helping to reduce the stagnant rate of reading failure leading into the third grade.

Furthering the notion of systems thinking in literacy and assessment, in creating systems of interventions, Dorn and Soffos’s (2002) work features school climate, classroom instruction and targeted supplemental support as requirements for continuous literacy improvement. Her system is also situated on the use of valid assessments for assessing the impact of intervention approaches on student achievement, adding to the impetus for overlaying leadership use of data with assessment measures as part of the process to evaluate the effectiveness of the school, the classroom, and the interventions. Data should be transparent and visible to school faculty, link formative assessments to instruction, and be used to find system trends as well as patterns within groups of students (Dorn & Soffos, 2002). Schools that consistently use common assessments developed collaboratively and scored by all teachers at a grade level were found among schools with the greatest gains in student achievement (Reeves, 2002).
Dorn, French, and Jones (1998), as experts in the field of literacy, began to weave together systems change or systems thinking, leadership, literacy learning and professional learning communities to create a catalyst for facilitating a change process (Dorn & Soffos, 2002) using reading measures to evaluate and guide practice. This intersection of concepts brings together leadership and literacy in decision making for reducing failure in reading.

**Assessment in Reading Instruction**

While the SOL is said to be a test of overall comprehension (Joint Legislative Audit and Review Commission, 2011), it is based on a standardized curriculum. Recognition of the standards based movement and framework is key to this research study because it has led the field of education to the widespread application of state mandated testing, which in turn, impacts what kinds of assessment practices teachers use and how they use it. As researchers Glasser and Linn (1993) asserted in the early 1990s, educators may not recognize the importance of the standards movement in America until it is in retrospect. This appears quite apropos in relation to the constructs presented here.

**Standards-based curriculum assessment.** Curriculum alignment involves bringing several aspects of education together to create a balanced, harmonious instructional environment where students successfully learn the intended learning outcomes. A brief look at the movement towards a standards-based educational environment provides the backdrop for the discussion about the rise in use of informal reading inventories and state mandated reading tests. It is pertinent to note the historical
significance of how education evolved into the present era of a standards-based curriculum and assessment.

The National Assessment of Educational Progress (NAEP), initiated in 1971 by the federal government, monitors student achievement in basic skills. The NAEP established absolute proficiency levels to report to the public (Allington & McGill-Franzen, 2004; Rothman, 1995), and as concerns over the failing status of American schools began to disquiet our nation’s leaders in the 1980s, the publication of *A Nation at Risk* took a closer look into educational practices, thus launching the reform of academic achievement and accountability issues (Education Commission of the States, 2002). A 1993 National Education Summit set the path for further reform efforts toward our status today by positing the need for higher academic student achievement, production of rigorous tests to ensure students are meeting the standards, and the introduction of a call for accountability systems. By 1998, almost all states were in the process of implementing academic standards in math and reading. Under George H. W. Bush’s presidency, the Goals 2000 reform ignited the use of content standards. However, the year 2000 did not see the actualization of the achievement of the goals set forth by the Bush administration’s Goals 2000 reform effort. Thus, in 2001, public education witnessed the birth of the No Child Left Behind Act, alongside the reauthorized Individuals with Disabilities Education Act legislation as well.

Taking a microscopic look into the importance of curricular and instructional alignment, a great push began for educators to identify intended learning outcomes for their grade level/subject matter, and school systems were charged with further expanding
on the state standards of learning in local curriculum guides. Policy makers and educational leaders thought standards would improve student achievement through clearly defining what was to be taught and the level of performance expected (Ravitch, 1995). District-wide curricular frameworks, scope and sequence guides, and pacing guides are designed to provide teachers with blueprints for long-range and short-range instruction, and schools are expected to plan for units of study and day to day individual lessons delivered in classrooms. Within the classroom setting, teachers historically have used the district’s framework, scope and sequence, and pacing guides along with adopted textbook series and publisher materials to plan and carry out units of study, teach lessons, monitor learning and enrich or remediate as necessary to ensure all students have improved student-learning.

When combined, the set of intended learning outcomes mirrored with the effectively planned experiences, activities, and interventions of good instruction is believed to provide alignment between the curriculum, the instruction, and the assessments, and thus, improved student outcomes (Soloman, 2009). Without alignment among the curriculum, instruction, and assessment, it is considered that instruction will most likely be hit or miss. Having standards, reports Soloman (2009), requires a consensus making process that explicitly represents statements of the American culture that is then reflected in written school curriculum. The standards are then transferred for alignment into curricular resources, assessment resources, and instructional activities. Even with successfully aligned curriculum, assessment, and instruction, the reduction of the number of students failing to learn to read by third grade still eludes U.S. educators
(Christie & Rose, 2012; Education Commission of the States, 2002; Hernandez, 2011; Marzano & Kendall, 1996). Pondering about the role of curricular, instructional, and assessment alignment as a whole then needs to narrow into a closer look at reading assessment, specifically.

**Historical overview of reading assessment.** Historically, only students with difficulties progressing were assessed and monitored with diagnostic reading tools such as an IRI. It wasn’t until the 1970s or later that the practice of using IRIs as a means to identify specific reading levels and monitor student progress became common practice. It has been a long-standing practice to use informal reading inventories, starting as early as the 1940s and becoming commonplace in the 1970s (Russell, 2013). Prior to the widespread use of IRIs, teachers tended to rely on the adopted textbook publishers for assessment components that consisted of graded curriculum texts, skills worksheets, and accompanying unit tests. Students were considered on-grade level if they were progressing through the published textbook series of books, skills worksheets, and unit assessments. Some of the main users of IRIs during this time were reading specialists and clinicians who primarily used reading inventories to diagnose students referred for special monitoring or interventions (McKeone, 2005). Teachers did not have formal assessment information about reading levels for students in their class that were not being monitored by the reading specialist (McKeone, 2005). The change really came into practice with a call for accountability and assessment systems from NCLB for state mandated testing. The IRIs then began to take hold as part of universal screening for all students in the later part of the 1990s and early 2000s.
Classroom evidence suggests the early detection and correction of reading difficulties can improve children’s reading achievement and, thus early assessment is a key factor in that process (Clay, 1993; Paris & Hoffman, 2004; Snow, Burns, & Griffin, 1998) and is essential for all stakeholders (Paris & Hoffman, 2004). Surveys of early literacy achievement began to show shifts over time reflecting the influences of developing methodologies and theories of children’s learning which included a shift in the 1990s to be more inclusive of contributions from emergent literacy theory, process writing approaches, and performance assessments (Paris & Hoffman, 2004). Guided reading methods, Reading Recovery®, and other early literacy methods contributed to the use of leveled books as part of an IRI process for assessing developing young readers. Survey results done by the Center for Improvement of Early Reading Achievement (CIERA) in an investigation of early literacy assessment practices revealed teachers have the most faith in teacher-designed and used assessments over standardized and commercial assessments, to include the use of informal reading inventories (Paris & Hoffman, 2004). The work done by CIERA added to the increasing legitimacy of IRIs for assessing not only student growth in reading, but also as program assessment tools for administrators (Paris & Hoffman, 2004), especially when combined with other forms of reading measures to create a holistic picture of children’s developing literacy knowledge and skills. They offer information in the form of diagnostic and summative means in authentic form, which teachers find appealing and believe to be valid in informing their instruction (Paris & Hoffman, 2004).
Systems of reading assessment. There are four areas of assessment: screening, diagnosis, progress monitoring, and outcomes based, according to the Standards for Educational and Psychological Testing (American Educational Research Association et al., 2014). Further, results used for a variety of purposes include evaluating student achievement and growth in a domain, diagnosing student strengths and weaknesses, planning educational interventions, designing individual instructional plans, and placing students in appropriate educational settings. The IRI and SOL assessments used across Virginia fit into each of the recommended categories by the Standards for Educational and Psychological Testing. IRIs can be used as a screening tool, for diagnostic purposes, and a progress monitoring tool, while the SOL assessment is an outcomes based criterion referenced assessment intended to measure mainly comprehension. Together they comprise parts of a system for assessment that includes not only a standardized state mandated assessment, but also classroom assessments as suggested by field experts on educational assessment (International Reading Association, 1999; Paris, Paris, & Carpenter, 2001).

Having multiple types of measures, such as the IRI and the SOL, meets the need for educators to have both high-quality classroom and large-scale assessments that can be used effectively as a total system. Not only can standardized reading data be used to compare the achievement of students against other students, it also can be used to provide teachers and district leaders purposeful information about reading strengths and needs of students (McKeone, 2005). Using IRIs in a variety of ways by both classroom teachers and reading researchers has practical implications when combined with other measures or
alone (Allen & Hancock, 2008; Nilsson, 2013). Theoretical models of reading provide
the backdrop for the use of assessments in the form of informal inventories as a means for
identifying student strengths and needs as developing readers. Models of reading that
subscribe to the usefulness of IRIs as part of their framework include developmental and
interactive models (Nilsson, 2013). Because it is believed that readers progress through a
series of developmental stages, IRIs provide insightful information pertaining to the
idiosyncratic nuances of learners progressing along a continuum of development.
Likewise, a balanced literacy approach to reading instruction includes a balanced
approach of components assessed and the purposes assessments are used for in education
children can and should include a variety of components across reading factors providing
all stakeholders with not only individual progress, but also normative standards of
achievement. Again, subscribing to the developmental theory of early literacy
acquisition, use of multiple assessments are believed to “reveal the most information”

Standards of Learning assessments. The reading SOL assessment are part of the
federal legislation requiring all public schools to assess students in grades three through
eight (Ravitch, 2009; Siegrest & Van Patten, 2007; Virginia Department of Education,
2015). The public school system in the state of Virginia utilizes the state Standards of
Learning to define the outcomes and expectations for all students in the K-12 setting.
These standards articulate the basic achievement levels for students in the areas of
English, mathematics, science, history, social science, technology, fine arts, foreign
language, health and physical education and driver education. The standards were created to encompass the united vision of parents, teachers, administrators, academicians, and business and community leaders for our students as they complete their public school careers. They provide the foundation for instructional programming in the public school setting. As a teacher in a public K-12 setting, the SOL provides the framework for the curriculum at each grade level. Essentially, any instructional lesson plans written and used in the classroom should meet the state standards of learning objectives and build student knowledge for successful obtainment of the essential knowledge spelled out in the SOL. Lesson plans and instructional learning objectives and activities should be correlated to the SOL (as a minimum). Thus, whatever materials, programs, approaches, or instructional activities and resources are used by teachers in their classroom should be aligned with the grade level specific objectives for the grade in which they are teaching.

The curriculum framework is designed as a tool to assist schools and teachers in appropriately planning for and executing instruction for the areas identified in the state Standards of Learning. Through the use of the curriculum framework, teachers craft lessons targeting the essential content knowledge and specific skills students need in their current grade and/or subject areas. The framework provides the basis of instructional planning, but does not limit the instructional programs of schools. Teachers and schools benefit from using the framework as the master resource for designing and implementing their instructional programs. Thus, ensuring their students are well prepared academically for mastery of the expected content and are able to demonstrate said mastery on the SOL assessments.
Student assessments in Virginia, as described in the Virginia Standards of Learning Assessment Technical Report (2014-2015), began during a period of “significant reform” of the educational system in 1994, yielding what the state now considers as three major elements: high academic standards, tests to measure progress, and accountability. The SOL assessments are standards-based for the purpose of measuring student performance in content areas. For this study, only the third grade reading SOL is included. The test construction includes not only multiple-choice (MC) items, but also what the state refers to as “technology enhanced items” (TEIs). The TEIs provide a way of allowing students to respond in ways besides the MC format. Test blueprints, item development specifications, multiple review committees, and field testing comprise the process for the development of the SOL assessments. Considered as a broad representation of what parents, classroom teachers and school administrators view as important content for students to learn, the SOL tests are reviewed and updated on a 7 year cycle. The tests results included in the database to be accessed for this study occurred during the 2010-2014 school years, in which the state was using what was called the 2010 standards and students were utilizing the online web-based version during this time frame versus the former version of paper/pencil used in prior years.

Schools in Virginia are guided in their planning for instruction through the use of the curricular frameworks and test blueprints which help educators align instruction with the SOL assessments. Teachers have an indication of the emphasis placed on different areas within the curriculum for third grade reading by the number and type of questions covered within categories.
Informal reading inventories. Informal reading inventories are used across the nation as the predominant measurement of student growth in reading acquisition. There are many types and varieties available for use. School districts typically adopt a system-wide commercially published IRI, train teachers, and begin collecting district-wide student performance on the assessment at regular intervals.

Emmett Betts in 1936 first developed a technique that simply used a set of graded readers to calculate children’s reading levels (Russell, 2013). While Betts was not the original creator of the IRI concept, he developed the word recognition and comprehension percentages for the categories of independent, instructional, and frustrational reading levels using the varied leveled book system. Karl Douglas Waldo in 1915 pioneered an informal assessment of both silent and oral reading patterns in children’s reading performance. This was the beginning of what would become a major method for teachers to assess students’ reading level and plan instruction accordingly. During the 1940s, Betts crafted the procedures and standards for determining the functional levels of reading, and since then, educators have been using informal reading inventories for a multitude of purposes. Ken and Yetta Goodman are widely credited as the ones to solidify the use of the IRI through their development of what they devised as “miscue analysis” in the 1960s, which is one part of the IRI process. By 1972, Yetta Goodman and Carolyn Burke produced their version of an IRI in the form of the Reading Miscue Inventory (Brown, Goodman, & Marek, 1996).

Another major contributor to the IRI movement included psychologist and educational researcher, Marie M. Clay, who began her work in the late 1960s publishing
her dissertation on the systematic errors of five year old readers (Brown et al., 1996). She went on to cement the term “running records” which were the cornerstone of her development of an early intervention program called, Reading Recovery®. While the intervention program used running records as a form of an IRI that could be used on any set of graded books, the procedure eventually carried over into the everyday use in some classrooms and schools as a part of daily observation and anecdotal record keeping of children’s reading behaviors. Thus, both running records and IRIs are rooted in the educational field. Interestingly enough, the use of the IRI process coincides with the historical push for accountability and a system for monitoring student achievement that was ushered in through the political arena with two federal reports in the 1980s — *A Nation at Risk* (National Commission on Educational Excellence, 1983) and *Becoming a Nation of Readers* (Anderson, Hiebert, Scott, & Wilkinson, 1985). This helped put the U.S. on a path for scientifically-based educational programming.

Even with the long-term established use of IRIs as a practice, some experts still debate both the reliability of IRIs and the original theoretical construct behind them as first published by Betts’ initial rendition in his 1946 *Foundations of Reading Instruction* textbook (Pondiscio, 2014). Regardless of any reservations experts in the field may have about the reliable nature of a teacher-administered IRI, the empirical data show there is predictive and cor relational value in the use of an IRI against state mandated end of grade reading assessments (Askew, 2011; Johnson, 2014; Morris 2004). Table 1 shows the highlights of some commonly used commercial informal reading inventories and the main features measured by each (Russell, 2013).
### Table 1

**Commonly Used Informal Reading Inventories**

<table>
<thead>
<tr>
<th>Title</th>
<th>Author(s)</th>
<th>Publisher</th>
<th>Copyright Date</th>
<th>Grade Levels</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Authors</td>
<td>Publisher</td>
<td>Publication Year</td>
<td>Grade Range</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>--------------------------------------------------------</td>
<td>----------------------------------</td>
<td>------------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Critical Reading Inventory: Assessing Student’s Reading and Thinking &amp; Readers Passages, 2nd Edition</td>
<td>Mary DeKonty Applegate, Kathleen Benson Quinn, &amp; Anthony J. Applegate</td>
<td>Pearson Prentice Hall</td>
<td>2007</td>
<td>Senior High</td>
<td></td>
</tr>
<tr>
<td>Informal Reading Inventory: Pre-Primer to Twelfth Grade, 6th Edition</td>
<td>Paul C. Burns &amp; Betty D. Roe</td>
<td>Houghton Mifflin Company</td>
<td>2002</td>
<td>12th Grade</td>
<td></td>
</tr>
<tr>
<td>Informal Reading – Thinking Inventory</td>
<td>Anthony V. Manzo, Ula C. Manzo, &amp; Michael C. McKenna</td>
<td>Thomson Wadsworth</td>
<td>1995</td>
<td>11th Grade</td>
<td></td>
</tr>
</tbody>
</table>

Word Lists, Narrative & Informational Passages, Oral Reading, Miscue Analysis, Comprehension: Text-Bound, Inference, Critical Response Items, and Retelling

Emergent Literacy, Sight Words, Phonics, Structural Analysis, Contractions Test, Word List Survey, Context Clue Use, Dictionary Use, Graded Word List, Reading Passages, Reading Interests

Interest Interviews, Leveled Sentences, Leveled Passages, Miscue analysis, Narrative & Expository Passages, Retelling, Intervention Strategies

Graded Word Lists, Graded Passages, Miscue Analysis, Comprehension Questions (Main idea, Inference, Sequence, Vocabulary), Retelling, Assessing Use of Context Clues

Graded Word Lists, Graded Passages, Prior Knowledge, Enjoyment, Literal Questions (Factual, Fund of Knowledge, vocabulary, Inferential, Inference, Abstract Concept, Analogical Reasoning, Concept- Based facts, Critical Evaluative, Explanation, Open-ended, Problem
<table>
<thead>
<tr>
<th>Title</th>
<th>Author(s)</th>
<th>Year</th>
<th>Grade</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qualitative Reading Inventory – 5, 5th Edition</td>
<td>Lauren Leslie, Jo Ann Schudt Caldwell</td>
<td>2011</td>
<td>PP – 12th Grade</td>
<td>Graded Word Lists, Graded Passages (Narrative &amp; Expository), Prediction Task, Retelling, Comprehension Questions (Explicit &amp; Inference), Miscue Analysis, Look-Backs, Think-Alouds, Notetaking</td>
</tr>
<tr>
<td>3–Minute Reading Assessments: Word Recognition, Fluency, &amp; Comprehension</td>
<td>Timothy V. Rasinski &amp; Nancy PadakExpress</td>
<td>2005</td>
<td>1st – 4th Grade</td>
<td>Word Recognition, Reading Fluency &amp; Performance, Comprehension (Fact, Main Idea, Detail, sequence, Personal Connection)</td>
</tr>
</tbody>
</table>
Informal reading inventories were designed to inform teachers for the advancement of students as better readers, not just benchmarking according to a set lock-step grade level standard. Askew’s (2011) study found the following,

The lack of information on IRIs could contribute to misunderstandings on the part of the teachers in terms of how to approach using the IRI in the classroom as a predictor of how well students will do on a state assessment (p. 91).

Knowing the predictive value exists is the first step, understanding the nature of the full scope of the categories for IRI levels and the potential outcomes on end of year state reading measures could truly inform leaders for decision-making to shape educational programming and work toward the state educational and societal goals for reducing reading failure by third grade. Leaders and educators could then adjust goals and pacing to better align with the objective to accelerate reading levels, not just maintaining curricular pacing, if that is not working currently as NAEP data suggest. Further understanding the ranges of how children score and what the predictive outcomes are would serve to inform leaders making decisions about instructional programming, benchmarking goals, pacing of curricular goals, implementation of acceleration goals, and assigning resources for elementary students reaching the third grade year.

Spector’s (2005) analysis of IRIs revealed inconsistent use of appropriate reliability information (Nilsson, 2013). Reliability was examined in Spector’s (2005) original work, but re-examined by Nilsson (2013) for an updated examination. Criticisms over the years, have rested mostly on the nature of reliability, rather than validity of IRIs. While validity denotes whether the test measures what it claims to measure, reliability
emphasizes the consistency of test scores (Creswell, 2005). Specifically, when considering the Standards for Educational and Psychological Testing, traditional methods of reliability in the form of test-retest, alternate-forms, internal consistency, and interrater reliability are limited in the technical manuals of the published IRIs studied by Spector. Ultimately, recognizing there are multiple purposes for the use of IRIs, Spector and others (Nilsson, 2013) settled on establishing recommended acceptable use ranges for decision-making based on aspects of the reliability information that is provided by various published IRIs.

Even noting the varying reliability information reported for some of the widely used IRIs, empirical research on the correlational qualities of the use of IRIs with other forms of reading measures, such as standardized assessments, criterion referenced tests or norm references tests indicates the value of IRIs persists in spite of the reliability debate. Interestingly enough, Manzo and Manzo (2013) remarked on the incredible fortitude of the staying power of IRI use since the original inception of Betts’ criteria in the 1950s era. Even as theoretical models of processing evolved into the 21st century, the primary structure measuring reading levels of young children remains constant in all various additions to the original decoding process. While there are a variety of commercially produced versions of IRIs, practitioners in the field must work to establish alignment across various tools used in schools systems. Each IRI system uses its own leveling system, and thus, educators frequently have to work to find cross-references for what one level may mean within different systems. (See Appendix A for an example of one school system’s solution to understanding the leveling system across a variety of tools.) This
type of resource helps educators make sure they understand the equivalencies when discussing levels from a variety of sources, as is often the case (Saint Paul’s Project for Academic Excellence, 2007).

Critics have rallied around a chief complaint of IRIs lacking in the ability to measure and capture the higher order thinking skills sought for students in the rigorous demands of 21st century learning. Applegate, Quinn, and Applegate (2008) completed a thorough review of eight published IRIs in a 2002 study that clearly notated the lack of higher order thinking questioning in the IRI process. Even with this clearly noted limited ability to capture higher ordering thinking within the basic structures of published IRIs, Manzo and Manzo (2013) suggest supplemental processes can garner this type of information adequately. Their procedures outlined in the Informal Reading-Thinking Inventory offers the additional comprehension information. Despite the acknowledged gap in higher order thinking skills on IRIs (Applegate, et al., 2008; Manzo & Manzo, 2013), some correlational studies have found predictive value in the use of IRIs on statewide comprehension measures to be useful (Johnson, 2014; Morris, 2004).

A final note about IRIs pertains to the arguments about reliability and technical rigor, as a particular focus has rested on reliability more so than validity (Nilsson, 2013). While this review of the use of informal reading inventories does not, in particular, focus on the major distinctions between the wide varieties available at this time, there are numerous in-depth reviews of the nuances of commercially published IRIs for the interested reader. Since the validity and reliability of informal reading inventories does rest on uniform administration procedures and scoring processes, training and teacher
adherence to the procedures can impact the reliability. Those who have done in-depth studies, have done thorough reviews on the pros and cons of various commercially published IRIs (Nilsson, 2008; 2013; Paris & Carpenter, 2003; Spector, 2005). Spector (2005), having been one the more critical evaluators of IRIs, cites poor documentation by publishers and weak methodology on the reliability as evidence of a disregard for the importance of reliability. Spector’s development of an analytical minimal criteria framework for determining reliability for IRIs made it difficult for them to be considered reliable by some. Since the initial reports, some makers of IRIs have updated their methodologies and reported on additional reliability information. Regardless of the stringent application of Spector’s original evaluation of IRIs, others still recognize the value of IRIs when adequately informed about making choices among the various ones available to choose, specifically which IRIs are well suited for their needs, and when proper training and support for procedures are followed (Fountas & Pinnell, 2014; Nilsson, 2013; Paris & Carpenter, 2003).

**Grade Three Reading Conundrum**

Much has been debated for decades about the methods and pedagogy of best practice in reading instruction, but field experts coalesce around the complexity of learning to read agreeing that it is the most complex skill young children are expected to develop in the primary grades. The challenge for educators of primary grade students starts in kindergarten as students enter classrooms with a wide variety of experiences and backgrounds creating a large spread of needs within each classroom. Educators then face the challenge of decision-making within the complexities of early literacy teaching and
learning in an attempt to bridge large differences in early developing literacy skills, but by third grade more than one in four students is already behind in the ability to comprehend written text at grade level expectations (Katzir et al., 2006). Christie and Rose (2012) suggest that “forty years of well-meaning state and national reading initiatives have not produced significantly higher student mastery” (p. 3) adding to a conundrum of the state of reform for literacy in the United States. Hernandez (2011) in his report on third grade reading scores and poverty implications toward graduation rates states:

Results of a longitudinal study of nearly 4,000 students find that those who don’t read proficiently by third grade are four times more likely to leave school without a diploma than proficient readers. For the worst readers, those who couldn’t master even the basic skills by third grade, the rate is nearly six times greater. While these struggling readers account for about a third of the students, they represent more than three fifths of those who eventually drop out or fail to graduate on time. (p. 3)

Even though the data found on long term trends from as far back as 1992 document a slight upward trend with an increase from 29% at or above proficient to the current level of 36%, proficiency as defined by the NAEP, is just above the most basic level of functional reading (U.S. Department of Education, 2016). While it shows some gains at a low level of proficiency, it simply is not sufficient (Christie & Rose, 2012). The proficiency level of reading, according to the NAEP (U.S. Department of Education, 2016), stops short of being able to analyze, think critically about what is read, and make
judgments supported by inferential understanding of text. Considering that only 36% of all students assessed by the NAEP in the nation are able to read at a proficient level or higher falls short of the sweeping large scale goals of federal and state reform efforts to increase higher achieving educational outcomes for students in the U.S. Proficient reading as categorized on the NAEP means students are able to draw conclusions and make evaluations by integrating and interpreting what is read. With so few students able to reach this benchmark, it leaves the large majority of students performing at or below the basic level which, at a minimum, includes a low level of comprehension skills emphasizing the ability to simply locate relevant information, understand and interpret meaning at the word level in context while applying some simple inferences to find details and draw simple conclusions. While recognizing some progress has occurred from 1992 to 2016 with seven percentage points in gains from 29% to 36%, decades of reform efforts to reduce reading failure have not reached even the 40% mark for students achieving at least a proficient level in reading.

**Acceleration in reading instruction.** Given the background, history, and status of reading achievement in our nation and the state of Virginia, the concept of accelerating student learning in early literacy skill development becomes essential to the discussion. When the NAEP refers to an achievement gap, they are referring to significant differences on assessment outcomes between various groups of students such as White or Black students, Hispanic and non-Hispanic, or second language learners and students with disabilities (U.S. Department of Education, 2016). To look at the concept of closing achievement gaps in another way is helpful in relation to this study. Experts in the field
of early intervention who specialize in the instructional methods of reducing reading failure often refer to the notion of accelerated learning. Accelerated learning is more helpful for those working directly with students to center a clear vision of how to catch lower performing students up to the levels at which their grade and age level peers are performing. So while policy makers at the state and federal level, as well as the school district level, may focus their reporting and decision-making around dialogue framed in closing gaps, direct service providers emphasize the critical nature of accelerated learning for all students. Therefore, this study refers to reducing the rate of reading failure to encapsulate all students in relation to the reading trends discussed in this context, and not about gap groups as is often the case when the phrase “closing the gap” is used.

Accelerative learning to interventionists in the early grades, such as those trained in Reading Recovery methods or guided reading methods, means that at each interval when a student or group of students show progress and achievement at a particular gradient of difficulty, they are then moved onward to the next level on a “ladder of progress” (Clay, 1991, p. 125). Clay’s (1991) work led her to what she called a literacy processing theory which formed the basis for an early intervention method called Reading Recovery. In this approach to reducing reading failure in the early years of literacy acquisition, she frames extensive complex literacy processing theory into methods or instructional practice designed to accelerate a child’s early learning with the aim of closing reading achievement gaps for any student performing below that of their age level peers. Clay (1991) describes the process as reading work that, when done effectively, can change the trajectory of a child’s early literacy learning progress. “The
reading work clocks up more experience for the network with each of the features of print attended to. It allows the partially familiar to become familiar and the new to become familiar in an ever-changing sequence” (Clay, 1991, p. 328). Table 3 shows a widely accepted trajectory for student growth over time. It also features a range of levels to indicate what constitutes a below grade level reader and an above grade level reader based on the grade level year of school (Fountas & Pinnell, 2012).
**Figure 3:** This figure shows an example of a widely used chart on the Instructional Level Expectations for Reading.
The distinction between accelerative teaching and learning and standards based curricular goals is noteworthy. In acceleration, the goal is to always be moving children once gains are made, regardless of pacing guides or intervals of time across a school year, such as quarterly benchmarking. This is a distinction not to be overlooked as part of this dialogue, given the reading conundrum. While interventionists operate under a framework that captures the nature of accelerated teaching and learning, classroom teachers are typically working under the auspices of standards based curricula with a prescribed teaching sequence for goals and objectives in that grade level that are often time bound by quarterly increments.

While acceleration models have historically been applied to small group or individual instructional settings, there is one example of a large scale approach to bringing an accelerated teaching and learning approach out of the intervention setting and into the classroom setting, which began in the late 1990s in New Zealand. This acceleration model for classroom use has only recently carried over into a small level of implementation in the United States (Duncan, 2016). Differences in diversity among linguistic, cultural, ethnic or poor socio-economic circumstances are often indicated in the data in the form of lower performance scores on literacy achievement measures than mainstream counterparts (Duncan & Johnson, n.d.b). Recognizing that as children enter into formal literacy instruction, they are coming from a diverse variety of backgrounds with differences that can influence progress in school is important to school leaders challenged with determining the resources and approaches necessary to reducing major differences in literacy achievement progress. This whole class approach based on an
acceleration model of teaching and learning has been researched, trialed, and studied for the very first entrants to school at age five in New Zealand, and now, is in operation in the Unites States through Georgia State University for the kindergarten year. Successful Start (Duncan & Johnson, n.d.a). was developed based on the work of Gwenneth Phillips and Pauline Smith through a major research project and report funded by the Ministry of Education in New Zealand, called *Picking up the Pace*, in which they designed the accelerative intervention for all entrants to formal schooling as “First Chance” (Duncan & Johnson, n.d.a). in an attempt to prevent the literacy gap from ever occurring for all children as they enter formal literacy instruction.

Phillips, in her work, first documented evidence that acceleration of identified low progress, at-risk five year old children could be accomplished (Phillips et al., 1997). The seminal findings from this research provided a way of discussing optimal learning opportunities and instructional conditions found to enable an accelerated type of learning to occur within a small group setting (Phillips et al., 1997). Having conducted trials and studied the specific effects, a framework around the practices took shape for it as a primary school intervention given within the classroom setting. This was achieved through additional major studies on the ability of the developer to replicate the benefits of accelerated teaching and learning methods with five year olds through professional development and training of other teachers.

What was found from the research in New Zealand was the ability to close literacy gaps for young children coming from low socio-economic communities with cultural and linguistic minority groups so they performed with the same level of
achievement at the end of the first year of school by age six (Phillips et al., 1997). The favorable outcomes of Successful Start (Duncan, & Johnson, n.d.b) have shown to increase the number of children achieving at expected levels, reduced the number of children needing intensive early interventions, and shown 80% of children reading at appropriate levels when attending school and being taught by a Successful Start teacher on a consistent basis.

Interestingly, the curricular basis for the instructional design of the accelerative teaching and learning approach is founded in research from several works by Phillips (as cited in Duncan, 2016) around storying reading practices of library using families, a 1997 study on the development of activity systems for reading and writing in high progress new entrants and those at risk of low progress after one term in school, and research on low progress older readers at seven years of age. All teaching is then based on observations of children’s behavior, a theory of socio-cultural and co-constructivist ideas about learning and development (Phillips et al., 1997) and core component parts of reading to and with children as well as writing with children. This is in contrast to a standards-based approach as described earlier that is used by public schools in the United States. This is featured prominently in the theoretical constructs in their studies (Phillips et al., 1997) when they state:

It clearly puts the onus on the teacher rather than the child to accelerate rates of progress and attain higher achievement. It demands that teachers accept and plan for multiple pathways to common outcomes (Clay, 1998; McNaughton, 1999) and
highlights the complexity of both the learning environment and instructional process within a classroom. (p. 49)

The authors go on further emphasizing social practices as central to a curriculum as “any recurring, patterned interaction in any setting, which reflects the group’s way of fulfilling goals, can develop into a social practice. Such practices can be idiosyncratic and unique to particular settings” (Phillips et al., 1997, p. 50). Thus, following the child’s strengths and needs and patterns of responding is distinctly different from the U.S. method of following a prescribed set of standards based learning goals paced out across the year for mastery of particular content within the very early years during the critical literacy acquisition stage (Joint Legislative Audit and Review Commission, 2011).

This is further reiterated by McNaughton (2014) in an article on the work and contributions of Marie Clay. As a pioneer in the field of emergent literacy, Clay referred to early learning that occurs even prior to formal instruction and used those observations to “direct our attention to developmental descriptions” (McNaughton, 2014, p. 53), help shape understanding of trajectories of development and our ability to alter them, as well as fine-tuning a processing model that responds to individuals allowing teachers to be responsive in teaching methods, not following any set prescription of standardized curricular goals, so much as achieving the over-arching goal of an early literacy self-extending system (McNaughton, 2014). The system enables the child to learn more and more by engaging in reading and writing acts designed for individuals or small groups, and thereby, “promotes emerging skill, allows for the child to work with familiar, introduces the unfamiliar in a measured way, and deals constructively with slips and
errors” (Clay & Cazden, 1990, p. 212). McNaughton (2014) sums up an important note of the application of these constructs by stating, “This view of developmental patterns which need to be understood at a personal level means that effective teaching requires highly knowledgeable, highly adaptable, and highly strategic experts,” (p. 89) which presents a challenge for the typical standards-based curriculum and assessment practices delivered at the classroom level.

**Role of leveled texts in the early grades.** Part and parcel in the acceleration framework is the use of leveled texts in the early grades. In order to accelerate student learning goals, a differentiated approach to instruction is paramount. A standards based approach tied to a graded curriculum does not serve the below level reader well and can constrain teachers to feel pressure in following a lock-step sequence when what is needed to accelerate learning is a differentiated approach. The differentiated method of instruction uses the leveled text system identified in the IRI assessment process to pinpoint the exact range of reading ability a student currently masters and allows the teacher to direct instruction at the exact level needed.

On the opposing side of the spectrum, another outgrowth of the era of accountability gave way to the movement for a common set of standards that could be used across states to give consistency and add rigor to lackluster state curricular expectations. This came in the form of the Common Core State Standards (CCSS; Council of Chief State School Officers, 2015). One way that CCSS focus on improving literacy is by introducing higher levels of textual complexity into the instructional mix, creating two very different sides, in some expert’s minds, about the nature of text
complexity. One side selects steadfast support for an anti-leveled text approach, citing the very data noted early on in this study with the unchanging NAEP scores as proof that years of over-application of a leveled text system has produced no better readers in the United States as compared to higher performing other countries. The other side holds tight to the view of matching leveled texts to readers, especially in the early grades, to provide a carefully selected gradient of text experience, and touts that anything else is risky for students in early developmental stages. This debate has sparked some experts to clarify their stance, such as Fountas and Pinnell (2014) who published an article clearly articulating they support a balance across a wide range of texts.

Nonetheless, to further explore the rigor behind common core, the Council of Chief State School Officers (2015) stated, “the research shows that while the complexity of reading demands for college, career, and citizenship have held steady or risen over the past half century, the complexity of texts students are exposed to has steadily decreased” (p. 1). Proponents for higher text complexity cite a faulty research foundation from Betts’ early work, to which the entire use of an instructional level theory is predicated upon. Shanahan (2011) stated that the seminal work of Betts in his Foundations of Reading textbook is not based on any empirical evidence and goes on to feature William Powell’s work across the 1960s, 70s, and 80s as revealing the data used in Betts’ work did not find any optimum levels for student learning, contrary to the entire construct proposed by Betts. Some consider 85% word accuracy (Shanahan, 2011) as a better predictor for student learning in literacy compared with the 95% word accuracy suggested in Betts’ work. This places the argument for common core’s higher rigor and
level of text complexity within a better research base than that proposed and widely accepted from Betts.

Along the same lines, Shanahan emphasizes a fundamental philosophical divide for the children to read with the most minimal amount of teacher support on easier leveled text, and instead purports readers, according to research from the opposing side, should have substantial teaching support on more challenging texts. The belief is that this practice maximizes student learning and the instructional level text theory is too constrictive and narrow of an approach, some going so far as to say “use of leveled text beyond the very first years of primary school yielded no achievement gains in students” (Pondiscio, 2014, para. 9). While critics of instructional level theory make their case for increased rigor, there are those who opt for a middle ground (Fountas & Pinnell, 2014) using a balance of independent, instructional and frustrational levels within core programs (Pondiscio, 2014).

Rounding out the discussion on a more middle ground approach, literacy experts Fountas and Pinnell (2014) suggest literacy programs should include a range of books within varying methods of instruction as well include use of whole-group methods, small-group methods, and individual teaching methods. These encompass a multitude of opportunities for students to engage and comprehend deeply across a wide array of text complexity within the various settings for whole, small, and individual settings in various genres (Fountas & Pinnell, 2014). Using the gradient of text difficulty, the goal is to continually adjust and move students up the gradient once they show gains on a level. This differentiated or personalized instructional approach claims to move students
forward at all times, continually increasing reading ability. Students are “soon able to do independently what they needed teacher support for yesterday” as the teaching moves on “to push the boundaries further” (Fountas & Pinnell, 2014, p. 3).

**Understanding the Role of Reading Measures**

The constructs of leadership, using multiple forms of data in a data assessment system, and the role of the reading assessments are all critical to the very early learner of literacy, as has been explored thus far. In a burgeoning territory of empirical research on the relationships of IRI and state mandated standardized reading tests, there is a growing body of evidence. This evidence suggests there is predictive value in the use of IRIs, primarily in relation to validating the types of assessments against the literacy processing models used in schools or as an evaluation of the function of high-stakes testing, redundancy in testing, or instructional decision making at the classroom level. This study seeks to overlap the growing body of research within the context of school leadership and a higher level of decision making for instructional programming, setting benchmark goals for the instructional reading levels in school-wide or district programming, and as critical information in school leaders’ goal to reduce the number of children failing to read by third grade. There are potentially additional layers of impactful and meaningful analyses and data within the datasets than has been studied and reported on to date. I posit that information, not yet explored as in-depth, is critical to the reading conundrum educators confront.

Even though schools have been using some form of an IRI at the classroom and school level, not just in clinician settings, for approximately the same number of years as
the federal government’s call for scientifically-based educational reform and accountability systems, the NAEP data continue to show few significant gains in reading performance (Christie & Rose, 2012; Education Commission of the States, 2011; Hernandez, 2011). Although the national reports call for accountability and the use of assessment systems for measuring reading levels of student progress on graded reading passages were meant to increase overall reading achievement, it has not translated into the data (Booker et al., 2007; Christie & Rose, 2012; Education Commission of the States, 2011; Hernandez, 2011; Joshi et al., 2009; Ravitch, 2009; Siegrest & Van Patten, 2007; National Center for Education Statistics, 2011). This is in contradiction to our overarching goals to reduce reading failure by third grade through accelerated learning measures and warrants a closer look at the actual data trends across these multiple measures of reading achievement.

In a dissertation study by Morris (2004), relationships between various reading measures and the Virginia Standards of Learning (SOL) assessment were examined from a cohort of 85 students in one elementary school in central Virginia. Similar research questions were posed to reveal moderate correlations in various reading measures. The study found correlations from fall (.49) and spring (.40) scores between second grade, the Phonological Awareness Literacy Screening and the Virginia third grade reading SOL at a significance level. Further outcomes demonstrated a 95% confidence interval of the ability of an IRI to predict passing the reading SOL. The IRI used in the Morris study was the Qualitative Reading Inventory-3 (QRI-3), which consists of pre-primer through high school graded word lists and graded reading passages with comprehension questions
(Leslie & Caldwell, 2001). Morris (2004) sought to determine the reading level necessary to pass the Virginia third grade reading SOL at the minimum level of a 400 scaled score as determined through the use of the IRI measures across the third grade year. Results indicated that by the end of third grade, students needed to be reading independently on fourth grade, but instructional on a sixth grade level. This was aligned further with the notion of teacher taken and scored informal running records that were also analyzed in the Morris (2004) study, which are records of miscue analysis of previously read texts from the instructional program, and not a set of published leveled assessment systems. An analysis of the levels on the running records of instructional program materials taken by teachers across the third grade year also indicated, similarly, students needed to be reading independently at fourth grade in the fall of third grade and instructionally at early sixth grade in the spring with a 95% confidence interval and a significant regression analysis at $p < .001$. This is important because it begins to get to the heart of what administrators need to know to lead schools in appropriate instructional planning and delivery. While Morris’s study begins to provide the starting point of reading levels at the minimum level for successful reading development, as deemed and measured by the VDOE, a full scope and framework of the range of reading levels that correlate with the varying designations for fail, proficient, and pass advanced were not established. Patterns of a full scope could potentially yield a more complete framework for understanding student reading levels needed for making critical instructional decisions among school leaders for instructional pacing, accelerative learning goals, and interventions in schools.
Johnson (2014) looked at the predictive quality of three commonly used assessment measures in Virginia for the purpose of informing teachers and instructional leaders in their data-driven decisions for reading instruction. His work added to Morris’s 2004 study, also based on Virginia’s reading assessment systems, as well as those of McKeone (2005) and Askew (2011) with studies in other states who utilize similar assessments to the SOL. These works contributed to the beginnings of an empirical base around the predictive value of reading assessments used in instructional and leadership data informed decision-making against state mandated high-stakes testing, and an IRI has shown to have a significant relationship with the SOL measure. There is prevalent and widespread use of IRIs as a reading measure, even with a gap in the professional literature on the predictive value against the state mandated SOL (Johnson, 2014).

Children who remain behind the same distance upon ending a school year as when they entered the school year, while having made some progress or movement in reading level, gives educators a false sense of success, and does little to actually close achievement gaps in reading among students. One example of this trend comes from a study done on the middle grade years of sixth, seventh, and eighth grades in Tennessee in the Memphis City Schools (Askew, 2011). The study used an archival dataset containing a large sample of data on 1,110 students from all 33 middle schools who were assessed three times across the 2008-2009 year with an IRI and took the Tennessee Comprehensive Assessment Program in the spring. The Scholastic Reading Inventory is another commercial form of an IRI that is commonly used in the United States. It has two options for use, which includes the traditional paper/pencil version and a computer
adaptive version, both of which have been found to provide reliable and valid results with IRI results. Focusing on the long term outcomes for students who failed to learn to read on grade level in the elementary grades, Askew’s 2011 study looked at student trends in the middle grades. The study showed that while growth did occur, there were no ranges set for what was considered successful growth towards closing achievement gaps, reducing reading failure, and accelerating students closer to grade level ranges needed to demonstrate minimal proficiency requirements by the state. Furthermore, one attention-grabbing finding from Askew’s study was that students did not grow a full year in reading. Students made progress in reading, but did not meet the Lexile growth goal defined as 75 to 100 Lexiles for a full year of growth in all of the middle grades for the Memphis City Schools in the 2008-2009 cohorts studied. Of note, from Askew’s 2011 study on middle school student reading outcomes, the results of the study indicated a .762 correlation for sixth grade students, a .824 correlation for seventh-grade students, and a .738 correlation for eighth grade students. The difference between the fall and spring IRI and the outcomes of the spring state TCAP test were also evaluated and were found to have significant correlations of .738, .469, and .517, for the sixth, seventh, and eighth grades respectively (Askew, 2011).

While predictive value was confirmed across the assessments, IRI data examined further for possible growth measures from beginning to the end of the school year did not show a full year’s growth for students. This is critical information when considering the entire cohort of middle grade students in a district did not statistically show a full year of growth, not only for the goal of a grade level year of growth, but also for students who
needed accelerated gains to reduce reading failure. Again, as a piece of valuable information for school leaders, this data set shows that while students made some growth, the goal for closing gaps and making accelerated growth was not met. The results from Askew’s middle grades study is central to the tenets of this study in that it conveys the linkage between failing to reach proficiency in the elementary years with lasting impact in the middle years. Data bears this out with alarming rates of students continuing to struggle in higher grades for whom the acquisition of basic reading knowledge eluded in the earlier years (National Center for Education Statistics, 2011). Delving a bit deeper into the mean averages reported by Askew, all of the average scores for the spring in each middle year grade fell within a category considered not proficient with the mean average only reaching the low end of the basic level (Askew, 2011).

Findings from Askew (2011), Johnson (2014), and Morris (2004) studies add to the importance for school leaders to use information from IRI reading assessments and provide validation that IRI scores can have predictive value against the SOL outcomes. The Johnson (2014) study is limited in providing more information about the range of categorical scores on IRIs and the SOL outcomes with only one reported data set. A mean score of 3.177 on the QRI equated with a mean score on the SOL of 448.2. Thus, a mean passing score of 448, with 400 being the lowest pass score achievable, correlates with a third grade reading level upon entry. In order to pass the spring third grade reading SOL, students must already be on or above the third grade level at the start of the third grade year to obtain a minimal passing score (Cramer, 2010; Morris, 2004).
Likewise, while these studies found strong correlations in reading measures, in a small study of 110 students in a small suburban district in Pennsylvania, only a modest correlation between the Developmental Reading Assessment for grades four and five in comparison to the Terra Nova and Pennsylvania State System of Assessment outcomes (McKeone, 2005) was found. This is the first study with a more modest correlational finding, and this information is equally pertinent to school leaders and district leaders for whom review and purchase of assessment systems is primary to their role. Outcomes such as these are pivotal in the decisions leaders make regarding instructional resources, especially when schools are evaluating and investing large amounts of financial resources in purchasing assessment kits for large scale use.

A limitation of the studies reviewed stems from the limited information reported on the range of categorized scores to link on, below, or above grade level status in reading as measured by an IRI with the categorized scores of pass, fail, or advanced on standardized test measures. The findings add to the importance for school leaders when considering the use of an IRI and provides validation that IRI scores can have predictive value against state standardized reading measures. Students must be reading both independently and instructionally at much higher levels, according to an IRI measure, when compared to the SOL categorical outcomes (Morris, 2004), but the research is limited in providing a full picture of the connections between the categories on IRIs and other outcomes on reading measures.

As additional findings are explored for the research in this arena, a summary of the overall pertinent finds may be a helpful visual aid. Table 2 shows a matrix of the
extant correlational literature reviewed along with a summation of the types of IRIs used, the criterion or norm-referenced assessments studied, the correlational findings in categorical form, not statistical (For the comprehensive statistical outcomes, please see original works of study by the authors), and the notation of whether or not the studies included any categorical reading level ranges as equated to categorical outcomes on the other assessment measures.
Table 2

Matrix of Extant Correlational Literature

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>IRI Used</th>
<th>Criterion or Norm Referenced Assessments Used in Study</th>
<th>Inclusion of Actual Reading Levels Equated to Categories of IRI and Other Assessments</th>
<th>Correlational Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Askew</td>
<td>2011</td>
<td>Scholastic Reading Inventory</td>
<td>Tennessee Comprehensive Assessment Program</td>
<td></td>
<td>•</td>
</tr>
<tr>
<td>Cramer</td>
<td>2010</td>
<td>Developmental Reading Assessment</td>
<td>Ohio Achievement Assessment</td>
<td>√</td>
<td>●</td>
</tr>
<tr>
<td>Hickey</td>
<td>2012</td>
<td>Developmental Reading Assessment</td>
<td>Ohio Achievement Assessment</td>
<td>√</td>
<td>●</td>
</tr>
<tr>
<td>Johnson</td>
<td>2014</td>
<td>Qualitative Reading Inventory - 5</td>
<td>VA SOL VA Standards of Learning</td>
<td>√</td>
<td>●</td>
</tr>
<tr>
<td>McKeone</td>
<td>2005</td>
<td>Developmental Reading Assessment</td>
<td>Terra Nova and Pennsylvania State System of Assessment</td>
<td></td>
<td>●</td>
</tr>
<tr>
<td>Morris</td>
<td>2004</td>
<td>Qualitative Reading Inventory</td>
<td>VA Standards of Learning</td>
<td>√</td>
<td>●</td>
</tr>
</tbody>
</table>
Hickey’s (2012) research with a large sample of 2,395 students in third grade classrooms from a sample population in a suburban Ohio district during a period of time from 2008-2010 who had been assessed using the Developmental Reading Assessment (DRA) in second and third grades is one of the only studies found to provide further insight into the use of IRIs. Investigating the relationship between the DRA and the third grade Ohio Achievement Assessment in reading (OAA), Hickey examined the correlation and regression analyses revealing the DRA as a strong predictor of the third grade ORAA raw scores with an even stronger correlation to students performing below grade level than those performing on or above grade level.

A meta-analysis of 250 studies by Black and William (1998) concluded the use of formative assessment improves student achievement with as much as a strong effect size of .07 in student gains when used. Supporting the use of such informal reading assessments as IRIs, Black and William (1998) also found the impact to be greater for students working below grade level than those on or above; thus, adding to the evidence of positive correlational effects of IRIs on potential student outcomes on other reading measures. Hickey’s (2012) study supports the Black and Williams findings.

Delving into the conceptual framework for what categorizations exist between the relationships of IRI scores and outcomes on state reading measures, Hickey (2012) features the results of outcomes based on on-grade level, below, and above categories, and concluded that the results of second grade DRA scores were strong predictors of students pass performance on the OAA, especially for low-performing students. Using Hickey’s calculations, only six students were missed as having been identified as at-risk
of failing the third grade state assessment. A DRA level of 24 was found to be the minimum score required for students in the spring of second grade in predicting a passing score of 400 or higher on the Ohio Achievement Assessment (OAA). There was a moderate relationship \((r = 0.47)\) using the Pearson correlation coefficient from the DRA to the OAA (Cramer, 2010). When several reading measures were combined with the DRA and correlated to the OAA (Cramer, 2010), there was a significant relationship found in the predictive value. In contrast to moderate correlations between the third grade DRA and the OAA, the second grade spring DRA level was found to have high and positive correlations \((r = 0.57)\) as also shown in Hickey’s (2012) study as well.

On the opposite side, Hickey’s (2012) results indicated that all students scoring at a level of 38 or higher in the spring of second grade passed the third grade OAA. Correlations were lower for students at or above grade level. There was a dip in the strength of the correlations as students were on or above grade level at the third grade. As a result of the findings from Hickey’s 2012 study, the school district used the information to re-evaluate their assessment system practices and decided to only use the DRA assessment once a year for all students with additional administrations of the assessments geared toward students who are identified as below grade level. Hickey (2012), in his role as an educational leader, substantiates the vital information gathered from reviewing the effectiveness of assessment practices, calculating results, reflecting on outcomes, and using the information to guide decision-making as crucial to the role of educational leader (Hickey, 2012). Hoy and Miskel (2008) stress the nature of leadership as involving a sense of breaking the status quo and initiating a process for adaptive
change to lead people through a united goal. Clearly, in Hickey’s (2012) work, this is evident and the data truly informed the decision making of district leaders, school leaders, and the work of classroom teachers, ultimately.

This study served to add to the empirical studies initially emerging in this realm by using the Rigby IRI, which is a different form of an IRI from those studied previously, with a significantly larger sample size, across multiple cohorts over a five year span, as well as by added layers of additional questions that sought not only to verify correlations with a differing IRI (Rigby), but to provide a more complete framework of the range of reported levels at various designator categories on the SOL (fail, proficient, pass advanced). It further expands on the current extant literature to evaluate the movement or progress of schools to demonstrate the level to which they are accomplishing the implied and stated goal of reducing reading failure across the school year as well as provides a barometer check on the systems and uses of these reading measures by school leaders.

Summary

While the United States and the Commonwealth of Virginia appear to overwhelmingly operate from a standards-based approach to instruction and use standardized assessment, New Zealand, alternatively, has enacted a framework that applies an acceleration model of teaching and learning that personalizes reading instruction in such a way as to accelerate reading progress and equalize the starting point for students of diverse backgrounds in the primary years. A large part of the studies found and explored for this literature review contain one main distinction and that is a focus on the predictive value of reading assessments to inform either instruction or
support a theory of literacy learning. In summary, Chapter 1 introduced the major constructs for the role of school leaders in using data to make instructional decisions along with a synopsis of reading assessment in the United States. Taking what the extant literature suggests for a data-driven school leader role with an emphasis on data based goal setting, developing teachers’ decision making capacity with regards to data usage, encouraging teachers to provide precise instruction to each student based on student data and providing a focus on improving instruction based on data (Sun et al., 2016), the impetus is present for exploring the implications of the IRI against the SOL and the leader practices of data monitoring systems specifically towards a literacy model that supports dynamic instructional decision making. To some experts in leadership studies, school leadership matters as much as teacher quality (Reeves, 2009; Sahlberg, 2013), and it is believed that “nothing much of significance happens related to improved schoolwide literacy achievement,” without strong principal leadership (Routman, 2014, p. 182).
Chapter 3: Methodology

Overview and Purpose

Situated within the “Use” branch of research methods (Mertens & Wilson, 2012), a pragmatic paradigm places emphasis “primarily on data that are found to be useful by stakeholders” and promotes the use of mixed methods (p. 88). By gaining knowledge in the pursuit of desired ends, the focus is placed on how research and data can be used as well as the results of that use. This study’s goals fit into the pragmatic realm by offering school leaders, as field practitioners, research that serves to enlighten their decision-making (Mertens & Wilson, 2012). While some theorists within the Use branch may lean more towards the program evaluation application of a pragmatic perspective, this study aims to meet similar aims as a program evaluation by researching an area that is critical to student achievement outcomes and school leaders’ decision making in the field of education.

The intended purpose of the study was to glean insight regarding the implementation and use of multiple measures to assess reading comprehension in third graders to better inform school leaders’ use of data to drive instructional decision making. This exploratory study included a correlational research design using a predictor variable (Rigby Informal Reading Inventory) with the criterion variable (Virginia Standards of Learning Assessment), and incorporated descriptive statistics of archival data for a specific student population.
Since there are scholars who make the case that school leadership is second only to classroom teachers as impactful in educational outcomes (Leithwood et al., 2008), and a data-driven leadership theory is prevailing in educational leadership, it is logical to overlay these concepts with those of being a literacy leader utilizing a dynamic approach to leading and literacy instruction to help close critical reading gaps by the third grade. The survey portion of the study explored principals’ use of data monitoring systems along with the level of leadership practices that foster a dynamic approach for the use of an informal reading inventory in either the quantitative and/or qualitative realms. It is, in essence, exploring if our espoused goals for closing literacy gaps by third grade are being born out in student outcomes, supported or undermined by our actual practices as implemented. This will be shown through the statistical analysis of the reading data and from the survey of how principals self-reported on their data-driven school leader practices in a literacy framework. The main question is are we implementing the use of assessment practices in a method that is counter-productive to our goals and outcomes? This study explored the sources of data pertinent to this issue and reports on findings to inform instructional leader practices.

**Research Questions**

1. How are the two main types of reading measures used in the state of Virginia correlated?

   a. To what extent are the Rigby Informal Reading Inventory (IRI) fourth quarter scores and state Standards of Learning (SOL) Reading scores correlated for a sample of third grade students in a midsized suburban district?
b. What is the range of Rigby IRI scores of students who passed the SOL at the proficient level? At the advanced level? At the fail level?

2. To what degree do outcomes on the Virginia third grade reading SOL show evidence of student acceleration towards closing a reading achievement gap?
   a. What are the proportions of change in categorizations of student reading by grade level (below, on, above) as measured by the Rigby IRI from beginning to end of the third grade year for the five years of cohorts sampled?
   b. To what degree do the number of reading levels that students move, as assessed by the Rigby IRI, differ among students classified as fail, pass, or pass advanced on end-of-year VA SOL Reading tests across the school year for a sample of third grade students in a mid-sized suburban district?

3. Do school leader data-driven practices reflect structures within a literacy framework that foster dynamic instructional practices?
   a. To what degree do school leaders identify practices within a data monitoring system in the district as fostering the use of the quantitative and/or qualitative information provided in an IRI?
   b. To what degree do school leaders identify their leadership structures as fostering the use of the fixed instructional practices through quantitative leveling information (benchmarking) provided by an IRI?
   c. To what degree do school leaders identify their leadership structures as fostering the use of dynamic instructional practices through personalization of reading instruction using qualitative information from an IRI?
d. To what degree do school leaders identify their leadership structures as fostering the use of dynamic instructional practices through regrouping or flexible grouping of students for reading instruction?

Research Design

The first two research questions and subsets of questions make use of correlational analysis methods investigating archival reading assessment data from five years of cohorts of third grade students from a midsized school district in Eastern Virginia. The third question and subset of questions used a survey to explore the school leader’s use of IRIs as a measure of reading progress to make data-informed decisions. By leveraging the correlational relationships in the archival data along with the school leader survey results, the study’s findings are strengthened by information obtained statistically to investigate the research questions as well as by investigating the practices reported by school leaders. Thus, the investigation of the statistical relationships between measures and how well they evidence reading gains among third graders is augmented by school administrators’ organizational leadership structures towards the use of IRIs as reading assessment data. From the survey, I gathered insight from participating principals regarding the use of multiple reading assessments and how their use informs their decisions as a school leader. These data contribute to understanding the overarching concept of tying the reading assessment measures to making school-level decisions.

Participants

Participants were drawn using a convenience sampling method. Archival data harvested from a district-maintained electronic database system over a five-year span was utilized for the statistical analyses. Sample archival data were comprised of reading
assessment scores for cohorts of third grade students who attended all ten elementary schools located within the selected mid-sized suburban school district from 2009-2014. These students participated in the VA SOL third grade reading assessments and IRIs from each of four quarters in the third grade year.

The sample of student cases for this study included 2,906 third grade records from five years of third grade classes across all 10 elementary schools in the Star School District. This was a robust sample size for the analyses. G*Power 3.1.9.2 was used to determine the minimum sample size necessary for statistical validity (Faul, Erdfelder, Buchner, & Lang, 2013). I calculated the minimum sample size necessary to conduct a Pearson \( r \) correlation and ANOVA. For the Pearson \( r \) correlation with an alpha of .05 and a power of .80, the minimum sample size for statistical validity is 67 students. For an ANOVA with a medium effect size (\( f_d = 0.25 \)), an alpha of .05, a power of .80, and 3 groups, the minimum sample size necessary for statistical validity is 159 students. The statistical analysis with the more rigorous sample size requirement (159) was used to set the necessary sample size for the students.

For the survey of school leaders, both principals and assistant principals were included and the convenience sample offered up to 22 possible participants within the leadership pool with a 41% return rate of nine fully completed surveys. An additional six surveys were started, but not completed, and therefore, were not usable in the analysis. Background questions captured a range of experience in school leadership roles for the respondant pool from 1 year to 15 with an average of 6 years, and a median of 5 years. School leader experience in the school district under study averaged at 13 years. All school leaders, except for one, in this respondent pool reported having some experience
teaching at the elementary grade level, and all but two as having been trained to administer an IRI. Seven school leaders indicated experience with administering an IRI previously, and two hold a reading specialist license.

Data Sources

Data from two standardized measures and surveys of elementary principals and assistant principals were used in this exploratory study. Additional data was gathered using a survey of elementary school principals and assistant principals.

The Rigby Informal Reading Inventory. The Rigby IRI provides a system of leveled fiction and nonfiction tests to assess each student’s reading level (Smith et al., 2008). They are considered a more formal and comprehensive assessment than just taking a running record. The texts are considered meaningful and tested to “guarantee the suitability and readability for a particular level” (Smith et al., 2008, p. 7). The assessment can be used as both a benchmark tool and a progress monitoring tool; thus, the publishers recommend obtaining three progress monitoring data points across the year. The sample in this study was assessed four times across the school year with one IRI testing session given at each quarter of the year.

The Rigby IRI is most commonly used by classroom teachers and reading specialists in the kindergarten through fifth grade at the elementary level of schooling. It has a multifaceted purpose of use within the educational setting. The assessment can provide individual student strengths and needs, allows teachers to gain information in placing students at appropriate reading levels for instruction, to monitor observable reading behaviors demonstrated during test administration, measure comprehension, oral reading, and fluency of reading. The instrument system includes 60 benchmarked books
at 30 different levels with two at each level. The instrumentation comes with corresponding retelling response sheets and reading record sheets, comprehension check pages, reading behavior analysis check sheets, reading progress portfolios, and a data management computer program tool. The kits are consistent from one user to another with the same assessment guides, testing materials and books. The types of books include a variety of genres across all levels for both expository and narrative styles. Much can be obtained about a student’s reading behaviors by noting the oral reading habits and behaviors, literal comprehension versus inferential comprehension questions, the level of accuracy of the reader, a self-correction rate on errors made and corrected during reading, reading speed, pace, phrasing, and intonation. The levels of the Rigby IRI are considered criterion referenced based on high-frequency words, sentence construction, meaning, logic, and the Fry Readability factors. The levels were further tested with “children of an appropriate reading age to guarantee the suitability and readability of the text” (Smith et al., 2008, p. 7). There is a kit 1 and a kit 2 which provides an option for test-retest reliability measurement by using two different assessments designed for the same level to help ensure proper identification of the student’s reading level and a reliable assessment of the student’s reading abilities.

There is limited information regarding the technical adequacy of this measure; however, some available information shows variability in the reliability and validity of the measure. The variance stems from each text’s own “level, structure, type, and topic” creating some variability in running record scores (Fawson, Ludlow, Reutzel, Sudweeks, & Smith, 2006, p. 124). To attempt to account for consistent administration and standardized assessment practice, the school district used for this study requires
participation in annual training for all teachers who are required to administer the IRI. The training is provided by the content area experts, who are certified reading specialists. The reading specialists use the assessment guide provided with the IRI kit to break out the information for a step by step presentation and procedural review each year with the staff. The reading specialists create the presentation and review of the assessment materials together as a group and then deliver it to each of their own schools using the common structure, language, and resources. Teachers can practice, ask questions, and be observed giving the IRIs as part of the training for consistent and standard application of the testing materials. New teachers or teachers requesting additional help in administration can receive individual support and coaching by the reading specialist as requested. Since the data are entered quarterly, reading specialists can follow up with teachers if outlier data are found. Teachers are then coached on ensuring they find the highest level read with success for instruction.

The texts used in the lower grades, including kindergarten, first and second, are very much like the early reader stories this age of student might read in school or at home. For the upper grades of third, fourth, and fifth, the texts are made to be much shorter stories than the typical chapter books, articles or picture books read at those grades. The Rigby IRI does not include a writing component. The assessment is done in a one-on-one setting with a child and teacher and begins by the administrator providing the book, title, cover illustration and a standard introduction to the story provided with each book. The student reads the story independently and then retells the story to the administrator who takes notes while the student talks. Standard prompts and questions are provided for systematic testing. Using the scoring criteria provided in the assessment
system, the administrator scores the retelling to obtain a score that indicates the reading at that level by the student is either instructional, independent or frustrational. A level deemed independent is the highest level at which a student can read without experiencing any frustration or a level of errors that would impede understanding of the passage read. An instructional level denotes the highest level at which a child can read with minimal errors and maintain comprehension and may require the assistance of an “expert other,” in terms of Vygotskian theory (Clay & Cazden, 1990). Lastly, a frustrational level is considered too hard for a reader to maintain accuracy, fluency and meaningful comprehension. Criteria for determining the levels of independence, instructional and frustrational include measuring word recognition by the word count of the passage read and the scoring of the reader’s response to the comprehension questions/prompts and retelling. An independent level requires the student to have a 98% to 100% word accuracy and a comprehension rate of 90% to 100%. To obtain the comprehension score, the number of questions in the retelling/question portion of the assessment session is rated using a rubric and the final percentage rate of comprehension is calculated. The teacher further probes the student who scores instructional with 3-5 comprehension questions about the text at both the literal and inferential levels. For the instructional level, a student may score within a 95% to 97% oral reading accuracy range and maintain a comprehension score that falls within the 75% to 89% range. Frustration level is earmarked by a breakdown in the oral reading accuracy with the calculation falling below 90%, and the comprehension breaking down at 50% or lower in understanding what is read. While this description includes a synopsis of the process for understanding the
scoring and analysis, a summary of the basic procedures for administering the Rigby IRI can be found in Table 3.
Table 3

Summary of Basic Procedures for Administering a Rigby IRI

<table>
<thead>
<tr>
<th>Step 1 Retelling</th>
<th>Teacher reads book title and the orientation provided and does not expand upon the content during the book introduction.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Student reads the text alone (silently or aloud). Teacher follows written prompts included to prepare the student for doing a retelling of the story read.</td>
</tr>
<tr>
<td></td>
<td>Student responds and is rated for the retelling/comprehension check.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 2 Reading Record</th>
<th>Teacher uses the standard procedures to obtain reliable assessment information and records the responses during the oral reading of the story.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Student reads aloud and teacher takes the record of reading.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 3 Comprehension</th>
<th>Teacher uses standard prompts to initiate oral comprehension questions and records student responses for rating.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Student responds and teacher records and rates responses using standards provided.</td>
</tr>
</tbody>
</table>

| Step 4 Analysis of Information | Teacher uses the standard procedures to score and complete an analysis of all the information collected during the reading assessment. |

*Note.* While this is a summation of the basic overall steps, the full scope of administering, scoring and analyzing the Rigby IRI is found in the 55 page manual (Smith et al., 2008).
The IRI process continues until the highest level readable at the instructional level is found, so students may be asked to do the process with multiple texts in one sitting or in multiple sittings until the instructional level is obtained. The testing books are leveled from a 1 to 30, which then correspond to a grade level range, as shown in Table 4. Students are assessed at the end of each marking period for a growth measurement across the grade level school year.

Table 4

*Rigby Informal Reading Inventory Levels and Grade Level Equivalences*

<table>
<thead>
<tr>
<th>Reading Inventory Levels</th>
<th>Grade Level Equivalency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – 6</td>
<td>Kindergarten</td>
</tr>
<tr>
<td>7 – 18</td>
<td>First Grade</td>
</tr>
<tr>
<td>19 – 22</td>
<td>Second Grade</td>
</tr>
<tr>
<td>23 – 25</td>
<td>Third Grade</td>
</tr>
<tr>
<td>26 – 28</td>
<td>Fourth Grade</td>
</tr>
<tr>
<td>29 – 30</td>
<td>Fifth Grade</td>
</tr>
<tr>
<td>30 +</td>
<td>Sixth Grade and beyond</td>
</tr>
</tbody>
</table>

*Note.* IRI levels and grade level equivalences are based on the widely accepted and used system of leveling from Fountas & Pinnell (2012).
**Virginia State Standards of Learning Test.** The Virginia Department of Education (VDOE) considers the SOL Test results as a method for assessing each individual child’s achievement. The SOL also measures communities of students’ degree of progress in meeting the commonwealth’s expectations in each content area tested. The state Board of Education uses the information to identify schools that are in need of assistance and support as well as using the assessments as an “objective means for measuring achievement gaps between student subgroups and for determining the progress of schools, districts, and the state toward closing these gaps” (Virginia Department of Education, 2015, p. 1).

Assessment schedules are set by the state department for all districts, and schools administer the tests during these times in the spring, summer, and fall. The third grade reading test is given each spring. Raw scores are initially obtained and then converted to a scale score. Raw scores initially indicate the number of points a student receives for correctly answering questions and the scale score is a conversion score to a common scale; thus, allowing for numerical comparison of student scores across different years and versions of the tests. Scores are then reported with performance levels that the state has established into three categories: fail/basic (399 or under), pass/proficient (400-499), and pass/advanced (500-600). These performance level descriptors feature what students should know and be able to do in the corresponding grade level being assessed. A standard setting committee works to establish the cut scores that align with the performance level descriptors. These committees contain educators recruited by the VDOE based on qualifying criteria.
Internal consistency reliability was assessed to establish the ability of the questions to work together to reliably represent a construct (George & Mallery, 2012). In the 2014-2015 technical report, there were 103,027 students assessed on the third grade reading SOL with 99% of those being done via online web-based assessments. The reliability coefficient established for third grade reading SOL core 1 and 2 on the spring 2015 constructed test were 0.90 and 0.88, respectively. These Cronbach’s alpha values indicate good and excellent reliability (George & Mallery, 2012). The VDOE also reports decision consistency and decision accuracy in accordance with Livingston and Lewis’s (1995) guidelines to account for decision misclassifications within the performance levels. Comparable levels of decision consistency and accuracy have been established in line with Livingston and Lewis’ work for a consistency rate in Core 1 third grade reading as 0.89 and 0.87 in core 2 from 2015 results. Further, the VDOE claims in the technical report that validity is established through multiple means based on test content, response processes, and internal structure. Given these multiple means as a whole, the VDOE suggests the SOLs as having valid measurement of the grade level content for third grade reading.

Principal data use survey. Surveys provide a method for collecting standardized information from a sample to draw generalizations back to the target population (Stern, Bilgen, & Dillman, 2014). Within this study, the survey provided information from school leaders serving as principal or assistant principal from the district providing archival IRI assessments and SOL reading scores. The survey provided an overlapping piece of information regarding the use and application of data-driven decision making within the school setting. The goal was to measure the degree to which principals in this
district rate their use of various data-driven monitoring systems, their use of specific practices that support literacy acceleration, and the qualitative features of the Rigby IRI used in their district.

The survey protocol used a standardized structure in which the same questions were presented to all survey respondents; thus, allowing for a more readily analyzed data set. Once a draft of the protocol was developed, a panel comprised of three reading content area field experts and three school leaders, in order to provide feedback on the content appropriateness of the questions, along with the readability and wording of the survey questions. The panel reviewed the survey to ensure the survey aligned with the research questions the survey was developed to address. The panel suggested improvements to the survey for consideration by the researcher. Utilizing feedback from the panel, edits were made to improve the survey. Five of the six panel experts asked to participate took the survey and provided feedback in the form of written or a phone call. The main edits suggested and utilized regarded improvements in grammar and wording for ease of understanding and measuring the intended information requested, movement of demographic or background items to the top of the survey, and considering a different response from “I do not know.” Four of the five panel members believed the “I do not know” response could be off-putting to school leaders as a reflection on them or their school. Even though the survey is anonymous, the panel feedback indicated that this response choice may not be received well and could skew the intended information. For additional clarity and information, several background questions were added to obtain a clearer picture of the background and experiences of the school leaders with the use of an IRI. Once edits were made from the expert panel feedback process, the survey was resent
to 2 of the expert panel members with one being a school leader and one being a literacy practitioner. Additionally, the survey was then reviewed by a professor of statistics and the director of e-learning at the university level who are versed in the Qualtrics online survey software package used by the college. This team of two provided tips and advice on the look and feel of the design of the survey within the Qualtrics software system. The resulting version of the survey is believed to be the most user friendly version that will more readily obtain input from the sample population. The time to take the survey was reduced from over 10 minutes to an approximation of 7 minutes.

The survey protocol (see Appendix) contains 27 questions across four main strands. There are six background questions, 20 questions with a Likert scale rating, and one open-ended question. One strand is designed to capture school leader use of data monitoring systems with five questions. The second strand assesses school leader practices that fall in a fixed instructional approach with five questions. The third strand contains six questions all pertaining to dynamic instructional practices that fall in the personalized instructional approach category through the use of the IRI qualitative data. The fourth strand contains four questions addressing dynamic instructional practices that include flexible grouping or regrouping practices supported by the school leader. The strands and the questions within each thread were crafted to provide an appropriate breadth and depth of data for analysis to sufficiently address the intended research questions. The open-ended question provided participants an opportunity to offer responses that may illuminate data trends not anticipated by the researcher. Table 5 shows a summation of the overall design matrix for the concepts in the survey.
Table 5

Survey Protocol Design Matrix

<table>
<thead>
<tr>
<th>Areas on Survey</th>
<th>Items Addressing Each Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of Data Monitoring Systems</td>
<td>7, 13, 19, 21, 25</td>
</tr>
<tr>
<td>Fixed: Benchmarking Practices based on IRI</td>
<td>9, 15, 17, 20, 26</td>
</tr>
<tr>
<td>Quantitative Information</td>
<td></td>
</tr>
<tr>
<td>Dynamic: Personalized Practices based on IRI</td>
<td>8, 10, 12, 14, 16, 23</td>
</tr>
<tr>
<td>Qualitative Information</td>
<td></td>
</tr>
<tr>
<td>Dynamic: Regrouping Practices based on IRI</td>
<td>11, 18, 22, 24</td>
</tr>
<tr>
<td>Qualitative Information</td>
<td></td>
</tr>
<tr>
<td>Other/Background</td>
<td>1, 2, 3, 4, 5, 6</td>
</tr>
<tr>
<td>Open-ended</td>
<td>27</td>
</tr>
</tbody>
</table>

Data Collection

Archival data for this research study were mined from an electronically maintained school district database and provided to the researcher with the assistance of the Manager of Application Software as well as the Superintendent, and the Public Relations Coordinator. The district database contains IRI scores from each of the four quarters of a school year and the Virginia SOL Reading outcomes for five years (2010-2014) of third grade cohort students from a mid-sized suburban district in the state of Virginia. Permission to access the information was granted by the Manager of Application Software, the Superintendent, and the Public Relations Coordinator. Student assessment data for the IRI and VA SOL was received in an Excel document for analysis. The dataset contained no student identification numbers and no additional identifying information was included in the file. Student names were not shared to protect the
confidentiality of student data. All data were previously collected by the school district for the purposes of student assessment. Ethical treatment of students was assured by the district prior to administration of the assessments. To further protect the rights of participants in subsequent handling of data, the data are maintained on a password-protected computer. The information was converted from Excel to SPSS (version 24) to conduct statistical analyses.

Principals and assistant principals received a recruitment letter via email. The recruitment email detailed the purpose of the research, the process of data collection, the benefits and risks of participation in the study, the voluntary nature of the study, and contact information for the researcher. A link was included to the survey instrument, which included a consent to participate segment within the online survey. A reminder email offered an additional opportunity to participate across a week long window for data collection. Studies on the use of web-based surveys indicate personalized email cover letters, follow-up reminders and pre-notification of the intent to survey in simple formats are factors that increase response rates. These methods were employed for the survey method. Utilizing the Qualtrics software survey application helped with providing easy access and responses by supplying a URL that was clickable and allowed respondents to fill out the survey online. Results were recorded immediately within the software and further analysis was then conducted beyond the software collection tool (Cook, Heath & Thompson, 2000). The survey content can be found in the appendix.

Data Analysis

Pearson $r$ correlations, chi-square test of independence, and Analysis of Variance were used within the study. Data were organized in Microsoft Excel and analyzed using
the Statistical Package for the Social Sciences (SPSS). Prior to analysis, descriptive statistics were calculated for the variables of interest and demographic information. To address research question 1a, “To what extent are the Rigby Informal Reading Inventory (IRI) and state Standards of Learning (SOL) Reading scores correlated for a sample of third grade students in a midsized suburban district?”, Pearson r correlations were calculated. The Pearson r correlation is used to measure the strength of association between two variables (Pagano, 2010). The relationship between the variables are represented by reporting r and p values. Correlation coefficients, r, represent the strength and direction of the relationship between the variables (Howell, 2010). A negative r value indicates an inverse relationship, where as one variable increases the other variable decreases. A positive r value indicates a direct relationship, whereas one variable increases, the other variable also increases. The significance, or p value, indicates if the results of the analyses are statistically significant. An alpha of .05 will be used in the analyses.

To address research question 1b, “What is the range of Rigby IRI scores of students who passed the SOL at the proficient level? At the advanced level? At the fail level?” descriptive statistics were used to describe the ranges of reading levels found at each of the levels for proficient, advanced, and fail.

For research question 2a, “What are the proportions of change in categorizations of student reading by grade level (below, on, above) as measured by the Rigby IRI from beginning to end of the third grade year for the five years of cohorts sampled?” a Chi-Square test of independence was conducted. The Chi-Square Test of Independence helps determine if there is an association between categorical variables and is considered a non-
parametric test. It was conducted to examine if differences exist in the proportion of students categorized as below, on, or above grade level as measured by the IRI. Additionally, the Chi-Square Test of Independence was conducted to assess differences in the proportion of students in each category, specifically, at the beginning and the end of the third grade year. An alpha value of .05 was set at .05 to determine if statistically significant differences exist.

To address research question 2b, “To what degree do the number of reading levels that students move, as assessed by the Rigby IRI, differ among students classified as fail, pass, or pass advanced on end-of-year VA SOL Reading tests across the school year for a sample of third grade students in a mid-sized suburban district?” an Analysis of Variance (ANOVA) was conducted. The ANOVA was used to determine if statistically significant differences exist in the number of reading levels students gained across the third grade year (Tabachnik & Fidell, 2012). Differences were assessed between groups, i.e. students categorized as fail/basic, pass/proficient, or pass/advanced. An alpha value of .05 was used to determine if statistically significant differences existed. Since statistically significant differences existed, a post hoc analyses of the Tukey’s test, was conducted to determine where differences existed between groups.

A survey approach was best suited for gleaning data related to research question 3 to facilitate an exploration of principal’s use of data systems and organizational structures with the use of the IRI data for instructional decisions unrestrained from any predetermined assumptions (Marshall & Rossman, 2011). To address question 3, “Do school leaders identify practices within a data monitoring system in the district as fostering the use of the quantitative and qualitative information provided in an IRI?” a
A web-based survey protocol was used to obtain school leader input on leadership activities. The survey contained strands to obtain responses that fall within identified categories of leadership practices. The questions were broken up into categories to identify practices toward a data monitoring system, fixed instructional practices, dynamic instructional practices, an open-ended response item and a few background questions. These were answered under the following subpart questions, “To what degree do school leaders identify their leadership structures as fostering the use of the fixed instructional practices through quantitative leveling information (benchmarking) provided by an IRI?”; “To what degree do school leaders identify their leadership structures as fostering the use of dynamic instructional practices through personalization of reading instruction using qualitative information from an IRI?”; and “To what degree do school leaders identify their leadership structures as fostering the use of dynamic instructional practices through regrouping or flexible grouping of students for reading instruction?” Findings are described in detail for school leader responses on the categories. Table 6 features a summary of the research questions, theoretical constructs, data sources, and analyses.
<table>
<thead>
<tr>
<th>Research Question</th>
<th>Theoretical Construct Alignment</th>
<th>Data Source</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. a. To what extent are the Rigby Informal Reading Inventories (IRI) and state Standards of Learning (SOL) Reading scores correlated for a sample of third grade students in a midsized suburban district?</td>
<td>Assessment theory and use of SOL tests and IRIs</td>
<td>Archival data on IRI and VA SOL scores</td>
<td>Pearson r correlations</td>
</tr>
<tr>
<td>1. b. What is the range of IRI scores of students who passed the SOL at the proficient level? At the advanced level? At the fail level?</td>
<td>Framework for reading level trends as categorized on the SOL</td>
<td>Archival data on IRI and VA SOL scores</td>
<td>Descriptive statistics</td>
</tr>
<tr>
<td>2. a. What are the proportions of change in categorizations of student reading by grade level (below, on, above) as measured by the Rigby IRI from beginning to end of the third grade year for the five years of cohorts sampled?</td>
<td>Acceleration or closing the 3rd grade reading achievement gap</td>
<td>Archival data on IRI</td>
<td>Chi-Square Test of Independence</td>
</tr>
<tr>
<td>2. b. To what degree do the number of reading levels that students move, as assessed by the Rigby IRI, differ among students classified as fail, pass, or pass advanced on end-of-year VA SOL Reading tests across the school year for a sample of third grade students in a mid-sized suburban district?</td>
<td>Acceleration or closing the 3rd grade reading achievement gap</td>
<td>Archival data on VA SOL</td>
<td>ANOVA, Tukey HSD for post hoc analysis</td>
</tr>
<tr>
<td>3. a. To what degree do school leaders identify practices within a data monitoring system in the district as fostering the use of the quantitative and/or qualitative information provided in an IRI?</td>
<td>School leaders making data informed decisions</td>
<td>Principal Surveys</td>
<td>Descriptive Statistics</td>
</tr>
</tbody>
</table>
3. b. To what degree do school leaders identify their leadership structures as fostering the use of the fixed instructional practices through quantitative leveling information (benchmarking) provided by an IRI?
3. c. To what degree do school leaders identify their leadership structures as fostering the use of dynamic instructional practices through personalization of reading instruction using qualitative information from an IRI?
3. d. To what degree do school leaders identify their leadership structures as fostering the use of dynamic instructional practices through regrouping or flexible grouping of students for reading instruction?

Assumptions of the Study and Ethical Considerations

Assumptions. When considering the quantitative methods for this study, it is assumed the training of teachers in using the IRI procedures reported and described were consistent in producing reliable results from the IRI, even though interrater reliability was not conducted by this researcher. In a study dated study from 1970, Dunkeld (1970) found a review of several informal reading inventories contained interrater reliability ranging from .92 -.99 when looking at oral reading accuracy and .92 -.99 for comprehension. His study further suggested reported reliabilities between standardized tests and IRIs are comparable if they are administered in a manner as prescribed and scored against objectively defined categories (Dunkeld, 1970). Further meta-analyses of IRIs done by Pikulski and Shanahan (1982) found consistent and structured training with post-training checks provided to all teachers yielded good interrater reliability among
most IRIs reviewed. Assuming these traits, as described for this school district are reportedly true, then the IRI results to be used in this study are presumed reliable as well. In the survey portion, considerations are given regarding the honesty and openness of participants in the survey process on school leaders’ perceptions of the use of IRIs and SOL predictions in making informed leadership decisions regarding instruction. Given the role of the researcher in this study as a former colleague having worked in a trusted leadership role amongst many of the leader participants included in the convenience sample for this study, it is assumed they provided honest responses without fear of evaluation pertaining to job security.

**Ethical Considerations**

All data for the correlational research component comes from an electronic database maintained by the Manager of Application Support for the school district. Given the approval from the dissertation committee and the College of William & Mary’s Institutional Review Board (IRB), all standards were followed to protect the participants’ identities. Once IRB permission was established, a formal written request, per the school district’s policy and procedures, was submitted for formal approval to work with the Manager of Application Support to obtain an electronic copy of the data for this study. Initial data in the database system contained student state testing identifier numbers. The state testing identifier number was removed and a randomly assigned numerical system replaced prior to conducting the statistical analyses outlined. No other personal or identifiable information was contained within the data obtained. Student information, school information, and school district information is not referenced at any time in this study, other than in the form of descriptive statistics about the district, and it will not be
possible to track results back to individual students within the study. Furthermore, survey data of elementary school principals and assistant principals were anonymous and protected.
Chapter 4: Results

In order to investigate leadership, literacy assessments, and acceleration for closing the reading achievement gap by third grade, multiple methods of research were employed.

Since the study focuses on reading outcomes from fall to spring, and the level of movement between the beginning of the year and the end of the year as part of the instructional decision making framework for literacy and leadership, it was important to isolate cases with quarter 1, quarter 4 data, and SOL outcomes. Additionally, outlier values were present in the data set. Scores containing numerical values outside of the range for the IRI or the SOL were removed to prevent skewed statistical analysis. Quarter 1 cases contained 3,252 in total with two removed for a reported score below the range available on the Rigby IRI, and 13 removed as scores above the range available on the Rigby IRI. Quarter 4 IRI scores contained 3,931 cases with one case entered below the available range, and 18 values entered higher than the maximum score of 30. Additionally, for the SOL scores, 24 scores were coded using state codes for invalid scores and one case contained a zero score, all of which were removed. Using the SPSS software features, out of zone values were set as missing values and all three variables were defined within the appropriate ranges for the Rigby IRI and SOL scores. This resulted in a total of 2,906 cases for further statistical analysis, which follows for each question posed below. The mean values were 24.38 (SD=3.156) for quarter 1, 26.78
(SD=2.610) for quarter 4, and 467.63 (SD=72.226) for the SOL as shown further in Table 7.

To evaluate whether or not missing data had an impact on the nature of the sample being studied a further analysis was conducted. Analyses of the data set with all available data included versus a restricted data set with list wise deletion due to missing variables were completed. There was no significant difference in the mean scores computed for Rigby IRI data in quarter 1 or quarter 4 between an analyses of the full data set versus the restricted data set. However, there was a seven point difference in the mean SOL score in the data set with the missing cells stemming from the loss of higher scores for cases where students were missing either quarter 1 or quarter 4 IRI data, but for whom an SOL score existed.

Table 7

Descriptive Statistics on Rigby IRI and SOL Scores for Three Variables

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIGBY_Q1</td>
<td>2906</td>
<td>3</td>
<td>30</td>
<td>24.38</td>
<td>3.15</td>
</tr>
<tr>
<td>RIGBY_Q4</td>
<td>2906</td>
<td>3</td>
<td>30</td>
<td>26.78</td>
<td>2.61</td>
</tr>
<tr>
<td>SOL_SCORE</td>
<td>2906</td>
<td>241</td>
<td>600</td>
<td>467.63</td>
<td>72.22</td>
</tr>
</tbody>
</table>
Research Question: 1: How are the two main types of reading measures used in the state of Virginia correlated?

More specifically, research question 1a. was expanded to consider to what extent the Rigby Informal Reading Inventory (IRI) fourth quarter scores and state Standards of Learning (SOL) Reading scores were correlated for a sample of third grade students in a midsized suburban district? The Pearson product-moment correlation coefficient (see Table 8) was computed to assess the relationship between the reading levels obtained on a Rigby IRI in quarters 1 and quarter 4 in third grade and the third grade reading SOL outcomes. There was a significant moderate correlation between quarter 4 and the SOL, \( r = .568, n = 2906, p < 0.01 \). Additionally, there was a moderate Pearson \( r \) correlation between quarter 1 Rigby scores and the SOL, \( r = .510, n = 2906, p < 0.01 \). Further analysis showed a strong correlation between quarter 1 and quarter 4 Rigby scores, \( r = .807, n = 2906, p < 0.01 \). Thus, reading scores on the Rigby IRI were significantly correlated not only with the reading SOL outcomes, but also between quarter 1 reading and quarter 4 reading outcomes.

Table 8

Rigby Quarter 1, Quarter 4 and Third Grade Reading SOL Pearson \( r \) Correlations

<table>
<thead>
<tr>
<th></th>
<th>SOL_SCORE</th>
<th>RIGBY_Q1</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIGBY_Q4</td>
<td>2906</td>
<td>.568**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.807**</td>
</tr>
<tr>
<td>SOL_SCORE</td>
<td>2906</td>
<td>.510**</td>
</tr>
</tbody>
</table>

** \( p < 0.01 \), two-tailed
Research question 1b investigated the range of Rigby IRI scores of students at each categorization for pass, pass advanced, and fail. Nearly half (49%, n = 1,417) of the 2,906 students fell into the pass proficient range on the third grade reading SOL assessment. The Rigby range for all students who scored at each level of proficiency on the SOL were associated with an array of levels. These equated to the following: Fail: Rigby levels 18-28; Pass: Rigby levels 25-30, Pass Advanced: Rigby levels 25-30. Within the pass advanced scores, the majority fell at the level 27-30 range, with a few (9%, n = 88) at level 25. While the range of Rigby IRI scores were the intended outcome for this research question, additional information found within the cross tabulations of the Rigby quarter 4 scores and the SOL level of categorization for scores provided additional clarity on the trends and patterns found within the data set analyzed for question 1.b. as discussed in further detail in the next section.

A Rigby IRI score of 25 for the end of the third grade year is considered on grade level. Of those who were considered performing at a reading level of 25 for the fourth quarter on the Rigby, 25% of those students were categorized as a fail on the state reading SOL test, 61% were categorized as pass proficient, and 14% were considered pass advanced. Of all students who scored a Rigby IRI level 26, which is entering grade four, 20% fell into the fail range, 60% fell into the pass proficient range on the SOL, and 20% in the pass advanced range. At the Rigby level of 27, which equates to quarter 2 of grade 4, 9.6% were in the fail range, 56% were in the pass proficient range, and 34% fell in the pass advanced range on the SOL reading measure for third grade. Thus, 90% of all students performing at the mid-year fourth grade level passed the SOL. Of all those who scored a Rigby IRI level 28 and level 29, which is considered end of grade four and grade
five equivalent, respectively, 96% passed proficient or higher. Once students reached a Rigby level 30, 99% passed proficient or higher. Table 9 further shows the breakdown for the cross tabulation of SOL categorization and reading level based on the quarter 4 Rigby IRI reading level.
Table 9

Cross Tabulation of SOL Categorization and Rigby IRI Reading Level

<table>
<thead>
<tr>
<th>Category</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Of Students (n = 57) Who Scored Rigby IRI Level 20, 21, or 22 (Second Grade Level)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75% Fail</td>
</tr>
<tr>
<td></td>
<td>21% Pass Proficient</td>
</tr>
<tr>
<td></td>
<td>4% Advanced</td>
</tr>
<tr>
<td>Of Students (n = 78) Who Scored Rigby IRI Level 23 (Beginning Year Third Grade Benchmark)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>67% Fail</td>
</tr>
<tr>
<td></td>
<td>32% Pass Proficient</td>
</tr>
<tr>
<td></td>
<td>1% Advanced</td>
</tr>
<tr>
<td>Of Students (n = 150) Who Scored Rigby IRI Level 24 (Mid-Year Third Grade)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>52% Fail</td>
</tr>
<tr>
<td></td>
<td>39% Pass Proficient</td>
</tr>
<tr>
<td></td>
<td>9% Advanced</td>
</tr>
<tr>
<td>Of Students (n = 520) Who Scored Rigby IRI Level 25 (End of Year Third Grade Benchmark Level)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>25% Fail</td>
</tr>
<tr>
<td></td>
<td>61% Pass Proficient</td>
</tr>
<tr>
<td></td>
<td>14% Advanced</td>
</tr>
<tr>
<td>Of Students (n = 511) Who Scored Rigby IRI Level 26 (Beginning Fourth Grade)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20% Fail</td>
</tr>
<tr>
<td></td>
<td>60% Pass Proficient</td>
</tr>
<tr>
<td></td>
<td>20% Advanced</td>
</tr>
<tr>
<td>Of Students (n = 481) Who Scored Rigby IRI Level 27 (Mid-Fourth Grade)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.6% Fail</td>
</tr>
<tr>
<td></td>
<td>56% Pass Proficient</td>
</tr>
<tr>
<td></td>
<td>34% Advanced</td>
</tr>
<tr>
<td>Of Students (n = 320) Who Scored Rigby IRI Level 28 (End of Fourth Grade)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4% Fail</td>
</tr>
<tr>
<td></td>
<td>50% Pass Proficient</td>
</tr>
<tr>
<td></td>
<td>46% Advanced</td>
</tr>
<tr>
<td>Of Students (n = 113) Who Scored Rigby IRI Level 29 (Beginning Fifth Grade)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.4% Fail</td>
</tr>
<tr>
<td></td>
<td>50.4% Pass Proficient</td>
</tr>
<tr>
<td></td>
<td>45.1% Advanced</td>
</tr>
<tr>
<td>Of Students (n = 639) Who Scored Rigby IRI Level 30 (End of Fifth Grade)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1% Fail</td>
</tr>
<tr>
<td></td>
<td>32% Pass Proficient</td>
</tr>
<tr>
<td></td>
<td>67% Advanced</td>
</tr>
</tbody>
</table>

Note. 37 students fell at level 19 or lower (first grade) with 27 in the fail category, while 10 were pass proficient and none pass advanced. The 10 that passed were spread across 13 different levels ranging from beginning kindergarten to end of first grade, and thus, are considered outliers in the data.
Research Question 2: To what degree do outcomes on the Virginia third grade reading SOL show evidence of student acceleration towards closing a reading achievement gap?

More specifically, research question 2a asked, what are the proportions of change in categorizations of student reading by grade level (below, on, above) as measured by the Rigby IRI from beginning to end of the third grade year for the five years of cohorts sampled? Using the Rigby reading level equivalences, student outcomes were investigated for level of movement or change in the categorizations for being considered on grade level, below grade level, or above grade level. A chi-square test of independence was performed to examine the relation between the changes in levels in Rigby quarter 1 and quarter 4 reading levels and the categorization of on, below, and above grade level reading status. There is a strong relationship between where a student ended third grade based and where they started ($\chi^2(4) = 11349.364$, p < .0.001).

This analysis revealed that 30% of students (n = 883) changed categorization between quarter 1 and quarter 4 of the third grade year. Of those that changed categorizations, 23% made upward movement, 7.5% made downward movement. A solid majority, 70%, stayed in the same category in which they began the year. Thus, 664 students went up, 219 students went down, and 2023 students stayed the same.

Of all those students who were categorized as below grade level in quarter 1, 52% stayed the same, 29% made a change to an on grade level status, and 19% made a change to an above grade level status. Considering only those students who were categorized at the on-grade level range in quarter 1, 9% moved to a below grade level status, 33% stayed on grade level, and 58% moved to above grade level. For those students
categorized in the above grade level category in quarter 1, 91% stayed above in quarter four and 8% dropped to on grade level, and 1% moved to below grade level.

Furthermore, research question 2b looked at to what degree the number of reading levels students moved, as assessed by the Rigby IRI, differed among students classified as fail, pass, or pass advanced on the end-of-year VA third grade SOL Reading test across the school year for a sample of third grade students in a mid-sized suburban district? A one-way between subjects ANOVA was conducted to compare the effect of categorization levels (pass proficient, above, fail) on the SOL test and the amount of change in reading levels between quarter 1 and quarter 4 on the Rigby IRI. There was a significant difference at the p<.05 level between being categorized as pass advanced, pass proficient, or fail, and the number of reading levels students moved from quarter 1 to quarter 4 on the Rigby IRI \(F(2, 2093) = 5.81, p = .003\). Post hoc comparisons using the Tukey HSD test indicated the mean score for the pass advanced scoring category (M = 2.25) was significantly different than the pass proficient and fail categories (M = 2.46, 2.55). However, the pass proficient category (M = 2.46) did not significantly differ from the fail category (M = 2.55). The data revealed that students classified as fail or pass proficient on the SOL moved at approximately the same rate or number of IRI reading levels between quarter 1 and quarter 4 while the students classified as pass advanced on the SOL moved at a statistically significant different rate, which was lower. The mean number of reading levels moved on the IRI for pass proficient was 2.46 and 2.55 for those in the fail category while those who scored in the pass advanced category moved an average of 2.24 IRI reading levels.
Research Question 3: Do school leader data-driven practices reflect structures within a literacy framework that foster dynamic instructional practices?

Research question 3 also considered to what degree school leaders identified practices within a data monitoring system in the district as fostering the use of the quantitative and/or qualitative information provided in an IRI? Further, it asked to what degree school leaders identified their leadership structures as fostering the use of the fixed instructional practices through quantitative leveling information (benchmarking) provided by an IRI or dynamic instructional practices through personalization of reading instruction or flexible grouping of students for reading instruction using qualitative information from an IRI?

A total of 15 respondents opened the survey, but only 9 completed useable data for a 41% rate of response. Each of the 20 individual questions on the survey matrix was analyzed for frequency and distribution. The 20 questions were further categorized into four main themes to explore the constructs within the study for literacy, leadership and data use for informed decision making. The four constructs included fixed practices that fall under data monitoring systems or benchmarking with IRI reading measure data and dynamic practices that fall under personalized instruction or re-grouping practices in reading instruction using IRI reading data. These constructs were used to garner a picture of how school leaders in this district rated their practices within these arenas. The ratings for use of practices fall within a response scale for always, sometimes, rarely, and never were converted to equate with a numbered scale from 4 to 1 for representation of a numerical average (see Table 10). Additionally, since the majority of the questions were asked in a positive frame with a 4 or always being the highest ranking, but two questions
were asked where the ideal ranking would be on the lower end of the 1 to 4 scale, those
questions were converted to give the same numerical weighting to the averaging process.
The option to select an unsure rating was provided. This would have been coded as a
missing data field; however, no survey respondents selected that rating in this
administration of the survey.
Table 10

*Averages on School Leader Use of Reading Measures Survey by Question & Theme*

<table>
<thead>
<tr>
<th>Item No.</th>
<th>Question</th>
<th>Theme</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I have a regularly scheduled process for meeting with teacher teams or individuals to review reading data results from an informal reading inventory (IRI).</td>
<td>Data Monitoring</td>
<td>3.3</td>
</tr>
<tr>
<td>7</td>
<td>When the IRI is administered, I review the reading level data.</td>
<td>Data Monitoring</td>
<td>3.1</td>
</tr>
<tr>
<td>13</td>
<td>I conduct regular data monitoring meetings across the school year with teachers.</td>
<td>Data Monitoring</td>
<td>4.0</td>
</tr>
<tr>
<td>15</td>
<td>I have a system for collecting IRI reading data across the school year.</td>
<td>Data Monitoring</td>
<td>3.5</td>
</tr>
<tr>
<td>19</td>
<td>I discuss the results of the IRI reading level data with teachers/teams in a data monitoring system after each administration of the IRI.</td>
<td>Data Monitoring</td>
<td>3.5</td>
</tr>
</tbody>
</table>

**OVERALL MEAN FOR FIXED DATA MONITORING PRACTICES** 3.5

| 3        | My teachers administer and record IRI results at the required intervals across a school year. | Fixed: Benchmarking           | 3.9  |
| 9        | Teachers use the reading text levels obtained on the IRI to monitor student benchmarks across the school year. | Fixed: Benchmarking           | 3.9  |
| 11       | I ask teachers about the amount of movement between reading levels for students across a school year based on text reading levels identified on an IRI. | Fixed: Benchmarking           | 3.2  |
| 14       | I ask teachers about text reading levels identified on an IRI.             | Fixed: Benchmarking           | 3.7  |
| 20       | I ask teachers about grouping students by text level identified on an IRI.   | Fixed: Benchmarking           | 3.2  |

**OVERALL MEAN FOR USE OF DATA FOR FIXED BENCHMARKING PRACTICES** 3.6
<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Dynamic:</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>I ask teachers questions that emphasize student strengths and needs identified on an IRI on the use of visual information and what actions they take using this information for reading group instruction or regrouping.</td>
<td>Personalization</td>
<td>3.2</td>
</tr>
<tr>
<td>4</td>
<td>I ask teachers questions that emphasize student strengths and needs identified on an IRI on the use of meaning and what actions they take using this information for reading group instruction.</td>
<td>Personalization</td>
<td>3.4</td>
</tr>
<tr>
<td>6</td>
<td>I ask teachers questions that emphasize student strengths and needs identified on an IRI on the use of structure and language information and what actions they take using this information for reading group instruction or regrouping.</td>
<td>Personalization</td>
<td>3.2</td>
</tr>
<tr>
<td>8</td>
<td>I ask teachers questions that emphasize student strengths and needs identified on an IRI on comprehension of what is read and what actions they take using this information for reading group instruction or regrouping.</td>
<td>Personalization</td>
<td>3.1</td>
</tr>
<tr>
<td>10</td>
<td>I ask teachers questions that emphasize student strengths and needs identified on an IRI regarding phrasing and fluency and what actions they take using this information for reading group instruction or regrouping.</td>
<td>Personalization</td>
<td>2.8</td>
</tr>
<tr>
<td>17</td>
<td>I ask teachers questions that emphasize strengths and needs identified on an IRI on concepts of print and what actions they take using this information for reading group instruction or regrouping.</td>
<td>Personalization</td>
<td>3.0</td>
</tr>
</tbody>
</table>

**OVERALL MEAN FOR DYNAMIC PERSONALIZATION PRACTICES**  
3.1

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Dynamic:</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Teachers move and change students regularly based on current reading data.</td>
<td>Regrouping</td>
<td>3.8</td>
</tr>
<tr>
<td>12</td>
<td>Teachers use reading data to consistently make adjustments to students in reading groups.</td>
<td>Regrouping</td>
<td>3.5</td>
</tr>
<tr>
<td>16</td>
<td>Teachers keep students in the same reading groups throughout the school year.</td>
<td>Regrouping</td>
<td>2.2</td>
</tr>
<tr>
<td>18</td>
<td>Teachers do not move students in or out of groups.</td>
<td>Regrouping</td>
<td>2.0</td>
</tr>
</tbody>
</table>

**OVERALL MEAN FOR DYNAMIC REGROUPING PRACTICES**  
2.9

*Note.* A score of 4.0 is the highest rating possible; 1.0 is the lowest.
Table 13 on averages for the questions on the survey also shows the four categorizations. When these are combined within the four constructs and rank ordered, the results indicate an average rating for fixed practices in benchmarking using IRI reading data as 3.6, an average of 3.5 for data monitoring systems, a 3.1 for dynamic personalized instructional practices, and a 2.9 for dynamic practices under regrouping.

The highest ratings by question overall were within the fixed construct. The school leader actions identified with the highest ratings within the fixed construct included conducting regular data meetings and discussing IRI results in data meetings or teams. Further within the fixed construct arena of highest ratings, school leaders indicated they believe teachers do administer the IRI consistently, as well as use the IRI results to numerically monitor benchmarks.

The lowest ratings by question overall were within the dynamic construct. The school leader actions identified within the dynamic construct included a personalized approach to the teaching and learning with practices for school leaders questioning and monitoring whether teachers use more qualitative data found within the IRI results. The lowest rated items within the dynamic construct were about leaders asking questions in data meetings centered around phrasing and fluency and concepts of print awareness information that can be found within the qualitative information on the IRI. Additionally, the lowest rated item on the survey was the concept of regrouping practices, where teachers move students flexibly and frequently to consistently align with the qualitative data provided on an IRI.

Analyzing the survey results holistically indicates an overall self-rating in the range of fairly consistent use of data monitoring systems and benchmarking reading data
with a more conservative rating in the dynamic leadership practices for personalized and regrouping instructional methods. While an opportunity was provided for respondents to provide additional information in an open-answer response format, the majority did not utilize that response option. Three respondents provided input. These were analyzed for themes around the topic studied. Overall, school leaders seem to indicate a pull towards fixed practices with a general use of IRIs for determining levels and grade level status, as a mandated part of reading practice, but that other topics were focused on at different times. Participants indicated that more training is needed for teachers on the benefits and information IRIs can provide. One respondent reported some of the ratings provided items rated as always or sometimes were only because it is mandated by the school division making it easier to do the right thing. Another respondent indicated teachers just see the IRI data as a means to determine reading levels and on-grade level determinations.
Summary of Results

Three main research questions with sub-sets of questions were analyzed for the current study in order to examine leadership, literacy assessments and acceleration in reading by the third grade. In regard to the relationship between two types of reading measures predominantly used in the state of Virginia, findings were significant with a moderately high correlation between the quarter 4 Rigby IRI reading score and the spring third grade reading SOL score ($r = .568, n = 2906, p < 0.01$). The quarter 1 IRI Rigby data also held a moderately high correlation to the SOL outcome in reading ($r = .510, n = 2906, p < 0.01$). When looking at both the quarter 1 and quarter 4 reading data, there was a strong correlation between the two quarters ($r = .807, n = 2906, p < 0.01$).

School leaders surveys pointed to an overall higher rating towards fixed practices for use of data monitoring systems and the use of quantitative reading levels on IRIs for benchmarking purposes than it did for dynamic practices of personalized instruction using IRI qualitative information and regrouping methods.

In sum, the results of these analyses indicated that:

- there is a significant relationship between the Rigby IRI reading measure and the reading SOL outcomes (Research Question 1);
- both quarter 1 and quarter 4 Rigby IRI scores were equally significant in predicting outcomes on the SOL reading assessment with no statistical difference between the two (Research Question 1a);
- where students performed on the Rigby IRI for reading in Quarter 1 correlated strongly to where students performed in Quarter 4 (Research Question 1a);
• Of all students reading at the first quarter fourth grade reading level 26 or higher on the Rigby quarter 4 IRI assessment, 70% passed proficient or advanced on the third grade reading SOL (Research Question 1b);

• of all students reading at a second quarter fourth grade level 27 on the Rigby quarter 4 IRI, 91% passed proficient or higher on the SOL (Research Question 1b);

• of all students scoring at the set Rigby IRI benchmark for end of third grade (level 25) 25% failed the SOL (Research Question 1b);

• there is a strong relationship between where a student ended third grade based on where they started (Research Question 2);

• there was an overall growth pattern for 23% of students between assessment in quarter 1 and quarter 4 on the Rigby IRI with movement going from either below to on or above grade level or from on grade level to above (Research Question 2a);

• seventy percent of all students stayed within the same grade level categorization of either below, on, or above from quarter 1 to quarter 4 (Research Question 2a);

• a small percentage (7.5%) dropped either from on to below or from above to on or below (Research Question 2a);

• there is a statistically significant difference between the number of levels moved based on the categorization of on, below, or above grade level status (Research Question 2b);
• there was no difference in the amount of levels moved between those classified as fail or pass proficient on the SOL with both moving approximately 2.5 levels (Research Question 2b);

• students in the pass advanced categorization made slightly less movement than those in the pass or fail categories with a mean of 2.2 levels of movement (Research Question 2b);

• school leaders rated themselves higher in fixed practices for using reading data under benchmarking practices and the use of data monitoring systems (Research Question 3); and

• school leaders rate dynamic practices for personalized instruction and regrouping as less consistent (Research Question 3)

Limitations

There are several limitations to consider for the scope of this study. With regards to sampling methods, a convenience sample was used which limits the generalizability of the results to other populations. Because all of the data obtained, as well as the administrators, were from a public school system in Virginia, the findings are not generalizable to school districts with different demographic of contextual factors, those outside of Virginia or those outside of the public school realm. Demographics of the Star School District include 12,697 students across 10 elementary schools, four middle, four high, and one charter school with an average class size at the third grade level being 23.25 students per teacher. Further, the district has approximately 21.5 % economic need, 2.2% limited English proficient students, a 10% special education population, and a
7.4% gifted population. The student population reported for the most recent school year includes: 61% White students, .25% American Indian, 5.5% Asian, 13% Black, 9.9% Hispanic, .27% Hawaiian Pacific, and 9% with two or more races.

Additionally, while the IRIreading assessment tool is part of the theoretical construct of the study as one form of a reading assessment, the reading inventory data set available for use in this study pertained specifically to the Rigby published reading inventory. Therefore, while the patterns and trends studied and reported herein are valuable for the field practitioner’s consideration in leadership and literacy assessment theoretical realms, they are limited to the Rigby assessment tool and not applicable to all versions of informal reading inventories.

Lastly, while the convenience sample used provided a robust quantitative data set for breadth of the correlational nature of the two reading measures studied, it provided a very limited data set for the survey portion of the study. The survey was designed to garner a better understanding of school leaders’ level of knowledge and use of reading measures data available from the tools studied. The pool available for this purpose, however, was limited by the size of the school district studied and limited to just ten elementary schools. Therefore, while the return rate of 41% would be considered fairly typical, it limits the depth of understanding to a rather small pool of cases. Additionally, the school leaders in this particular sample population have undergone an overhaul of the school district’s literacy model and participated in a five-year long range plan that included a heightened understanding of overall literacy constructs and the use of assessment tools. Consequently, consideration must be taken regarding their self-reported ratings on the survey measures as indicative of that particular population’s
experiences and environment, and therefore, may not be reflective of other school leader populations in different localities. With the timing of the principal survey being more recent than the archived data set of reading assessment scores, it is important to note the principals have received various training experiences on the literacy model overhaul during the intervening timeframe. Likewise, among the principal participants, it is unknown as to what the difference may be between those who chose to participate and those who did not choose to participate. Thus, the scope of understanding for all leaders even within the school district studied is limited.
Chapter 5: Conclusions

The purpose of this exploratory study was to investigate school leaders’ practices in a data-informed system around two widely used forms of reading measures. State mandated grade-level tests are required, but schools also commonly use some form of an informal reading inventory to monitor the early reading acquisition process. It is important and timely to consider the correlational nature of these two main measures of reading as well as obtain a sense of school leaders’ level of leadership practices in an ever more urgent battle against performance gaps in reading among children in the U.S., the state of Virginia, and the region in which the study was conducted. An era of accountability and assessment solidified the widespread use of multiple measures of student outcomes, but school leader data informed decision-making is dependent upon the reliability and validity of the measures used as well as the degree of understanding and application of the information garnered from the assessment tools. It is possible to be data rich, and yet information poor.

As reading growth remains stagnant across the U.S., and lags behind in Virginia (U.S. Department of Education, 2016; Virginia Department of Education, 2014), this study seeks to answer some long-held questions among educators, and the results are more timely than ever. District and school leaders continue to grapple with critical decisions in the types of assessment systems they use, and the methods of implementation in their work with schools. They seek to make advances in reducing reading gaps and
level the playing field for all learners in the early years of literacy acquisition. This is all predicated upon what assessments they use and how they use them.

The first section discusses findings for the correlational nature between the SOL and one form of a published IRI, the Rigby. It also looks into the levels of reading students fell into based on the classifications of pass, pass advanced, and fail on the Virginia third grade reading SOL, along with the proportions of change based on reading group status (on, above, below) between quarter 1 and quarter 4 and rounds out with the amount of movement across levels by classification on the SOL. This section will feature the critical constructs introduced in Chapter 2 on acceleration and data-informed decision making, which will be followed by the limitations of the study, implications of the findings, future research, and conclusions.

**Discussion of Findings**

**Reading Measures**

An analysis of the descriptive statistics revealed the mean Rigby IRI score for the Star School District was 24.38 ($SD=3.16$) for quarter 1. This score was interesting because according to the widely accepted reading level ranges assigned to grade levels by field practitioners, such as Fountas and Pinnell (2012), and agreed upon by the school district, the expected IRI equivalent for the Rigby Quarter 1 data would be a level 23. According to the quarterly goals for grade three in the district, a level 23 is the end of quarter 1 with a level 24 for the end of quarter 2, and level 25 for both quarters 3 and 4. Thus, the five-year cohort mean score sampled for this study performed at a higher than expected level with regard to the benchmarks set by experts. Likewise, the quarter 4 mean for the end of third grade was a level 26.78 ($SD=2.61$), which is the equivalent of
quarter 2 in the fourth grade. Thus, students in the Star School District, overall, were above grade level at entry and end of year based on the Rigby IRI data for third grade.

Furthermore, the mean score for the third grade reading SOL assessment was 467.63 ($SD=72.23$). This places the cohorts studied well above the pass cut score of a 400, but under the pass advanced score of 500. In considering the categorizations for students on, above or below grade level, approximately 15% ($N=457$) of students were considered below grade level at the quarter 1 IRI assessment period. Of those performing below grade-level expectations, 48% were able to change grade-level categorization by quarter 4 reassessment on the IRI. Evaluating this information in light of the increasing urgency and stagnating proficiency levels on the NAEP, as reported earlier in this study, the Star School District appears to be performing better than the NAEP data suggest for the nation (37%) and the state (47%; U.S. Department of Education, 2016; Virginia Department of Education, 2014).

This study took a robust sample size of just under 3,000 students and investigated the relationship between the state’s third grade reading SOL and an IRI, the Rigby. The outcomes indicate a significant relationship exists between the level of reading obtained by quarter 4 as assessed on the Rigby and the SOL outcome. Thus, where students end the year in terms of their IRI reading data has a moderately high correlation to SOL outcomes. This was also true for the quarter 1 data, which indicated a moderately high correlation as well. However, there was a strong correlation between quarter 1 and quarter 4 Rigby scores. In general, these findings support the use of the IRI Rigby assessment as an appropriate tool for measuring reading progress and predicting student outcomes on the SOL assessment within a holistic data-monitoring system.
School leaders and division leaders benefit from having this confirmation on a form of an IRI that has not previously been rigorously examined. Past studies included a look at the Developmental Reading Assessment, the Qualitative Reading Inventory, and the Scholastic inventory (Askew, 2011; Cramer; 2010, Hickey, 2012; Johnson, 2014; McKeone, 2005; Morris, 2004). While just two of the studies were done in Virginia (Johnson, 2014; Morris, 2004), all of the studies found modest to strong correlations between the use of an IRI and various state reading assessments (Tennessee, Ohio, and Pennsylvania) as predictors for outcomes. The addition of this study, with a moderately strong finding for the Rigby IRI and the Virginia state third grade reading assessment, augments school leader decision making in the continued use of some form of an IRI as a reading measure that does offer value in predicting student achievement at the state level assessment measure. School leaders making critical decisions to keep or do away with reading measures, specifically in Virginia, can now use this information to buoy decision-making for keeping an IRI in an assessment system for reading measures. Keeping in mind that an IRI is just one form of measurement in reading, leaders can consider other assessment measures within an entire data monitoring system. While this study focuses on the Rigby IRI, the type of analysis done here could be replicated with other forms of IRIs and the Virginia state reading measure. This alone provides important data-based evidence for leaders and is helping to create a preponderance of support in the general construct for their use. Moreover, the additional questions posed in this study, delve deeper providing further insight and clarity for school leader consideration, as will be discussed in subsequent sections. Moreover, considerations must be given to the limitations of a moderate correlational finding for the Rigby to the SOL because there is
still some variability in the prediction of likely outcomes based solely on the Rigby IRI toward the SOL spring assessment. Thus, a well-rounded toolkit and use of the qualitative data provided within the IRI, aside from just the numerical benchmark provided on IRIs, is important to note. School leaders benefit not just from understanding more about the predictive value of the Rigby to the SOL, but also can benefit from realizing the untapped potential with using the dynamic approaches to instruction featured within this study that include personalized instruction and frequent regrouping based on the additional qualitative information found in the IRI.

Research question 1b further investigated the range of Rigby IRI scores of students at each categorization for pass, pass advanced, and fail. Almost half (49%) of students in the study fell into the pass proficient range on the reading SOL. Considering that an equivalent Rigby IRI score of 25 for the end of the third grade year places a student as on grade level for reading progress, it is important to look further and see that a quarter (25%) of the students performing at grade level on the Rigby were categorized as a fail on the state reading SOL test. Looking further into the trends between the IRI scores and the categorizations on the SOL, of all students who scored a Rigby IRI level 26, which is entering grade four, 60% fell into the pass proficient range on the SOL, 20% in the pass advanced range and 20% into the fail range. At the Rigby level of 27, which equates to quarter 2 of Grade 4, the fail rate drops to approximately 10% on the SOL. Thus, 90% of all students performing at the mid-year fourth grade level passed the SOL. Of all those who performed at the end of grade four and beginning grade five IRI range (levels 28 & 29), 96% passed proficient or higher. Once students reached a Rigby level 30 (end of Grade 5), virtually all of the students (99%) scored proficient or higher.
Overall, these data show that higher reading levels on the Rigby IRI seem necessary for passing the state reading SOL. As the reading levels reached fourth grade or higher, the percentage of those students passing the SOL went up, increasingly, with 99% of those students topping out at a fifth grade reading level on the Rigby passing the SOL.

**Acceleration**

There is a strong relationship between where a student ends based on where he or she started the year in reading level. Using the Rigby reading-level equivalencies, student outcomes were investigated (question 2a) for level of movement or change in the categorizations for being considered on grade level, below grade level, or above grade level, and it revealed a majority of those studied (70%) stayed within the same categorization from quarter 1 to quarter 4. Another way of viewing this information is from the perspective that the majority of students stayed within whatever group they came in on—either on, below or above grade level. Drilling down further into each categorized group (on, above, below) helps uncover additional patterns. From all those students who were categorized as below in quarter 1, 52% stayed the same, 29% made a change to an on grade level status, and 19% made a change to an above grade level status. Considering only those students who were categorized at the on-grade level range in quarter 1, 9% moved to a below grade-level status, 33% stayed on grade level, and 58% moved to above grade level. For only those students categorized in the above grade-level category in quarter 1, 91% stayed above in quarter four with an 8% drop to on grade level, and 1% moved to below grade level.

From these results, it is appropriate to surmise that students who started the third grade above level tended to stay above grade level. Of those who started out below grade
level, about half were able to make a change upward, suggesting there is evidence of student acceleration to the expected grade level performance band and closing reading gaps between quarter 1 and quarter 4. These data suggest the concept for acceleration and closing reading gaps is happening to some extent in the school district utilized for this study. This is essential evidence for the work being done in schools towards ensuring all students become literate by the third grade. A small number of students slid backwards from being above or on-grade level to below grade level. However, it is important to also recognize that over half of the cases making upward movement were found within the on-grade level grouping who were able to make movement reaching the above grade level band by quarter four suggesting instruction is most successful for the average student, and less impactful for those below or above.

Furthermore, research question 2b looked at to what degree the number of reading levels students moved, as assessed by the Rigby IRI, differed among students classified as pass, pass advanced, or fail on the end-of-year VA third grade SOL Reading test across the school year for this sample of third grade students. There was a significant difference between being categorized as pass proficient, pass advanced, or fail, and the number of reading levels students moved from quarter 1 to quarter 4 on the Rigby IRI. The data revealed that students classified as pass proficient or fail on the SOL moved at approximately the same rate or number of IRI reading levels between quarter 1 and quarter 4 while the students classified as advanced on the SOL moved at a lower rate. In addition, the mean number of levels moved for both the pass proficient and fail groups was approximately 2.5 levels while the advanced was lower with a mean of 2.2 levels moved. Given that the expected amount of movement is three Rigby reading levels (from
23 to 25) across the year, this information is surprising because no group mean demonstrated a full three level movement pattern. Furthermore, in considering acceleration, it might be considered noteworthy that while the fail group kept a similar pace of movement to those at the on-grade level status, the number of levels in movement do not necessarily equate to accelerated growth. To this point, Allington & McGill-Frazen (2013) emphasizes that while below grade level students make the same amount of progress throughout the school year, a claim supported in the findings of this study, it is actually summer reading loss that accounts for roughly 80% of the reading achievement gap that exits between children of varying economic status (Alexander, Entwisle, & Olson, 2007). Furthermore, this is exacerbated by the on-going buildup of summer loss over multiple years, such that by the time students near graduation, there is a 4 year accumulated reading gap (Allington & McGill-Frazen, 2013). This supplicates another important question to ponder. Is a 2.5 level of movement across the school year acceptable for either group, on-grade level or below grade level, knowing the yearly growth expectation is a 3 level change, and is it especially enough for those who are impacted by compounding summer reading loss?

Thus, while initial data revealed a positive outcome for some changes in grade level status for percentage of students moving up and out of below grade level performance, a deeper look into the number of levels actually moved exposed further information for consideration as school leaders. At the rate of levels gained found in this study, students are not catching up to on-grade level peers, and thus, acceleration and closing reading gaps is not being achieved. This is a critical outcome from the study and one that should be deliberated thoughtfully by school leadership. If summer reading loss
for low performing students is one of the most important mitigating factors, as Allington & McGill-Frazen (2013) and other researchers suggest, more needs to be done to address summer reading methods for students to maintain the gains from across the school year, and provide opportunities for additional reading maintenance or growth as their on or above grade level peers experience.

**Data Informed Decision Making in Leadership**

In recent years, the field of educational leadership has burgeoned to incorporate leadership practices that inspire and transform schools beyond the years of just managing the schoolhouse (Darling-Hammond et al., 2007). The disciplines that school leaders institute and utilize can have direct impact on student outcomes second only to that of the front-line instruction provided by classroom teachers (Bernhardt, 2003; Hallinger & Heck, 1998; Lashaway, 2002; Marzano, 2002; Marzano et al., 2005; Reeves, 2002; Taylor, 2010). The National Policy Board for Educational Administration (2015) puts an impetus on this portion of the leadership realm through the required standards for use of student performance data by principals, highlighting the potential increase in effectiveness in schools and student outcomes. The types of data principals have and how they use it also plays a critical role in the data informed aspect of school leadership (Marzano et al., 2005), and this study aims to add to that understanding around the use of informal reading inventories and in relation to the state mandated reading test. The findings within this study amplify the stance found in literature on the critical nature of not only the kinds of data principals use, but also the way in which they use it as a tool for decision making (Dorn & Soffos, 2002; Halverson et al., 2007).
The school leader use of reading measures survey to address question 3 was designed as a means for obtaining collective feedback to begin to explore the depth and level of their use around one of the most widely used reading measures in the field in the form of an IRI, and in this study in particular, the Rigby published version. While the applicant pool was limited in size to the school district examined in this study, it provides a small window into the arena of how this sample of leaders rate themselves in their daily practices towards a set of fixed practices around the use of the district selected IRI versus a dynamic set of practices around the qualitative information also available within the Rigby IRI. This is the lens in which the study is situated with practices that foster basic data-monitoring systems, the use of IRI data in a static benchmarking form or dynamic practices that lean toward personalized instruction and frequent regrouping for instruction.

The results of the survey, while limited in number, show school leaders in this district reported a fairly consistent use of data monitoring systems and benchmarking reading data with more conservative ratings in dynamic leadership practices for personalized and regrouping instructional methods. The highest ratings by question overall were within the fixed construct for conducting regular data meetings, discussing IRI results in data systems or teams, teachers administering the IRI consistently, and teachers using the IRI to monitor benchmarks. These findings on the survey results suggest a high level of implementation for school leaders in the Star School District and their use of data driven systems.

The lowest ratings by question overall were within the dynamic construct for personalized instructional practices around phrasing and fluency, personalized
instructional practices around concepts of print, teachers keeping students in the same
groups for the year, and re-grouping practices. These suggest principals were not as
confident in structures that emphasize the deeper level and use of IRI qualitative data for
driving day to day instruction. Keeping in mind that part of what principals do to set high
expectations is through leading by example and asking questions to guide teacher growth
and practice (Coburn & Turner, 2011; Deike, 2009; Farley-Ripple & Buttram, 2014;
types of questions and structures for dialogue around literacy instruction that principals
use can foster either a dynamic or fixed instructional practice mindset with teachers.

In narrative open-ended responses, school leaders indicated a pull towards fixed
practices with a general use of IRIs for determining levels and grade level status as a
mandated part of reading practice, and more training is needed for teachers on the
benefits and information IRIs can provide. This is not surprising. As discussed in
Chapter 2, while a great deal of work has been done to move the field toward leadership
practices that transform schools (Darling-Hammond et al., 2007) and work to close
reading gaps, the high stakes accountability era solidified a great deal of emphasis on
standards based instruction and the pendulum swung toward a strong emphasis on
measuring student outcomes. This has worked against the level of dynamic practices
needed to truly get at the level of personalized instructional decision making that meets
children where they are and creates the environment for accelerating and frequent
regrouping of students for constant movement and growth. To restate Routman’s (2014,
p. 38) sentiments, although standards and curriculum can guide our efforts, what and how
we teach, must include interconnected responsive practices that lead to high student engagement, achievement, and independence as learners.

**Implications**

Exploration of the topics studied in this research around leadership practices, literacy assessments, and acceleration in reading by the third grade revealed interesting implications that school leaders could use to inform their leadership practice. Reducing reading proficiency gaps for students lagging behind their grade-level peers is critical for redirecting the trajectory of success for all students in literacy. While Table 11 at the end of the chapter features a synopsis of the main results, findings, and implications from the study, and an expanded dialogue is provided below on the reading measures, acceleration, and data informed decision making.

**Reading Measures**

When considering the two main widely used reading measures, it is important to know if these predominant measures are valuable toward the goals for ensuring all students are literate by third grade. Since the state reading measure for the SOL are a mandatory part of our accountability systems in education and because informal reading inventories are widely used and accepted as measures of student reading growth, it is beneficial for leaders to know how these main reading assessments can inform one another and the degree of relationship in their use for predicting successful reading mastery according to the state assessment by the end of third grade. The outcome of this study did find the Rigby IRI offers a predictive value for outcomes on the reading SOL at third grade. It aids the field in knowing that this form of an IRI, along with those studied by previous researchers, is building a base of evidence toward the value of an IRI against
state reading measure outcomes (Askew, 2011; Cramer, 2010; Hickey, 2012; Johnson 2014; McKeone, 2005; Morris, 2004). Schools in Virginia can implement the process detailed in this study to determine if other forms of IRIs used in school districts correlate to the third grade reading SOL. Other state measures could also be used against the more widely used forms of IRIs to bolster the findings as well. Schools using a Rigby IRI can be confident that the time, energy, resources, and efforts to assess students using the Rigby is valuable time and beneficial to measuring reading growth.

**Closing Reading Gaps**

In terms of reducing the reading failure rates and increasing students reaching the average bands for reading performance by the third grade, this study found that the Star School District data showed movement of 23% of the students between quarter 1 and quarter 4. This was surprising and good news in terms of closing gaps in performance between students below grade level and those on or above grade level. As this concept was central to the theoretical constructs to this study, it was surprising to see the degree of students who were able to move either from below grade level to on-grade level or from on to above. The data indicate instruction impacted students who came in on-grade level the most with the amount of movement seen in that group as 58% moving from on-grade level to above grade level by quarter four. This suggests instruction helps the on-grade level group the most. While this is good news for moving the on-grade level students to greater success in the above grade level band, it does not indicate a mastery of adaptive instruction for those students coming in below grade level.

Thus, it is helpful to muse about what these implications may look like moving from theory to practice. When considering the nature of the reading gap among students
below grade level and thinking beyond just the quantitative numerical level that the
typical use of an IRI score provides, it is helpful to contemplate exactly what a more
dynamic approach looks like from the qualitative information available on IRIs. This can
include the analysis of what a student does when reading and teacher observations of
reading behaviors. The literacy cueing system provides answers to the strengths and
needs of each child’s literacy processing system, revealing tendencies students have
towards use of the types of information available to them while reading. Analyzing the
cueing systems within a meaning system, a structural system of how language works, and
the visual information within printed text provides the teacher with valuable evidence
about what a student uses well, and what a student needs next to improve the process for
the very next reading opportunity. This provides a window into what students use and
neglect and offer the teacher the chance to zero in with laser precision towards
scaffolding the student learning needs to move them forward in the literacy process. This
type of focused instruction leads to an ever-improving literacy cycle that allows students
to continuously grow, which is the goal for all students, regardless of grouping status, but
especially for students lagging behind their age appropriate peers.

This dynamic personalized approach to instruction and the use of frequent
regrouping practices, as students continuously grow and change, provides the basis for an
acceleration model. The notion of an acceleration model, as discussed in this study,
challenges the status quo practice of a standards based curricular framework for the
youngest learners from kindergarten to third grade. I posit it is more important to
consider the dynamic instructional practices needed to constantly shift and move students
in a literacy framework to persist in closing wide differences among the youngest
learners in reading performance than it is to follow a lock-step curricular map for discrete skills and content in the early grades. The data found in this study amplifies what experts in the field have reiterated for years now. If reading gaps are not closed and students are not on grade level by third grade, students are very likely to continue to struggle, remain behind in literacy development, and have long-term impact across their lives.

**Data Informed Decision Making**

The school leader use of reading measures survey indicated leaders in this study had more confidence in the implementation of overall data monitoring systems, administering assessments at the school building level, and using literacy assessments for benchmarking and leveling students. The implications of this finding includes a positive level of implementation for administrators within the district studied. School leaders should continue to employ these systems, and work to expand their breath and depth of knowledge in further application of the qualitative information found within the IRI assessments utilized. Less confidence was shown in the realm of dynamic practices that feature the qualitative data available in the IRI measures offered. This includes the level of questioning and dialogue around strengths and needs of students for growth in the total literacy process for areas such as phrasing and fluency, reading rate, comprehension, and the use of the cueing systems for meaning, structure, and visual information. IRIs provide this level and depth of information, but school leaders rated themselves more conservatively in employing structures and questioning in their data monitoring systems within these realms. This implies an area of growth for a school leaders exists, especially if the goal of closing reading gaps by third grade is to be actualized.
One recommendation for consideration in the area of IRI reading measures is for school leaders to determine if the particular IRI tool being used meets the needs of the goals for assessment. In the case of the Rigby used in this study, it has a ceiling effect for ending at a level 30 or the fifth grade level. Since a large number of students in this study were well beyond the grade level expectations at the start of the third grade and continued to grow to the end of fifth grade benchmark by the end of the third grade, it advisable to use an IRI that goes beyond the fifth grade and allows for a more accurate picture of growth for students working above grade level. Without a tool that measures growth for all students, including those who are performing well above the grade level expectations, there is no real way to monitor a year of growth for all students, even if they come in above grade level already. This is a key consideration since it is expected that students working above grade level would still benefit from dynamic and adaptive teaching as well as those benefiting within the average bands of performance, as the data from this study indicated.

Further, since both quarter 1 and quarter 4 IRI reading data indicate equal moderately high correlations, it is advisable for school leaders to consider actions that foster a sense of urgency within the first 45 days of the third grade, prior to the quarter 1 assessment. Any growth and advancement that could be made within the first quarter to increase student reading levels on the Rigby IRI for quarter 1 could bolster the number of students meeting with more success by quarter 4 and on the SOL. It is recommended for school leaders to create a sense of focused instructional goals in the first quarter, along with academic strategies for quarter 4 in the second grade, strategies for preventing summer slide (Allington & McGill-Franzen, 2013) prior to the third grade, and highly
focused immediate in-depth re-introduction to the reading process in quarter 1 of third grade. Using the valuable information provided on the IRI as predominant reading behaviors coming up from the second grade, teachers could zero in on student strengths and needs for a dynamic and personalized approach to reading instruction immediately upon the start of third grade. These recommendations need to be led by the school leader who can create the catalyst for implementation from the data, and who can apply a more persistent level of academic press for accelerating reading. Hence, the implications call for the exploration of the optimal level of training that is needed to prepare school leaders to lead this charge. It is not enough for a reading specialist to lead a data-monitoring system for literacy benchmarking and quarterly assessments. The school leader has an equally important ability to make an impact on the literacy outcomes. Exploration of just what kind and to what degree of professional development is required for school leaders is a critical missing component in the data monitoring system. The field of school leadership and leadership preparation programs need to know more about what it takes to equip leaders on how to ask the right kinds of questions towards a dynamic and personalized set of instructional practices using all of the IRI data available to truly accelerate learning and close reading gaps among students in the early grades.

Further, the outcomes in this study, which indicate where students start out in quarter 1 and where they end in quarter 4 are strongly correlated, expands the constructs that acceleration of any kind that can be made prior to the start of third grade and within quarter 1 of the third grade and could improve outcomes for students in reading. On the other side of the issue of positive changes in reading group status, a small percentage of students did move backwards from being either on grade level to below or above to on-
grade level. Leaders would benefit from learning more about these trends and take steps to reduce this pattern.

Indeed, the findings around the self-ratings for school leader use of a more dynamic approach to leadership and instructional decision making through personalized instruction from using the qualitative information available within the IRI process as well as frequent regrouping practices indicated there is room for growth. This is an area where school leaders could benefit from additional professional development in fostering this type of in-depth dialogue and questioning to steer teachers towards the dynamic approach. As it stands now, school leaders rated themselves higher in the application and use of a fixed approach leftover from the era of a standards-based approach to instruction whereby IRI data is collected for the main purpose of checking benchmark levels, entering scores into data monitoring systems, and moving along a pre-planned scope and sequence for grade level learning. However, a complex theory of literacy learning demonstrates that student reading growth can be accelerated through the use of more detailed information found in IRIs, if applied by classroom teachers regularly and used to group and regroup students to continually meet their reading needs (Clay, 1991; Duncan, 2016; Phillips et al., 1997). The research on the impact of school leaders, secondary to classroom instruction, is ever-increasing and indicative of great potential; thus, it is incumbent on school leaders to begin to model, foster, and monitor this dynamic and in-depth use of qualitative information available within the reading measure tools they already use and implement.

One area of particular interest for district and state level administration is the notion of the higher reading levels exhibited by the cases examined within this study for
reaching a pass proficient or above status on the third grade reading standard of learning test. It is interesting to note that the higher levels obtained on the Rigby IRI, those on or above a fourth grade level, tended to yield the increased likelihood of passing the third grade state reading assessment. If this trend were to hold true for other IRI measures studied and other localities around the state, it would be important for state level administrators to review and recalibrate the state assessment tools and for school leaders to consider the alignment issues inherent in this discrepancy with informal reading measure benchmarks widely used by schools. Moreover, if state leaders had a centralized database for collecting and equating IRI reading data, as they indicated in their study on third grade progress (JLARC, 2011), they could add clarity, increase outcomes, and raise expectations statewide, potentially. Especially, if further studies revealed a menu of appropriately valid and reliable IRI forms and their use for predictability in state reading success. On the opposite side of this spectrum, likewise, schools may need to evaluate their benchmarking goals. If a higher level of reading performance is demanded by the state expectations for reading levels, then schools may also need to consider the benchmarks are off point and should raise the bar to meet the increased demands at the state level.

Accordingly, school systems really would benefit from evaluating what it would take to bring school leaders up to speed in leading the way for this level of dynamic and personalized approach to the use of informal reading inventories so the benefits found in the predictive value from this study can be expanded toward closing reading gaps by the critical stage of third grade. Typically, school organizations presume the reading specialist as the most likely person to manage literacy measures and oversee the overall
implementation of reading programming in schools. However, as has been shown in the research reviewed in this study, the school leader can and does have a large impact on school outcomes. Reading specialists may have limited supervisory abilities and skillsets, and the school leader role is better suited for setting academic press and accountability, more so than the specialist. Given the critical nature of learning to read by the third grade, it is more advantageous for school leadership preparation programs and school districts to invest in an exploration of what it takes to educate elementary school principals for the optimal use of qualitative reading data to refine teaching and learning and adjust pacing for a truly accelerated model of learning. Without exploring the training needed to help elementary principals better guide the literacy model components and instruction, there is risk of little change in below grade level student status, and an unchanging gap in literacy development for some students.
### Table 11

**Results, Findings, and Recommendations**

<table>
<thead>
<tr>
<th></th>
<th>Results</th>
<th>Findings</th>
<th>Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>There is a significant relationship between the Rigby IRI reading measure and the reading SOL outcomes (Research Question 1);</td>
<td>The Rigby IRI offers predictive value for student outcomes on third grade reading SOL.</td>
<td>Schools can use this process to determine if the reading measures they use, such as an IRI like the Rigby, also correlate and offer predictive value.</td>
</tr>
<tr>
<td>2</td>
<td>Both quarter 1 and quarter 4 Rigby IRI scores were equally significant in predicting outcomes on the SOL reading assessment with no statistical difference between the two (Research Question 1a);</td>
<td>Quarter 1 Rigby data offers the same predictive value as quarter 4 for the outcomes on the SOL.</td>
<td>Use quarter 1 data to tackle differentiated instructional goals for third grade students as marked changes in reading level within quarter 1 could increase student outcomes by quarter 4 and on the SOL.</td>
</tr>
<tr>
<td>3</td>
<td>Where students performed on the Rigby IRI for reading in Quarter 1 correlated strongly to where students performed in Quarter 4 (Research Question 1a);</td>
<td>The level at which students enter third grade strongly predicts where they will be at the end of the year.</td>
<td>Efforts to increase student entry level into third grade could increase likelihood for student outcomes by the end of the third grade.</td>
</tr>
<tr>
<td>4</td>
<td>Twenty-five percent of all students scoring at the set Rigby IRI benchmark for end of third grade (level 25) failed the SOL (Research Question 1b);</td>
<td>Reaching the current end of year third grade reading benchmark does not guarantee passing the SOL.</td>
<td>The current benchmark for end of third grade is possibly too low for what is needed to increase the likelihood of successful outcomes on the SOL.</td>
</tr>
<tr>
<td>5</td>
<td>Seventy percent of all students reading at quarter 1 fourth grade reading level 26 or higher on the Rigby quarter 4 IRI assessment passed</td>
<td>The likelihood of passing proficient increases when students reach fourth grade quarter 1 reading level.</td>
<td>Leaders might reconsider if the fourth grade quarter 1 level is a better goal for end of third grade reading benchmarks.</td>
</tr>
</tbody>
</table>
proficient or advanced on the SOL (Research Question 1b);

6. Ninety-one percent of all students reading a quarter 2 fourth grade level 27 or higher level on the Rigby IRI quarter 4 passed proficient or advanced on the SOL (Research Question 1b);

   The likelihood of passing proficient on SOL increases greatly when students reach fourth grade quarter 2 reading level.

   Policy makers at the state level might benefit from evaluating other IRI measures to see if this trend occurs across more measures. If so, the state test may need to be recalibrated to align with end of third grade expected reading levels.

7. There was an overall growth pattern for 23% of students between quarter 1 and quarter 4 on the Rigby IRI with movement going from either below to on or above or from on to above status (Research Question 2a)

   There is some evidence of upward movement for students to change their grade level status to on or above between quarter 1 and quarter 4.

   Evaluate who and why students are making forward movement and seek to increase this trend.

8. A small percentage (7.5%) dropped either from on to below or from above to on or below (Research Question 2a);

   There is evidence of a small percentage of students who fall backwards in their grade level status.

   Evaluate who and why students may be dropping back in grade level status.

9. There is a statistically significant difference between the number of levels moved based on the categorization of on, below, or above grade level status (Research Question 2b);

   Above grade level students moved less levels than those on or below grade level (mean of 2.2 levels across the year).

   Evaluate if this is due to the ceiling effect from the Rigby IRI tool and consider other measures to continue to determine if above grade level students are making more gains than shown on the Rigby tool.

10. There was no difference in the amount of levels moved between those classified as fail or pass proficient on the SOL with both moving

    On grade level and below grade level students made the same amount of changes in number of reading levels (mean

    Consider if the same amount of movement between these two groups is enough and/or appropriate for closing reading gaps
approximately 2.5 levels (Research Question 2b); of 2.5 levels across the year). and an acceleration model.

11. School leaders rated a mean of 3.6 for fixed practices with benchmarking of reading data, and an average of 3.5 for data monitoring systems (Research Question 3);

School leaders rated themselves higher in fixed practices for using reading data under benchmarking practices and the use of data monitoring systems (Research Question 3);

Overall, school leaders have cursory data driven system in place and monitoring the administration of reading assessments, data collection, and meeting standard benchmarks.

12. School leaders rated a mean of 3.1 for dynamic personalized instructional practices, and a 2.9 for dynamic practices under regrouping. (Research Question 3).

School leaders rate dynamic practices for personalized instruction and regrouping as less consistent (Research Question 3).

School leaders may have a lower level of mastery for guiding teachers in working with qualitative IRI information to inform and accelerate learning instruction for students in reading.

---

**Future Research**

To amplify the present findings, more research is needed in leadership, data monitoring systems, literacy assessments, and accelerative methods for closing reading gaps by the third grade. There are several potential worthwhile lines of research that could add to the repertoire explored in this study, such as:

- A replication of the current research design on the correlational nature of other informal reading inventories against the third grade state reading assessment.

- A replication of the current research design around the state to further calibrate to the level of appropriateness and/or difficulty of the state third grade reading assessment to the Rigby and other IRI measures.
• An expansion of the school leader survey tool to larger pools to get a better barometer of the confidence levels and proficiency of school leaders’ use of qualitative data measures on an IRI.

• A more in-depth qualitative interview process to capture more of school leaders’ thinking around literacy assessment measures and leadership practices within a data-informed decision making system.

• Replicating the study for the purpose of expanding the understanding around the level of movement in reading levels across the third grade year to determine if the same amount of movement is happening across the state in other districts.

• A study that expands a look at the same concepts and correlations at the second grade level to better inform the work done prior to third grade and the ways in which efforts for leadership and use of IRI data could be increased prior to the third grade year.

**Conclusion**

In conclusion, increasing reading levels and the number of students reaching reading proficiency by third grade is crucial for our societal goals of producing literate and productive citizens. School leaders can and do have significant impact on the learning environment in schools and there are a limited number of assessment tools for the reading acquisition stages, in addition to the high stakes state reading assessments schools are mandated to use. It is important for school leaders to understand more about the purposes, connections, and capabilities between two of the most widely used forms of reading assessments: an informal reading inventory and the third grade Virginia Standards of Learning reading assessment. Moving away from a standards-based
approach and fixed practices in the use of the valuable information available within informal reading inventories toward a dynamic approach that adapts to student strengths and needs is required for an accelerative approach to student progress in literacy development. School leaders should be trained to have a better skillset, knowledge, and application for using the critical qualitative information found in informal reading inventories within their data driven instructional systems. The third grade reading conundrum need not be an insurmountable mountain as evidence shows combinations of effective leader practices coupled with advanced knowledge and application of informal reading measure data can foster an environment ripe for acceleration, and ultimately, the reduction of the reading gap for students by the third grade. More research must be done to further expound on leadership practices, literacy assessments, and acceleration. Further research, such as this study and those suggested for future investigation, will add to our knowledge base and expand our understanding and practice for implementation in the field.
# Appendix A

## Text Level Correlation Chart

<table>
<thead>
<tr>
<th>Reading Recovery Levels</th>
<th>Rigby Catalog Reading Level</th>
<th>Grade Level Equivalent</th>
<th>Rigby PM PM Plus Levels</th>
<th>Fountas &amp; Pinnell Level</th>
<th>Dominie Levels</th>
<th>Wright Group McGraw-Hill</th>
<th>Success for All</th>
<th>DRA2 Levels</th>
<th>Lexiles</th>
<th>DRP Degrees of Mastery</th>
<th>SAT10 Scaled Score</th>
<th>STAR Reading Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1-3</td>
<td>A</td>
<td>BR</td>
<td>NA</td>
<td>&lt; 450</td>
<td>0.3</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>Emergent</td>
<td>Beginning Kindergarten</td>
<td>Starters 1</td>
<td>A</td>
<td>1</td>
<td>A</td>
<td>BR</td>
<td>NA</td>
<td>NA</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>Early</td>
<td>Middle K</td>
<td>Starters 2</td>
<td>B**</td>
<td>2</td>
<td>B</td>
<td>2</td>
<td>BR</td>
<td>NA</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>End K</td>
<td>3-4 red</td>
<td>C</td>
<td>3</td>
<td>C</td>
<td>3</td>
<td>BR</td>
<td>NA</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td>Beginning Grade 1</td>
<td>5-6 red/yellow</td>
<td>D</td>
<td>4</td>
<td>D</td>
<td>4</td>
<td>BR</td>
<td>NA</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>Middle Grade 1</td>
<td>7-8 yellow</td>
<td>E</td>
<td>5</td>
<td>E</td>
<td>5</td>
<td>BR</td>
<td>NA</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>End Grade 1</td>
<td>9-10 blue</td>
<td>F</td>
<td>6</td>
<td>F</td>
<td>6</td>
<td>BR</td>
<td>NA</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>Early Fluent</td>
<td>Beg. Grade 2</td>
<td>11-12 blue/green</td>
<td>G</td>
<td>7</td>
<td>G</td>
<td>7</td>
<td>BR</td>
<td>NA</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td>Middle Grade 2</td>
<td>13-14 green</td>
<td>H</td>
<td>8</td>
<td>H</td>
<td>8</td>
<td>BR</td>
<td>NA</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td>End Grade 2</td>
<td>15-16 orange</td>
<td>I**</td>
<td>9</td>
<td>I</td>
<td>9</td>
<td>BR</td>
<td>NA</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td>Early Fluent</td>
<td>Beg. Grade 2</td>
<td>J</td>
<td>10</td>
<td>J</td>
<td>10</td>
<td>BR</td>
<td>NA</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>Fluent</td>
<td>Middle Grade 2</td>
<td>17-18 turquoise</td>
<td>K</td>
<td>11</td>
<td>K</td>
<td>11</td>
<td>BR</td>
<td>NA</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td>End Grade 2</td>
<td>19-20 purple</td>
<td>L</td>
<td>12</td>
<td>L</td>
<td>12</td>
<td>BR</td>
<td>NA</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td>Grade 3</td>
<td>21 gold</td>
<td>M</td>
<td>13</td>
<td>M</td>
<td>13</td>
<td>BR</td>
<td>NA</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td>22 gold</td>
<td>N</td>
<td>14</td>
<td>N</td>
<td>14</td>
<td>BR</td>
<td>NA</td>
<td>2.8</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td>Grade 4</td>
<td>23 silver</td>
<td>O**</td>
<td>15</td>
<td>O</td>
<td>15</td>
<td>BR</td>
<td>NA</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td>24 silver</td>
<td>P</td>
<td>16</td>
<td>P</td>
<td>16</td>
<td>BR</td>
<td>NA</td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td></td>
<td></td>
<td>Grade 4</td>
<td>26 emerald</td>
<td>Q,R</td>
<td>17</td>
<td>Q</td>
<td>17</td>
<td>BR</td>
<td>NA</td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td>28 emerald</td>
<td>S,T</td>
<td>18</td>
<td>S</td>
<td>18</td>
<td>BR</td>
<td>NA</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td></td>
<td></td>
<td>Grade 5</td>
<td>29 sapphire*</td>
<td>U,V,W</td>
<td>19</td>
<td>U</td>
<td>19</td>
<td>BR</td>
<td>NA</td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td>30 sapphire</td>
<td>X,Y</td>
<td>20</td>
<td>X</td>
<td>20</td>
<td>BR</td>
<td>NA</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td></td>
<td></td>
<td>Grade 6</td>
<td>31</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>BR</td>
<td>NA</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td></td>
<td></td>
<td>Grade 7</td>
<td>32</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>BR</td>
<td>NA</td>
<td>6.0</td>
<td></td>
</tr>
</tbody>
</table>

| 149 |
This table shows how these levels correlate to each other.

1. Reading Recovery™ is a registered service mark of Ohio State University
2. Rigby PM & PM Plus (Rigby, Barrington, IL)
4. Rigby PM & PM Plus (Rigby, Barrington, IL)
7. The Wright Group/McGraw Hill (Bothell, WA)
8. Success for All Foundation (SFAF), Baltimore, MD
11. DRP Degrees of Reading Power ©Touchstone Applied Science Associates (TASA), Inc.
12. SAT10 (Standford Achievement Test) scaled scores correlated to reading levels
13. STAR Reading, Renaissance Learning, Inc.

*Note: This is an example of one school division’s chart to manage the correlational leveling systems used by various publishers and educators. (Saint Paul’s Project for Academic Excellence, 2007).*
Appendix B

School Leader Use of Reading Measures Survey & Online Consent Form

SURVEY – SCHOOL LEADER USE OF READING MEASURES

Introduction:
This is a survey about data monitoring systems and the use of reading data on an informal reading inventory (*see additional information below). Consider the use of data monitoring processes and the use of an informal reading inventory given in your school to answer the following questions. Please select the answer that most closely reflects your current practice. This survey is anonymous. You will not be identified, so we ask that you provide your candid perspective. The survey will take about 7 minutes to complete.

*An informal reading inventory is an informal testing instrument which consists of leveled reading passages which are used to determine a student’s reading level. Each passage is to be read orally by the student who subsequently attempts to answer accompanying comprehension questions asked by the teacher after reading. The teacher takes a record of the reading behaviors while the orally child reads and codes the record for analysis to inform instructional planning. Examples of commercial reading inventories include published materials such as the Rigby Informal Reading Assessment, Developmental Reading Assessment (DRA), and the Qualitative Reading Inventory (QRI), just to name a few. Thank you for participating in this survey. Your feedback is important to educational research.

Consent Informed Participation Consent:
By completing this online survey, I agree to participate in a research study with an emphasis in the area of school leadership, data driven decision-making, and reading measures. The study will explore school leaders' self-ratings of actions on these topics and the use of reading measures in schools. This research is part of a doctoral candidate’s dissertation research study in the Educational Policy, Planning, and Leadership program with an emphasis in General K-12 administration at the College of William and Mary.

As a participant, I understand that my participation in the study is purposeful in that elementary principals and assistant principals volunteered and were selected with the intention of providing a representation of elementary school leader perspectives about the use of reading measures. Specifically, the study will focus on the use of an Informal Reading Inventory (IRI) in correlation with the predictive value on the Virginia State Standards of Learning Reading assessment in the third grade. I understand that as a participant I will be asked to participate in an online survey that will take approximately 7 minutes. I understand the responses I provide are confidential and that my name and the school name will not be associated with any results of this study. Further, I understand the survey results will have no identifying information as well. I understand there is no personal risk or discomfort directly involved with this research and that I am free to withdraw my consent and discontinue participation at any time. If I have questions or problems that may arise as a result of my participation in the study, I understand that I should contact Laura Estes, the researcher, at l1estes@cox.net or Dr. Tom Ward at
tjward@wm.edu. My completion of the survey signifies that I am at least 18 years of age and that I consent to participating in this research study.

THIS PROJECT (EDIRC-2017-12-13-12570-mxtsch) WAS FOUND TO COMPLY WITH APPROPRIATE ETHICAL STANDARDS AND WAS EXEMPTED FROM THE NEED FOR FORMAL REVIEW BY THE COLLEGE OF WILLIAM AND MARY PROTECTION OF HUMAN SUBJECTS COMMITTEE (Phone 757-221-3966) ON DECEMBER 19, 2017 AND EXPIRES ON DECEMBER 19, 2018.

A Please share the number of years you have worked as a school leader (either as an Assistant Principal, Principal, or combined years as both).

B Please share the number of years you have worked in this school district.

C Have you previously ever taught at the elementary level?

   ○ Yes (1)
   ○ No (2)

D Have you been trained on how to administer and interpret an Informal Reading Inventory?

   ○ Yes (1)
   ○ No (2)

E Have you ever administered an Informal Reading Inventory?

   ○ Yes (1)
   ○ No (2)

F Do you hold a reading specialist license/endorsement?

   ○ Yes (1)
   ○ No (2)
   ○ Not, but currently enrolled and completing coursework. (3)
Survey Please rate the following statements with the response that best describes your current individual practice as a school leader.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Always (1)</th>
<th>Sometimes (2)</th>
<th>Rarely (3)</th>
<th>Never (4)</th>
<th>Unsure (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have a process for regularly scheduled meetings with teacher teams or individuals to review reading data results from an informal reading inventory (IRI). (1)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>I ask teachers questions that emphasize student strengths and needs identified on an IRI on the use of visual information and what actions they take using this information for reading group instruction or regrouping. (2)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>My teachers administer and record IRI results at the required intervals across a school year. (3)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>I ask teachers questions that emphasize student strengths and needs identified on an IRI on the use of meaning and what actions they take using this information for reading group instruction. (4)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>Teachers move and change students in and out of groups regularly based on current reading data. (5)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>I ask teachers questions that emphasize student strengths and needs identified on an IRI on the use of structure and language information and what actions they take using this information for reading group instruction or regrouping. (6)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>When the IRI is administered, I review the reading level data. (7)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>I ask teachers questions that emphasize student strengths and needs identified on an IRI on comprehension of what is read and what actions they take using this information for reading group instruction or regrouping. (8)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>Teachers use the reading text levels obtained on the IRI to monitor student benchmarks across the school year. (9)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>I ask teachers questions that emphasize student strengths and needs identified on an IRI regarding phrasing and fluency and what actions they take using this information for reading group instruction or regrouping. (10)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>I ask teachers about the amount of movement between reading levels for students across a school year based on text reading levels identified on an IRI. (11)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
</tbody>
</table>
Teachers use reading data to consistently make adjustments to students in reading groups. (12)

I conduct regular data monitoring meetings across the school year with teachers. (13)

I ask teachers about their students' text reading levels identified on an IRI. (14)

I have a system for collecting IRI reading data across the year. (15)

Teachers keep students in the same reading groups throughout the school year. (16)

I ask teachers questions that emphasize strengths and needs identified on an IRI on concepts of print and what actions they take using this information for reading group instruction or regrouping (17)

Teachers do not move students in or out of groups (18)

I discuss the results of the IRI reading level data with teachers in a data monitoring system (data teams, literacy teams, data walls, etc.) after each administration of the IRI. (19)

I ask teachers about grouping students by text level identified on an IRI. (20)

**Open Response:**

Please share any other information you believe would be helpful in knowing how you work with teachers regarding the use of Informal Reading Inventory information.
REFERENCES


[electronic version]. Retrieved from the College of William & Mary’s Mental Measurements Yearbook online database.


doi:10.1080/10573569.2013.789783


doi:10.1207/S15324818AME1503_2


(375548771)


http://www.learningpt.org/pdfs/literacy/nationalreading.pdf


VITA
Laura Frances Estes

Birthplace: Buffalo, New York

Education

2018 College of William & Mary
Williamsburg, VA
Ed.D., Educational Policy, Planning and Leadership, General
Education K-12

1999 Virginia Commonwealth University
Richmond, Virginia
Master of Education, Reading Specialist K-12

1992 Radford University
Radford, Virginia
Bachelor of Science, Early Education and Middle Grades

Experience

2015- present Federal Title I Coordinator, Williamsburg, Virginia
2013-2015 K-12 Instructional Specialist for English, Reading/Language Arts
& Title I, York County, Virginia
2011-2013 Assistant Principal, York County, Virginia
2001-2011 Reading Recovery® Teacher Leader, Williamsburg, Virginia
2002-2010 Adjunct Faculty, College of William & Mary, Virginia
1999-2001 Reading Recovery® Teacher Williamsburg and Hanover County, Virginia
1993-1997 General Education Classroom Teacher, Hanover County, Virginia