PREDICTING STUDENT SUCCESS ON THE THIRD GRADE READING ACHIEVEMENT ASSESSMENT IN OHIO

Todd M. Cramer

A Dissertation

Submitted to the Graduate College of Bowling Green
State University in partial fulfillment of
the requirements for the degree of

DOCTOR OF EDUCATION

December 2010

Committee:

Patrick Pauken, Advisor

Susan Peet
Graduate Faculty Representative

Jamie L. Hollinger

Paul Johnson

Rachel Vannatta Reinhart

ABSTRACT

Patrick Pauken, Advisor

Since the passage of the first Elementary and Secondary Education Act in 1965, increasing reading achievement for all students has been a focus for our nation. Unfortunately, our country still has over 30 million citizens who are illiterate (Mukherjee, 2007). One of the challenges for schools is to accurately identify students in need of early intervention in an efficient and effective manner before students participate in high stakes reading assessments.

This study aimed to assist school leaders in identifying which diagnostic reading measures will best identify students at risk of not passing the Ohio Third Grade Reading Achievement Assessment. The purpose of this study was to examine the relationship of students' performance on five district diagnostic reading measurement tools – (1) KRA-L; (2) AIMS probes from second grade administration; (3) DRA literacy assessment; (4) STAR Reading Assessment from second grade; (5) standards based report card indicators – and their performance on the October administration of the Ohio Third Grade Reading Achievement Assessment.

Grounded in the theoretical framework of Teale and Sulzby (1992), as supported by Gunning (2006), the study examined four research questions focusing on student literacy data. The study revealed strong relationships between each of the instruments examined and the Ohio Third Grade Reading Achievement Assessment. Pearson Correlation coefficients indicated that the STAR SS and STAR PR have the strongest relationships with the Ohio Third Grade Reading

Achievement Assessment. Multiple regression results indicated the STAR PR and the AIMS best predict passage on the Ohio Third Grade Reading Achievement Assessment for the total sample. The final research question examined the predictive nature of each tool for each subgroup based upon disability, SES and race. The STAR PR was found to be the most accurate for students without disabilities, both SES groups (Low SES and Non-Low SES), as well as White students and minority students. AIMS was found to be additional predictor for minority students and students without disabilities. No tool emerged as predictive for students with disabilities.

While schools need to select an instrument that best meets the needs of their population, this study clearly found strong relationships between each of the instruments examined and the Ohio Third Grade Reading Achievement Assessment. With the exception of students with disabilities, significant models were generated for each subgroup. Overall the STAR PR was shown to have the strongest relationship with the Ohio Third Grade Reading Achievement Assessment.

ACKNOWLEDGMENTS

Although my journey through the Leadership Studies program has been longer than planned, it has been one worth taking. This journey would not have been possible and may have ended prematurely without the help and support of family, friends, co-workers and Leadership Studies professors.

I would like to thank Dr. Patrick Pauken for his time, patience and ability to balance being an advocate for my success while providing the honest, straightforward feedback needed to ensure the final product was a document I can take great pride in. I am grateful for your friendship and being able to have had the privilege of knowing you since your first semester at BGSU.

My wife Beth for the countless times you ensured me that I could finish this process and the sacrifices you made to provide me the time needed to complete my degree. I could not have accomplished this without your love and support.

My two wonderful daughters, Ali and Riley. I appreciate your understanding when dad needed time in the back room to work on his university homework.

Dr. Rachel Reinhart for your patience and being an exceptional educator as I returned to the program only to find myself in your class wondering if I had made the right decision. You assured me that if I worked hard and applied myself, you would ensure I would succeed. Your passion for statistics is evident and contagious.

Dr. Paul Johnson for your support and ability to view issues from multiple angles. You opened my eyes to viewing issues with a new perspective. I appreciate the resources you shared as well as your expertise and insightfulness.

Dr. Jamie Hollinger for sharing your practitioner lens and subject area expertise. I value your opinions and am grateful for the advice you gave throughout the process.

Dr. Susan Peet for sharing your perspective and asking thought provoking questions. Your advice concerning content enriched my experience. I am thankful for your involvement.

Dr. Cynthia Beekley for constantly reminding me that I needed to finish. For being my mother at work and pushing me each time we met.

Kathy Hott for your encouragement and flexibility which were of great assistance in finishing in a timely manner. You have always been there with a supportive shoulder.

My family for your support and interest over the past two years as I worked to finish. I appreciate your understanding during difficult times and the phone calls to check up on the progress being made.

Lastly, to the Leadership Studies students and faculty. I appreciate the friendliness of the four cohorts I had the privilege of taking classes with and the memorable dinner conversations. I cannot say enough positive things about the Leadership Studies faculty. From revalidations through my dissertation, everyone has been there offering their support. It has been a privilege getting to know each of you and I appreciate what each of you has done to make my journey a success.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I. INTRODUCTION

Background of the Problem

Since the passage of the first Elementary and Secondary Education Act in 1965, increasing reading achievement for all students has been a focus for our nation. This act, which was part of President Johnson's Great Society initiative, provided public schools additional funding including over 1 billion dollars to assist children living in high poverty areas. Before the passage of this legislation and without doubt since then, educators have focused countless hours on the teaching of basic literacy skills in an effort to close the literacy achievement gap. Even given the modern technology and teaching strategies that have been utilized over the past 75 years, over 30 million illiterate citizens currently reside in the United States (Mukherjee, 2007). Further illustrating this point, many prison systems plan for the number of beds needed in the future based on third grade literacy rates across their state (Shethar, 1993). For this reason, amongst many others, on January 8, 2002, President Bush signed Public Law 107-110, an act designed to close the achievement gap with accountability, flexibility, and choice, so that no child is left behind (NAGB, 2009). This legislation, which later gained notoriety as "No Child Left Behind", mandates states to adopt regulations and assessments designed to measure the yearly performance of all students enrolled in grades 3-8, in the areas of reading and mathematics. States were also charged with the responsibility of creating an accountability system whereby schools that received failing marks on their state generated report cards, which reflect the results of state mandated tests would be subject to consequences including school choice options. If a school is forced to offer school choice it can often have a devastating effect on a district as Ohio school funding permits state and local funding to follow children to their school of choice, which can impact a district not only financial terms but also in the need to

adjust staffing levels based on declining enrollment. Hence the importance of being successful on the state report card a tool many parents use when reviewing school choice options. Later in 2002, the state of Ohio passed Ohio Revised Code (ORC) Section §3301.07, setting forth the state requirements that would be enacted at the district level as well as guidelines for student assessment. In doing so, the aim was to ensure that "no child is being left behind" in the areas of literacy and mathematics. Various accountability measures were also established to help students, families, and educators in Ohio focus and achieve literacy goals.

One such accountability measure, the Ohio Third Grade Reading Achievement Assessment (OAA) is part of this enactment. The OAA and Ohio Graduation Test (OGT) are currently the only high stakes student achievement tests requiring passage in Ohio. Since they are high stakes tests, students must earn a "proficient" score in order to be promoted to the next grade level on the OAA and demonstrate proficiency on the OGT in the subject areas of Reading, Writing, Math, Science and Social Studies in order to graduate from high school. The OAA is written based on the academic content standards set forth by the Ohio Department of Education (ODE) and parallel assessments are administered twice a year to all third grade students attending public schools in the state. Although the practice varies from district to district, typically students are required to participate in both the fall and spring administrations of this test regardless of passage on the first administration. Since ODE only reports the highest score for each child, there is no risk involved in having a student who earned proficient or higher retake the assessment in the spring.

Beginning in 2009, Ohio changed the name of this testing instrument from the Ohio Achievement Test to the Ohio Achievement Assessment. The change was in name only as the content and skills assessed and scoring remain the same. This marked second time the name had

been changed as Ohio began in the 1990's by referring to state mandated tests as Proficiency Tests. With the change in name and legislation, Ohio has increased the accountability including public reporting of data at both the school and district level through annual report cards.

This accountability measure is critical to both schools and districts as the report cards are available to the public through ODE's website. Also, buildings and districts are mandated to hold yearly meetings to discuss the academic achievement levels which include reviewing the state issued report card. These report cards share basic demographic and academic information including the passage rate of students for each academic indicator assessed by ODE. To earn an indicator, a school and/or district must have at least 75% of the students participating in the OAA earn at least a proficient score on the given assessment. In this age of school competition, it is important for both schools and districts to have earned an excellent rating from ODE. Students in Ohio have the option of attending private, parochial, charter or public schools. Since private and parochial schools are not rated in this manner, it is critical for public and charter schools to have this ranking as an indicator of the quality of education they provide for their students. In order to earn this rating, a school must have earned at least 94% of the available indicators (ODE, 2009). Without an excellent rating, schools lose their competitive edge not only when comparing public schools against one another, but more importantly, when comparing public schools against their private and parochial counterparts who are not assigned school ratings. Hence it is important to ensure that each school and the entire district earn all the indicators and receive an excellent rating.

In Ohio all public school districts as well as individual school buildings receive annual report cards. While building report cards are based on the building results, district report cards are the accumulation of building level data from within the district. According to ODE, all

public schools and school districts are held accountable on the report cards for their performance in two ways (2009). One accountability measure is adequate yearly progress (AYP). AYP measures the performance of groups and subgroups (Black non-Hispanic, American Indian or Alaska Native, Asian or Pacific Islander, Hispanic, Multi-Racial, White non-Hispanic, Economically Disadvantaged, Limited English Proficient, Students with Disabilities) of students who have participated in the Ohio Achievement Assessment (OAA). If a school or district fails to make the projected goal set forth by ODE, the school or district is then put into year one of the school improvement process. If a school/district continues to not meet the goal level set forth by ODE for two consecutive years, the school/district is then subject to layered interventions that are in place by ODE to support the Ohio Improvement Process (OIP). OIP is designed as a framework that engages school staff, parents and community members in reviewing district and building data, then forming an improvement plan based on the data. These interventions include a second measure of accountability, schools and districts are responsible for meeting passage rate goals which are set by ODE for all students in order to earn indicators on the district and building report cards. Although the state requires 75% of students participating in the assessment to earn at least a proficient score, it should be noted that there is inconsistency as to the number of points a student must earn in order to achieve the proficient status. According to the ODE Department of Assessment (2010), in order to earn a proficient score on the October 2009 administration, a raw score of 33 must have been attained. However, just one year earlier it took a raw score of 30 to be designated as proficient. Thus, a student earning just 63% of the available points will be labeled as proficient on the 2008 administration. This scoring system has become a point of discussion with educators across the state as accountability for such results continues to be the focus of state and national media as many question the formula values and test validity (STEPS,

2004). If a school or district does not earn the necessary indicators, as defined by ODE, this is another avenue by which a district will become part of the Ohio Improvement Process (OIP).

The Ohio Improvement Process is a cyclical procedure containing four stages: (1) identify critical needs, (2) develop a focused plan, (3) implement the focused plan, and (4) monitor the improvement process. Once a plan has been formulated, districts can elect to make it their continuous improvement plan, on which progress is reported to ODE on a yearly basis. Districts are categorized as needing a low, medium, or high level of support from the State Support Team (SST) or Educational Service Center (ESC) depending on the level of complexity based on the percentage of student groups not meeting AYP (ODE, 2009). Districts can voluntarily use this process as a school improvement model even if their students are making AYP. OIP requires districts to participate in three full days of training which include teachers, administrators and community members, then continue with monthly follow up meetings as districts implement and monitor their plan. It is a time consuming process designed to focus district resources and open up communication between districts and stakeholders in an effort to improve the quality of education.

By utilizing the OIP process, districts are able to engage their constituents in a data driven process that encompasses all of the accountability measures ODE has implemented for schools and school districts. At the building level, individual schools are responsible for earning indicators, which primarily focus on achievement test scores. The individual building scores are then combined, forming the district report card. Both the building and the district report cards are published on an annual basis and available to the public via the ODE website. Based on the number of indicators earned, each building and district in the state is assigned a rating ranging

from Academic Emergency to Excellent with Distinction. This accountability system holds all school and school districts publicly accountable for the academic achievement of all students.

Rationale for the Study

Beginning in kindergarten, students in Ohio are given state-mandated literacy assessments designed to help identify children at risk of not achieving state set grade level literacy standards. Districts often assess students with additional instruments, above and beyond those identified as mandatory assessments by the state department of education. These assessments, both state-mandated as well as others that have been implemented at the district level, are key in providing data to principals and central office administrators to aid in the allocation of resources, selection of curriculum materials and the evaluation of the effectiveness of classroom instruction. More importantly, classroom teachers use these assessment data to inform individual student instruction and make classroom level decisions. A difficulty with these assessments is that they often overlap one another as several tools may be used to measure the same literacy skill. Also, the administration and scoring of these instruments often take up valuable time that a classroom teacher would otherwise be spending planning lessons or teaching students. Finally, other than the OAA, all other district-administered instruments are costly, with some districts spending well over $100,000 on a yearly basis to collect student literacy data. Each of these measures are administered in an effort to identify those at risk of not making grade level standards and to allow for early literacy interventions to take place as well as to identify students who may need differentiated instruction to ensure their academic needs are met. However, even given these efforts, 37% of fourth grade students in the United States are not reading at grade level (NAEP, 2007). If educators could correctly identify within the first few

weeks of school the students in need of literacy remediation, effective early intervention would be provided.

The positive impact of early intervention can greatly increase the likelihood that students not making adequate progress in reading will receive the additional support needed to reach grade level expectations. If early intervention is not implemented, and students remain lacking the needed early literacy skills, consequences include academic underachievement, poor self esteem, behavioral difficulties and low motivation (Sloat, Beswick & Willms, 2007). School districts have also implemented their own measures to help identify student who are at risk of not passing the third grade OAA. Although intervention itself is a difficult task given the complexity and behaviors involved, in order to become a good reader, early intervention remains critical to ensure student success (NAEP, 2009). The challenge of identifying which students are in need of these early interventions continues to be difficult for many schools and districts.

Currently, there is a very limited body of research available to guide the decision making of school personnel when considering the implementation of the diagnostic tools that were examined in the present study. While various studies (Hollinger, 2009; Rescorla & Rosenthal, 2004; Riedel, 2007; Rouse & Fantuzzo, 2006) have been conducted examining reading instruments such as DIBELS, KRA-L, and standardized tests (Terra Nova, Iowa test of basic skills, MAT – 6), to date, there have not been any studies examining the predictive nature of the Developmental Reading Assessment (DRA), STAR, AIMS fluency probes, or standards based report card indicators in comparison to the Ohio Third Grade Reading Achievement Assessment. These district diagnostic reading measurement tools are important since educators rely on them as a key source of data when examining the skill sets of each child. Since they are diagnostic in nature, the results are used to inform instruction and ensure the academic needs of each student

are met as opposed to achievement tests which are summative in nature. This study will fill this gap and enable school leaders to make an informed decision based on research before deploying the use of limited financial and human resources and possibly implementing a new diagnostic assessment instrument that may not help guide classroom instruction and improve student literacy achievement.

Purpose of the Study

The purpose of this study is to examine the relationship of students' performance on five district diagnostic reading measurement tools and their performance on the October administration of the Ohio Third Grade Reading Achievement Assessment. This study was designed to assist school leaders in identifying which diagnostic reading measure will best identify students at risk of not passing the Ohio Third Grade Reading Achievement Assessment. In doing so, it will allow leaders to begin early intervention with students in need, knowing the proper students have been identified by an instrument that is predictive of the Ohio Third Grade Reading Achievement Assessment. The five district diagnostic reading measurement tools used in this study include: (1) Kindergarten Readiness Assessment- Literacy (KRA-L); (2) Second grade AIMS fluency probes; (3) Developmental Reading Assessment (DRA); (4) Second grade district created standards-based report card; and (5) STAR reading assessment at the second grade level.

The sample of this study is composed of 341 students in grades three and four who attend suburban public elementary schools in an Ohio school district. All of the participants in the study have taken part in a minimum of four of the five reading instruments being explored.

Research Questions

1. What is the relationship between the score attained on a district diagnostic reading measurement tool and students' Ohio Third Grade Reading Achievement Assessment?

2. What is the minimum score needed on each district diagnostic reading measurement tool (DRA, AIMS WEB fluency probes, KRA-L, standards based report card, STAR Reader) to predict passage on the Ohio Third Grade Reading Achievement Assessment?

3. Which district diagnostic reading measurement tool (pass/no pass) best predicts passing the Ohio Third Grade Reading Achievement Assessment?

4. What district diagnostic reading measurement tool (DRA, AIMS WEB fluency probes, standards based report card, STAR Reader) best predicts the passage on the Ohio Third Grade Reading Achievement Assessment for children in each subgroup (African American, Multi-Racial, Hispanic, Low SES, IEP)?

Theoretical Framework

This study will be grounded in Teale and Sulzby's (1986) theoretical framework for literacy development and Gunning's (2006) study of closing the reading achievement gap. Teale and Sulzby (1986) propose that literacy development originates from five primary roots, each of which builds upon the prior. The five roots are as follows: Root One: Development of print awareness in situational contexts; Root Two: Development of print awareness in connected disclosure; Root Three: Development of the functions and forms of writing; Root Four: The use of oral language to talk about written language; Root Five: Metacognitive and metalinguistic awareness about written language (Teale & Sulzby, 1986, p. 6-11). Further examining this framework, Root One is based in the observation that print awareness has been developed by the student.

It is typical for students who are beginning kindergarten to function at the Root One stage, being able to read environmental print in context. An example of this would be recognizing "McDonald's" or referring to toothpaste as "Crest" (Teal & Sulzby, 1986). During Root Two children understand the orientation of books, and demonstrate general understanding of terms such as read and story. In application, students understand that print flows directionally from left to right and words go to together to form thoughts. Root Three begins to focus on the development of written language, and the receptive process involved as children learn to read and write. These processes though largely unmeasured, are critical to the development of the academic skills of children. Root Four continues to focus on language development; however the focus of this root is upon the translation of reading skills to the spoken message. During this stage children make the connections between letters and sounds they can read and begin to apply this knowledge in their writing. Root Five, the pinnacle of Teale and Sulzby's work concerning literacy development, focuses on metacognitive and metalinguistic features of language development (1986). Children focusing at this higher level are able to analyze and explain language (Teale & Sulzby, 1986).

Teale and Sulzby's work has been included in various other literacy studies (Anderson & Matthews, 1999; Deloache, 1991; Nathan & Stanovich, 1991; Pinnell, Lyons, DeFord, Bryk & Selzer, 1994; Purcell-Gates, 1996; Whitehurst & Lonigan, 1998) examining emergent literacy and the subsequent development of literacy skills. Similar to these authors, Thomas Gunning (2006) has developed a parallel theory on language development and expands upon it by addressing the need to close the achievement gap that exists between white students and their underrepresented peers.

Gunning (2006) focuses on language as the basis for all reading and writing skills. The challenge is overcoming the deficits that are created at the earliest ages, before children enter formal schooling. As Hart and Risley (1995) discovered, by the age of four, children from professional families have been exposed to 45 million words while those from a low socio-economic status (SES) have only heard around 13 million words. Thus even before entering school a gap of over 32 million words exists. Gunning (2006) addresses classroom strategies that should be utilized to assist students in overcoming their barriers. It all begins with the development of vocabulary. The difference between children entering school with a vocabulary at the 25th percentile and those at the 75th percentile is the equivalent of two years of vocabulary development (Gunning, 2006). Unfortunately, schools tend to widen this gap as students participate in the educational process and go without the remediation necessary to help them acquire the same literacy skills as their peers. Continuing, his framework highlights, students who have a vocabulary gap as defined by the number of words they know and understand, which does not close during their first several years of instruction often develop a comprehension gap, meaning students do not have the skills to understand a piece of writing, due to their vocabulary weakness.

To close this gap, Gunning (2006) suggests four key steps: (1) Instruction in words as needed; (2) Planned introduction of words; (3) Individual word study; (4) Learning to learn words. Ironically these concepts have been incorporated into the publication of nearly every language arts curriculum offered for elementary age students (Houghton Mifflin, 2010; Literacy By Design, 2009; Scott Foresman, 2010). By engaging in these activities daily, students will make the gains necessary to close the achievement gap. Although other strategies and methods may accomplish the same goal, these four key steps to build student vocabulary, have been the

focus of many supporting studies (Beck, McKeown & Kucan, 2002; Biemiller, 2004; Swanborn & de Glopper, 1999).

The research based on the theoretical framework of Teale and Sulzby (1986), as well as Gunning (2006), demonstrates the importance of acquiring literacy skills in a logical, sequential manner that is critical in order for children to be successful in elementary school. If this foundation is not properly built, children will be at risk of not developing the adequate literacy skills necessary to perform on state mandated assessment measures.

Significance of the Study

Through a thorough examination of the selected assessments designed to measure literacy skills, school leaders and educators will identify and implement the best use of each of the assessments studied. Leadership will then be able to use the outcomes of this study to support their decisions regarding which instrument to administer to effectively and efficiently identify students for early intervention. Those instruments that do not have a strong relationship or predictive validity to the Ohio Third Grade Reading Achievement Assessments may no longer need to be administered, saving valuable district resources including both human and monetary capital. Educators will focus on instruments that are proven to have a strong positive relationship and predictive validity to help all students maximize their learning potential. In doing so, limited resources could potentially be spent on other expenditures.

The study will enable school leaders to make decisions regarding the administration and interpretation of data gathered from the reading instruments used in this study. Consequently, students learning in schools that implement the findings of this study, combined with a research based literacy program will be able to reach excellence as defined by ODE. In turn, schools concerned with meeting AYP will also be able to move forward and reach their goals for each

subgroup ideally leading to the assignment of an excellent rating by ODE. These goals will only be accomplished through the appropriate use of the diagnostic tools as they provide information to teachers that inform instruction. As part of this informed instruction, teachers will be able to provide students who are performing below the identified standard with intervention, therefore strengthen their literacy skills and enabling them to have the competency level necessary to pass the achievement tests.

Definition of Terms

Adequate Yearly Progress (AYP). A measure implemented by the State of Ohio aligned with a Federal mandate that holds schools and school districts accountable for the performance of all students including those in federally defined subgroups (ODE).

AIMS Web Fluency Probes. A diagnostic reading instrument designed to measure the number of words a student correctly reads in one minute.

Curriculum Based Measurement (CBM). A classroom assessment tool designed to measure skills based on curriculum taught at the classroom level.

Diagnostic Reading Assessment. A reading assessment that identifies student strengths and weaknesses to inform classroom reading instructional decisions (ODE, 2009).

Developmental Reading Assessment (DRA). A reading assessment which requires a child to read a selection, then answer comprehension questions both orally and in written form.

High Stakes Assessments – Assessments developed by the Ohio Department of Education that require a student to earn a proficient score.

Kindergarten Readiness Assessment for Literacy (KRA-L) - A mandated assessment tool in the state of Ohio which must be completed within the first six weeks of school. It is designed to measure the literacy readiness skills of students entering kindergarten.

No Child Left Behind (NCLB). A federal law passed in 2002 designed to hold school and districts accountable for performance based results, focus resources on effective programs and provide school choice for parents (ODE).

Ohio Department of Education (ODE). A governmental agency created to oversee the education of children residing in the state of Ohio.

Ohio Reading Achievement Assessment (OAA). A required reading assessment designed to measure what a student knows and can do beginning in grade 3 (ODE, 2009).

Report Card Indicators. State of Ohio indicators that have been selected as points of emphasis and on which progress is reported at least four times per year. These include achievement tests given in grades 3-8, the Ohio Graduation Test, attendance rate and graduation rate.

Socioeconomic Status (SES). A measure of an individual or family's relative economic position. A child's SES status in Ohio public schools is determined by the family's yearly/monthly income in relation to the number of members in the household. It is typically a minimal estimate due to the documentation required for identification (ODE, 2009).

STAR Assessment. A computer based assessment designed to determine a proximal reading level for individual students.

Student Achievement Subgroups. Students are grouped by a demographic or other characteristic such as race/ethnicity, gender, economic status, language background and or disability.

Limitations

While this study is a comprehensive examination of the predictive ability of five reading measures in comparison to the Ohio Third Grade Reading Assessment, it does contain limitations. This study is limited geographically, as the sample comes from only one district in the state of Ohio. School practitioners should use caution in their interpretation of this study's results given its limited generalizibility. Another factor that is not controlled for is teacher quality. While the literature is strong connecting teacher quality to student performance (Defour, Defour, Eaker & Many, 2006; Ingle, 2009;Marzano, 1996) each student in the study will come from a unique progression of teachers as well as buildings and districts due to the transient nature of the student population of the district involved in this study. Also, a uniform measure of teacher quality is not in existence in conjunction with the study site. Consequently, teacher quality is a variable that is not be accounted for as part of this study. The final limitation which needs to be considered is the changing nature of the Ohio Third Grade Achievement Assessment. The Ohio Department of Education has informed districts that the Reading Achievement Assessment as well as the content standards upon which the OAA is based will be replaced by 2014. Therefore this study will need to be applied with caution as it only examines current Ohio Third Grade Reading Achievement Assessment data.

Organization of the Remaining Chapters

This study is organized as follows: Chapter 1 is comprised of the background of the problem, rationale for the study, purpose, research questions, theoretical framework, and significance of the study, definition of terms and limitations of this study. Chapter 2 is a literature review on reading assessment instruments as well as the developmental aspects of reading as they pertain to this study. Chapter 3 is an explanation of the research methodology

used, data collection, and procedures. Chapter 4 presents the descriptive narrative of the study's results and an analysis of the data. Finally, Chapter 5 summarizes this study's major findings and includes recommendations for future research and policy implications.

CHAPTER II. LITERATURE REVIEW

Since the adoption of the No Child Left Behind Act in 2002, every state in the United States has developed a series of policies, including standards, assessments, and accountability measures designed to gauge the literacy skills of all students attending publicly funded schools. Beginning in third grade in the state of Ohio, students must pass the state mandated reading assessment to be promoted to fourth grade. While some districts choose to exercise other options as provided by the state law, others tightly adhere to the law and strongly consider retention for those students not successful on this assessment. Consequently, it is vital for schools to provide high quality, effective intervention for students at an early age to assure they acquire the skills necessary to pass the Ohio Third Grade Reading Achievement Assessment. It is known that students enter schools with differing levels of literacy readiness skills and these initial performance levels have an impact on third grade reading performance (Foster & Miller, 2007). Thus it is imperative for districts to have confidence in the diagnostic tools that are being used to identify which students are in need of remediation as well as those in need of acceleration. Data taken from these diagnostics should be at the core of the decision making process. In the absence of quality diagnostic instruments, students are not able to be identified and provided the aggressive support they need, and the result can often be a phonics/decoding gap translating into a text comprehension gap (Foster & Miller, 2007; Gunning, 2007). Equally important, professional educators must be engaged in continuous professional development to remain current on literacy practices as well as to maintain a base level of knowledge which is necessary to make decisions regarding student progress and possible identification of literacy challenges. School leadership must embrace the use of data in all aspects of their field including using data to monitor student progress, form and implement professional development opportunities as well

as evaluate the effectiveness of the instructional programs at their respective school and districts as well as remain current regarding classroom instructional practices and beliefs.

This chapter will examine the existing research on issues that must be considered by leadership before making decisions that will impact student achievement. To begin, No Child Left Behind is explored to provide a framework explaining the basis for school accountability. Several predictive models follow and diagnostic measures will be identified and examined, each with decisive benefits as well as areas of weakness. Next, studies focusing on assessment that guides instruction will be examined. Issues specific to subgroups will be explored giving attention to students in subpopulations such as SES, African American, and students with disabilities. Lastly, the impact of teacher quality and efficacy will be presented focusing on the impact of the classroom educator on the effectiveness of identifying students in need of intervention as well as providing effective intervention. Following the presentation of the research studies, critical findings will be summarized followed by a discussion of the limitations and gaps that exist in the current research which are addressed in this study and future studies.

No Child Left Behind

No Child Left Behind is the current incarnation of one of the principle pillars of President Lyndon Johnson's War on Poverty, The Elementary and Secondary Education Act of 1965, which created the Title I federal aid program aimed at reducing achievement gaps between rich and poor and among the races (Mantel & Greenblatt, 2007). On January 8, 2002, President Bush signed into law Public Law 107-110, an act designed to close the achievement gap with regard to accountability, flexibility, and choice, so that no child is left behind (NAGB, 2009). This legislation which later gained notoriety as “No Child Left Behind” (NCLB) mandates states to adopt regulations and assessments designed to measure the yearly performance of all students in

the areas of reading and mathematics. Although this was a new mandate from the federal government, Ohio already had an assessment program in place whereby students participated in math and reading assessments beginning in fourth grade. States are also charged with the responsibility of creating an accountability system whereby schools that received failing marks would be subject to consequences including needing to offer students the option of attending another learning institution. Later that year, Ohio passed legislation setting forth the state requirements and mandates that would be enacted at the district level as well as guidelines for student assessment. As part of this enactment, Ohio created the Third Grade Reading Achievement Assessment. Other than the Ohio Graduation Test (OGT), this test is the only other high stakes test given in Ohio. Since it is a high stakes test, students must earn a "proficient" score in order to be promoted to the next grade level. This assessment is designed using the academic content standards set forth by the Ohio Department of Education (ODE) in 2001.

In alignment with the standards based assessments, Ohio also created a multi-tiered accountability system whereby district/schools are assigned an initial designation depending on the number of state indicators earned. Table 1 displays the indicators each building/district was able to earn during the 2009 – 2010 school year.

Table 1

State of Ohio Report Card Indicators (2009 – 2010)

Grade	Test(s)	Standard
3	Reading and Math Achievement	75%
4	Reading and Math Achievement	75%
5	Reading, Math and Science Achievement	75%
6	Reading and Math Achievement	75%
7	Reading and Math Achievement	75%
8	Reading, Math and Science Achievement	75%
OGT (10th)	Reading, Writing, Math, Social Studies and Science Achievement	75%
OGT (11th)	Reading, Writing, Math, Social Studies and Science Achievement	85%
	Non- Test Indicators	
K-12	Attendance Rate	93%
12	Graduation Rate	90%

(ODE, 2010)

Although the name has changed three times, the Ohio Achievement Assessments continue to be the benchmark assessment against which the effectiveness of classroom instruction is marked. These state mandated assessments are not a new concept in Ohio. Since 1993 students in Ohio have been participating in state mandated assessments. That year the state implemented a series of assessments labeled as proficiency tests, designed to measure whether a student was competent in the areas assessed. Since this time, despite numerous legal challenges

(Draper, 1994), the state has continued a state wide assessment program. However, it should be noted that the test has been transformed from one designed to measure minimal competencies to one that is intended to not only measure a minimal skill base, but also examine the growth in individual student's skills. This value added growth model, which is the focus of Ohio Revised Code §3302.021, provides another element of concentration and accountability for Ohio educators.

Value added focuses on comparing the achievement scores, which for the purposes of mathematical calculations are converted to Normal Curve Equivalent (NCE) scores for a given student from one year to the next. In doing so, ODE calculates a growth score for each child, classroom, building and district. A value added score of zero equates to a year of growth. This is an important accountability measure for educators as positive and negative value added scores can also impact their school and district ratings. If school/district scores exceed the standard error in a positive direction for two consecutive years, the school/district will earn having their state designated performance level improved by one rating level. Consequently an effective school can be promoted to excellent through the value added calculations. In order to earn the highest rating available from ODE, Excellent with Distinction, it is necessary for a district to have earned a positive value added score for two consecutive years in addition to earning at least 94% of the indicators (ODE, 2010). For those schools/districts that exceed the standard error in a negative manner for two consecutive years the result is the demotion of one level in regard to the state rating system. Due to these implications, value added has become a point of focus for all schools and districts in Ohio. Educators must not only ensure that every child reaches the minimal competency level, but also make certain that those who enter their classrooms with adequate and advanced grade level skills continue to experience academic growth.

Another key component of NCLB focuses on the requirement that all public schools assess all students in grades 3-8 annually in mathematics and reading (Ravitch, 2009). It also stipulates all students must be "proficient" in reading and math by the end of the 2013 – 2014 school year. This same law allows for states to create their own assessments and set their own guidelines in defining proficiency (Ravitch, 2009). Consequently each state has devised their own assessment practices and procedures as well as methods used to calculate passage on these high stakes assessments. Until a national assessment is devised, states will continue to research their own assessment questions and administer their own assessments to ensure compliance with NCLB..

NCLB mandates each state to create an accountability system which must include components such as adequate yearly progress (AYP). Districts must set targets that result in 100% passage on reading and math achievement tests by 2014. Ohio, like other states submitted a plan to the federal department of education along with subsequent revisions. NCLB, which was passed by a bipartisan Congress, was modeled in a large part after legislation which Texas had passed and implemented between 1993 and 2002. Every few years it is mandatory that Congress reauthorize the Elementary and Secondary Education Act. However, in 2002, the political climate was unique in that Republicans controlled both the House and the Senate, thus giving President Bush the ability to more aggressively push issues central to his campaign. However, even in these strong economic times, the United States was engaged in military endeavors which were consuming any economic surplus, leaving little financial support for the mandates set forth in NCLB. This lack of financial support has been a persistent point of criticism. While no piece of legislation addresses all concerns, NCLB was crafted in a manner

that provides structures, but does not mandate standardization from state to state or provide additional funding for the assessment programs that each state is accountable for producing.

Educational policy prior to NCLB was written in broad terms, intended to allow states to regulate their own educational systems. However, with the dynamics of our country changing and special area groups gaining power, it was clear that change was on the way. Our education system was and continues to have a traditional mindset while students continue to struggle to make the progress necessary to be successful. As our country becomes more diverse every day, our educational system is struggling to find ways to best educate all children. Discrepancies continue to exist when examining factors that attribute to a quality education. According to government statistics, only 70% of our schools are meeting AYP (retrieved from http://ed.gov/nclb/accountability/results/progress/nation.html). Table 2 presents the 2007 (National Assessment for Educational Progress) NAEP scores for fourth grade reading.

Table 2

NAEP Reading Achievement Test Results

NAEP Reading Achievement for 2006 – 2007		
	% 4th Graders Basic	% 4th Graders Proficient
All	66%	32%
White	77%	42%
Black	46%	14%
Hispanic	49%	17%
Low Income	50%	17%

(NAEP, 2007)

Unfortunately, Ohio's scores are often at or below the nation's averages. These data have had a profound impact on our country as even our students deemed successful often enroll in college and need remediation. It has been found that over 60 percent of first-time community college students in the National Education Longitudinal Study took at least one remedial course, and 29 percent of first-time students in public four-year institutions also enrolled in at least one such course (Bailey, Jenkins, & Leinbach, 2005). Employers also share that as many as 40% of high school graduates do not have adequate skills to function in the workforce (retrieved from http://www.principalspartnership.com/feature207.html). Although it has not been formally measured, the financial impact of these factors on our economy is devastating. It is clear that in an age of global competition, the United States is not positioning itself in a desirable manner.

Predictive Measures of Literacy Skills

In order for early intervention to occur, students must first be identified as needing such assistance. Since teacher referral is the single most powerful variable in determining whether a student will be identified as at risk and receives specialized instruction, it is critical to investigate the accuracy of teacher perception of student progress (Algozzine, Ysseldyke, & Christenson, 1983; Ysseldyke, Vanderwood, & Shriner, 1987). These studies suggest that teachers' ratings of student achievement are not strongly correlated with actual student achievement levels as measured by standardized tests further substantiating the need for an unbiased measure to identify at risk students. Compounding this issue, teachers tend to overestimate the amount of academic progress students actually make (Graney, 2008). Thus it further emphasizes the importance of standard assessment measures being used to assure adequate student progress is being made.

Ohio has implemented a series of assessments for this purpose to collect unbiased data on student academic skills. One such assessment, the Kindergarten Readiness Assessment for Literacy (KRA-L), must be administered to all kindergarten students within the first 6 weeks of school. This section will examine some of these diagnostic reading assessment instruments as well as published literature associated with each instrument.

Kindergarten Readiness Assessment – Literacy (KRA-L)

The Kindergarten Readiness Assessment – Literacy (KRA-L) is a diagnostic instrument developed by the state of Ohio designed to measure the early literacy skills of kindergarten students prior to or within six weeks of entering school. It is designed to be a quick screening instrument to measure oral language, rhyming, letter identification and alliteration. These elements were selected as they are essential to the reading development of children (ODE, 2008). Teachers can utilize these data to inform their instructional decisions and provide engaging lessons that lead to student success. Since it only examines four qualities of reading readiness it should not be used as the sole measurement of a student's academic abilities. However, oral language, rhyming, letter identification and alliteration are well researched concepts that correspond to early literacy success (Adams, 1990; Donaldson, 1978; Snow, Burns, & Griffin, 1998; Whitehurst & Lonigan, 1998).

These elements are each essential to the development of a solid literacy foundation. Oral language is typically developed through verbal interactions with an adult. As demonstrated, many researchers found there is a significant positive relationship between a child's ability to correctly pronounce words and their ability to acquire literacy skills (Jarmulowicz, Taran & Hay 2007; Bishop & Adams, 1990; Catts et al., 2001; Scarborough, 1990). Consequently, the KRA-L measures oral language through a series of interactions with students. These interactions

include recognizing sound patterns in words as well as the ability to use words out of context. The skills are measured through two separate activities, each designed to assist in measuring the oral language skills of children entering kindergarten.

Phonemic awareness, which is the ability to distinguish sounds, was the focus of Foster and Miller (2007) as well as Nation and Snowling (2004). Specific developmental trajectories for phonics and early text comprehension have been examined with students in grades kindergarten through third. At the conclusion of these studies it was found that students in the low group did not match the scores attained by the higher two groups in first grade until they were in third grade. It was also found that although the phonics gap was virtually closed by third grade, a significant comprehension gap or lack of ability to understand what they have read, had formed. It is conclusive that the development of early literacy skills is significantly influenced by the development of early phonemic awareness skills. The KRA-L measures this skill by asking the student to identify the beginning sound of two words.

Letter identification is also a critical skill in learning to read as it has also been found to be a predictor of early learning success (Snow, Burns, & Griffin, 1998). There is evidence that these early learning skill deficits that are evident at age four can carry over to second grade and beyond (Molfese, Modglin, Beswick, Neamon, Berg, Berg, & Molnar, 2006). The KRA-L assesses this early academic skill by presenting a series of upper and lower case letter children should be able to identify by name. The proctor scores each response given by the child as the letters are presented. Although no one foundational skill is vital in order for a child to obtain adequate literacy skills, the identification of letters is often considered a reflection of a child's exposure to literacy concepts and often foretells the early academic success or struggle a child will have, thus making it a critical skill to acquire.

Just as it is important to provide a continuum of instructional materials for children to obtain the necessary literacy skills, it is also essential to have a continuum of assessments to help guide instructional decisions. In their study, Good, Simmons, and Kame'enui (2001) focus on assessment in a framework of early intervention. In order to provide the intervention, reliable and valid measures must be in place that can be trusted to identify children at risk of not being successful. Many districts in the state of Ohio use a copious number of instruments for the same goal. These instruments include but are not limited to AIMS Web, DIBELS, STAR Reader, DRA, Developmental Reading Inventories, and other district created assessments. Although these instruments each have qualities that differentiate themselves from one another, there are common qualities that emerge as each is examined. Case in point, oral reading fluency (ORF), measured by the number of minutes a student can correctly read in one minute, is one of these components (Wolf, 1999).

AIMS Web

AIMS Web Fluency Probes are orally read passages taken from Edformation's standard reading assessment passages, which contain well written, connected text. These ORF probes are designed to be an easy to use, quick measure that provides information around which educators can develop targeted lessons and provide students appropriate instructional materials. These probes provide data that is easily understood by both educators and parents and can be monitored in an efficient way. AIMS Web probes only take one minute to complete, thus making them a viable ORF instrument to use as a progress monitoring tool.

ORF has also been the subject of other various studies (Baker, Smolkowski, Kratz & Fien, Seeley, Kame'enui & Beck, 2008; Good, Simmons & Kame'enui, 2008; Reschly, Busch, Betts, Deno & Long, 2009; Taub, McGrew & Keith, 2007; Wang, Porfeli & Algozzine, 2008) as

it has been found to be an accurate predictor of success on reading achievement tests (e.g., Buck & Torgesen, 2003; Crawford, Tindal & Steiber, 2001; Hintze & Silberglitt, 2005; McGlinchey & Hixson, 2004; Shapiro, Keller, Lutz, Santoro, & Hintze, 2006). According to the meta analysis of Reschly, Busch, Betts, Deno & Long, a statistically significant estimate of the z-transferred correlation was found with respect to state specific achievement tests (N=70), $\gamma o=0.77$, $t(139)=46.92$, $p<.001$. Based on these statistics it can be concluded that ORF is a significant predictor of success on state achievement tests and should be taken into consideration when making instructional and evaluative decisions regarding student progress.

Further supporting this meta analysis, fluency has been found to be a predictor of success on subsequent reading skills (Good, Simmons & Kame'enui, 2001). Without adequate fluency, or the ability to read written words quickly with accuracy and expression, students are not able to progress and develop the comprehension skills necessary to pass state mandated achievement tests. In a study conducted by Fuchs, Fuchs and Maxwell (1988) it was found that ORF correlated higher with reading comprehension (.92) than any other measure. Expanding on the notion that ORF is an accurate predictor of standardized tests, other research has found that ORF correlates with the Oregon high stakes test as well as the SAT – 10 high stakes test were typically between .60 and .80 (Baker, Smolkowski, Katz and Fien, Seely, Kame'enui & Beck, 2008).

Although ORF is a well researched and documented method by which educators gain information to guide their instructional decisions, there are several other tools and methods commonly used.

STAR Reading Assessment

One such instrument is the STAR reading assessment, an online instrument produced by Renaissance Learning. Although frequently used in conjunction with Accelerated Reader, the STAR Reading Assessment is designed to measure the reading level of students beginning in first grade and continues through middle school. This assessment is computer based and is designed to last ten minutes in administration time. According to the National Center for Response to Intervention (2008), this instrument has been found to be able to be generalized to a broad population as well as a reliable and valid instrument for minority populations.

In completing this instrument, students are first assigned an instructional grade level, typically the grade level in which they are enrolled. The assessment begins by asking the student to read a short passage then respond to multiple choice questions. Depending on whether the answer given was correct or incorrect the following question will be more complex or less complicated. This process of question engagement and adjustment continues until the student arrives at their instructional level as determined by their responses. Educators can then access the results which provide a color level which can be used with Accelerated Reader, a sister program as well as a reading level span indicating the results of the assessment. The STAR Reading Assessment is designed to be administered up to three times per academic school year.

Report Card Indicators

Aligned to the State of Ohio language arts course of study, educators use a standards-based report card as a tool when making important instructional decisions as well as a reporting mechanism to demonstrate individual academic progress to parents. A standards based report card is comprised of state standards that have been selected by teachers as the most critical for academic success for each subject. Standards based report cards are a key provider of

information when deciding whether a student is promoted placed or retained at the end of each school year. The indictors that are currently used on the second grade report card include: (1) Integrates strategies to read unknown words (visual, structural, context, phonics, decoding); (2) Applies word attack skills; (3) Recognizes high frequency words; (4) Identifies vocabulary that is critical to the reading; (5) Retells a story that includes: story elements, characters, setting, problem, sequence of events, and solution in a fiction selection; (6) Identifies topics, main idea, and details of a nonfiction selection; (7) Answers literal, inferential, and evaluative questions to demonstrate comprehension (Springfield Local Schools, 2010).

Developmental Reading Assessment (DRA)

The Developmental Reading Assessment (DRA) enables educators to observe, record, and evaluate changes in students reading performance (Beaver, 2006). It is administered at least twice a year and results are reported to parents via progress reports. The DRA requires a student to first read a selection aloud, permitting the teacher to perform an oral reading fluency assessment. If the student demonstrates adequate reading fluency, he/she is asked to answer a series of scripted comprehension questions. Following this oral assessment, beginning in second grade students are then required to answer additional questions through a written response. This process continues until the student is unable to pass the assessment at which point the last successfully completed reading level becomes their instructional reading level. Individual student results are plotted on a growth chart and are part of the student's assessment portfolio that accompanies them throughout their school career.

Third Grade Ohio Reading Achievement Assessment

Since October of 2003 the Ohio Department of Education has administered a Reading Achievement Assessment to all students who participate in third grade reading instruction. This assessment is designed to measure a student's literacy skills specifically in the areas of acquisition of vocabulary, reading process, informational text and literary text. Scores achieved on this measure are then placed into one of the five state categories (below basic, basic, proficient, accelerated, advanced). A score of 400 has been set as the minimum score necessary to earn a proficient on the third grade reading achievement assessment. Each year the state sets the cut points for each of the categories (below basic, basic, proficient, accelerated, advanced) however a 400 is always the minimum score necessary to earn a proficient designation.

Due to the fact that this assessment is identified as a high stakes test, school districts have been required to adopt policies regarding students not able to earn at least a proficient score on this assessment. The Ohio School Board Association (2010) currently suggests the following three options be adopted into district policy for students who score in the basic or below basic category:

(1) promotion to the next grade if the principal and reading teacher agree that other evaluations of the student's work indicate the student is academically prepared for the next grade;

(2) promotion to the next grade with "intensive intervention" in that grade or

(3) retention in the current grade.

Given these three options, most schools choose to provide "intense intervention" to students in order to ensure they are able to continue on to the next grade level.

Subgroup Effects

Numerous studies have documented the impact belonging to a subgroup has on the reading achievement levels of those belonging to the group. However there is variance within subgroups, thus all subgroups cannot be unilaterally examined. An in depth analysis of the evidence reveals the effects vary according to the subgroup (i.e. African American, Low Socio Economic, Students with Disabilities) one identifies with, however the research clearly indicates that belonging to any subgroup has an impact on the acquisition of literacy skills (Chatterji, 2006; Entwisle & Alexander, 1993; Hedges & Nowell, 1998; Jenks & Philips, 1998)

Socioeconomic Status

Historically, students who qualify for the government free/reduced lunch program and those who come from an underrepresented population have performed poorly on achievement test measures when compared to their white, non-low SES classmates (Hedges & Nowell, 1998; Jenks & Philips, 1998). Schools are faced with the reality of then providing every child regardless of their prior knowledge and background experiences, a free and appropriate education.

In examining students who qualify as low SES, Barton and Coley (2009) examined the performance of Black and Hispanic students against their white classmates. They discovered sixteen factors that were predictors of academic success. Although a few of their factors appeared to not be discriminatory in nature such as birth weight, and lead poisoning, they shared that Black and Hispanic children are nearly twice as likely to live in households identified as below the poverty line (2009). These researchers also found differences in the exposure that Black and Hispanic children have in regard to having someone reading to them at home as only 50% were read to daily compared to 68% of their white counterparts.

The results substantiate the results of other studies, which have found that children from lower SES homes begin school with weaker language skills than those from higher SES households (Chatterji, 2006; Entwisle & Alexander, 1993). Along these same lines Neuman and Celano (2001) found that those in high-income neighborhoods had roughly 10 times greater access to reading materials than those in low-income neighborhoods. Without quality literacy materials available free of charge, those who live in low-income areas may not have the necessary resources to provide rich, high quality reading materials for their children. Regardless of race, Chatterji (2006) found that children from high poverty households enter school approximately 1 standard deviation behind children from middle or upper income households in regard to their early reading skills.

Underrepresented Populations

It is important to note, however, that race does influence the results of many studies focusing on the early literacy skills of young children (Chatterji, 2006; Jeynes, 2008; Stiefel, Schwartz & Chellman, 2007). These researchers have all found that children from Hispanic and African American cultural heritage tend to enter school behind their white peers. It has also been found that there is variance in how children from these underrepresented backgrounds are educated which concurs with the work of Chatterji (2006), as she explains the importance of having a highly qualified teacher (Kim, Zabel, Stiefel & Schwartz, 2006).

Test scores across the nation continue to reflect these factors (Chatterji, 2006; Kim, Zabel, Stiefel & Schwartz, 2006; Stiefel, Schwartz & Chellman, 2007). It was found that in the state of New York, schools that were identified as segregated tended to perform more poorly than those that were integrated. Ironically, this gap does not exist when examining gender, only race. The findings of Chatterji are also worthy of attention as it was established that Blacks were

less likely than all other subgroups to be proficient in more complex reading skills such as making inferences, being able to derive meaning from text and interpret a reading selection beyond the facts presented (2006). Thus it can be concluded that students who qualify for a subgroup are at risk of not attaining comparable skills to their white counterparts.

Students with Individualized Education Plans

Another student group that schools must monitor the progress of is students who qualify for an Individualized Education Plan (IEP). An IEP is a legal document which specifies specific academic/behavioral goals and objectives for children who have been identified as having significant deficiencies in these areas. An IEP also provides information as to accommodations and specifies the services needed to ensure the student will progress toward the agreed upon goals. The document is good for one year and either party (school or parent) can request a case review at any time during which the document can be amended. In order to qualify for an IEP a student must first receive a multifactored evaluation (MFE). An MFE is comprised of the results of standardized cognitive and achievement testing as well as antidotal information that has been shared by a member of the MFE team. This team typically consists of parent(s), teacher(s), specialist(s), school psychologist and a school administrator. At an MFE meeting the team considers all the data and information that has been collected, and then determines whether or not a student qualifies as having a disability. For those who do qualify this process is then completed again every three years. For those who do not qualify, the results can be revisited and are considered relevant for up to one year. Given the rules and regulations that have been enacted at both the federal and state levels, the identification of students who may qualify as having a disability through the MFE has become a prohibitive process (Operating Standards for Ohio Education Agencies Serving Children with Disabilities, 2008). As the number of students

who qualify for an IEP decreases, the staffing levels to provide the assurances contained in these plans have also become minimal. This has led many schools to limit their service delivery options and often returning to a model that promotes providing literacy instruction for students qualifying for an IEP in a separate, isolated environment. Compounding this complex issue is the mandate to make AYP each school year.

Similar to underrepresented populations, students with disabilities also experience a strong score discrepancy between themselves, subgroups, and their non-disabled counterparts (http://ilrc.ode.state.oh.us/PublicDW/asp/Main.aspx). For years, many educators and parents alike have argued that students with disabilities should be educated in classrooms with their typical peers, a practice known as inclusion. In a study conducted by Gandhi (2007), it was found that those educated in the regular education classroom did as well as those educated away from their peers in a separate setting. Unfortunately given limited financial resources research such as this is often difficult is not impossible for educational leaders to embrace and implement.

Leadership Implications

Schools, like society continue to become more complex, demanding more sophisticated leadership that can handle rapid change an innovation (Fullan, 2001). Today's school leaders must be able to engage others in a data based decision making process. Leaders must also be willing to be held accountable for results that are a manifestation of the decisions made. Hence leadership continues to play a vital role in determining the success of our children (Marzano, 2003).

Data Driven Decision Making

The use of data in the decision making process has become an everyday practice for school administrators. However, it is important to note that data alone is not enough. While data represents facts and figures, it does not necessarily lead to making informed decisions. Only when data is put into context, and reflected upon, often through the use of comparisons and trends does it provide information (Dufour, Dufour, Eaker & Many, 2006). This is most likely why many educational institutions would describe themselves as data rich and information poor (McNulty, 2009). Districts collect data on a frequent, ongoing basis; however, what happens to the data once it is collected? This is the challenge leadership must embrace. Classroom educators need to be provided with district wide professional development opportunities whereby they can learn about effective use of data to improve student instruction which will optimally lead to increased student performance. Educators also need to be provided with opportunities to collaborate and share useful data practices with one another. It is the responsibility of leadership to ensure that this time is provided and structured in a manner that will be productive for classroom educators who have the greatest impact upon the amount of academic growth children experience (Marzano, 2003). Additionally, parents need to be offered the opportunity to participate in learning opportunities and engaged in learning how to interpret their children's data. Without actively involving this group of stakeholders in the data analysis process, schools will not be able to maximize the academic performance of their children. Schools need to offer opportunities for parents to learn about the various tools available to them to monitor and track their students' progress and be comfortable asking questions about the education their child is receiving. Such training is part of the OIP process, a process in which many districts in the state of Ohio have engaged over the past two years.

The OIP process provides a framework for this type of professional development with a core group of district stakeholders. Beginning with a core group of district administrators, teachers and community members, OIP facilitators lead the group through a four day process of examining and interpreting data allowing for each individual to share their input and help the participants put the data in context. This data driven process can then be replicated and the building level as well as at grade level/content meetings held in each building. The end goal is that a plan for which the district will be held accountable for implementing will be formulated.

Allocation of Resources

School leaders across the nation make difficult decisions on a daily basis regarding the best use of public funds in order to maximize student achievement while maintaining quality extracurricular programs for all children. While federal and state funding resources continue to decline every year (Daggett, 2009), and the expense associated with educating children such as fuel costs and salary and benefits for staff continue to rise. This financial stress has led many districts to examine every aspect of their organization and leverage each dollar in an effort to sustain academic excellence. School leaders must not only address the financial needs and challenges of today, but prepare staff to be as efficient as possible with the time allotted each day to spend engaged with students. In doing so leaders are challenged to develop quality, job imbedded staff development that will prepare educators for success in their classrooms. One way to maximize the time educators spend engaged in providing high quality educational experiences to children is to eliminate practices that are not effective and aligned to the goals of the school/district. It is the responsibility of the leadership to be able to listen and respond by reviewing data then taking action to eliminate barriers to making the best possible use of every instructional minute. School leadership can do this through a variety of existing avenues

including district and building leadership teams as well as business advisory councils. In doing so, school leaders will allow for classroom educators to get the most out of the time they have available to educate our youth. In doing so, it will ensure students are more academically prepared and ready to participate in state mandated assessments which are at the core of most state accountability systems.

Accountability is a natural extension of data for most school leaders. Since the inception of proficiency tests in the 1990s Ohio's school leaders have been held accountable for the results of students. This accountability is evidenced by the name of the district superintendent and building principal being printed on the front cover of the building and district report cards. In Ohio the accountability system has several key components that while connected, must be taken into consideration individually when making decisions regarding resource allocations.

State Report Card Indicators

In Ohio report cards are issued to both schools and school districts that receive public funding. This past year there were twenty six indicators at the district level. Table 3 shows the twenty four academic tests beginning at the third grade level and two non test related indicators, graduation rate and attendance that are reported on the district and building report cards as applicable depending on the grade levels attending each building (ODE, 2010).

Table 3

2009 – 2010 State Indicators

Grade	Test(s)	% Proficient or Above to earn a Standard
3	Reading and Math	75%
4	Reading and Math	75%
5	Reading, Math and Science	75%
6	Reading and Math	75%
7	Reading and Math	75%
8	Reading, Math and Science	75%
10 - OGT	Reading, Writing, Math, Social Studies and Science	75%
11 – OGT (cumulative)	Reading, Writing, Math, Social Studies and Science	85%
Graduation Rate	N/A	90%
Attendance Rate	N/A	93%

(ODE, 2009)

The number of indicators earned is the starting point for the state rubric which determines the label (excellent with distinction, excellent, effective, continuous improvement, academic watch, academic emergency) that will be place each school and district. After the number of indicators is determined, the performance is then examined.

Performance Index

Performance index is determined by weighting each score category then dividing by the number of scores (ODE, 2009). Currently Ohio uses the information contained in Table 4 when calculating performance index.

Table 4

Performance Index Weighting

Performance Level	Weight
Untested Student	0.0
Below Basic	0.3
Basic	0.6
Proficient	1.0
Accelerated	1.1
Advanced	1.2

(ODE, 2009)

Earning a high performance index (≥100) is important to leaders as this is one way a school/district can earn an excellent rating designation from ODE. Performance index is also an important value as it represents the number of students your building/district has performing at a high level, exceeding the minimum standard of proficient. Another important category school leaders must examine is Adequate Yearly Progress.

Adequate Yearly Progress

Adequate Yearly Progress (AYP) is another factor taking into consideration when determining the rating of each school/district. In examining AYP, Ohio examines the scores of each subgroup of students that participated in the assessments (ODE, 2009). In order for a

subgroup to be calculated, Ohio has set a minimum number of students. The percentage needed to earn the indicator varies based on the past passage rates as well as the length of time each assessment have been in place. Table 5 contains the AYP targets that Ohio set for the 2009 – 2010 school year.

Table 5

AYP Targets

Grade	Reading Target	Math Target
3	77%	68.5%
4	74.6%	73.7%
5	74.6%	59.7%
6	80.6%	64.1%
7	74.9%	57.8%
8	79.0%	58.0%
10 - OGT	77.4%	68.0%

(ODE, 2009)

If a school/district do not make adequate yearly progress for three consecutive years, the school/district can be rated no better than continuous improvement. Hence it is important to ensure that all subgroups are meeting the desired target each year. School leaders are challenged with finding instructional strategies and resources necessary to ensure each subgroup is on target to make the progress necessary and allow their students to pass the reading and math portions of the assessments given.

Value Added

Value added is the last factor taken into consideration when examining the performance levels of schools/districts. Value added is designed to measure the growth each student experiences over the course of a school year. It is calculated by converting the scaled score achieved into a Normal Curve Equivalent Score (NCE), then comparing the NCE from one year to the next. The individual student scores are combined to form school and district value added scores. If the growth index falls within the standard error, a district/school meets value added expectations. If the total exceeds the standard error in a positive manner, then a district/school exceeds value added expectations. However if a district/school exceeds the standard error in a negative manner then it failed to meet expectations. There are consequences associated with this negative growth as well as rewards for exceeding growth (Table 6). Depending on the percentage of indicators earned, the performance index AYP status and value added scores, schools and districts are assigned a rating. Table 6 provides a flow chart as to the determination of school and district ratings (ODE, 2009).

Figure 1. Rating determination table.

2009-2010 RATINGS TABLE

Indicators Met (district numbers in parenthesis)		Performance Index Score		AYP Status*	Preliminary Designation	Did the Preliminary Designation increase or decrease to Continuous Improvement based on AYP Status?**		Amount of Growth Using Value-Added Calculation	Final Designation
94% - 100% (25-26)	or	100 to 120	and	Met or Not Met	Excellent	IF YES STOP HERE No additional change to the designation can occur based on the value added calculation. The preliminary designation of Continuous Improvement becomes the FINAL DESIGNATION	and	Above average growth for at least 2 consecutive years	Excellent with Distinction
								Below average growth for at least three consecutive years	Effective
								Otherwise no effect on rating	Excellent
75% - 93.9% (20-24)	or	90 to 99.9	and	Met or Not Met	Effective		and	Above average growth for at least 2 consecutive years	Excellent
								Below average growth for at least three consecutive years	Continuous Improvement
								Otherwise no effect on rating	Effective
0% - 74.9% (0-19)*	or	0 to 89.9	and	Met	Continuous Improvement	IF NO CONTINUE Value-added MAY affect a designation when it has not been increased or decreased to Continuous Improvement because of the AYP Status	and	Above average growth for at least 2 consecutive years	Effective
								Below average growth for at least three consecutive years	Academic Watch
50% - 74.9% (13-19)	or	80 to 89.9	and	Not Met				Otherwise no effect on rating	Continuous Improvement
							and	Above average growth for at least 2 consecutive years	Continuous Improvement
31% - 49.9% (9-12)	or	70 to 79.9	and	Not Met	Academic Watch			Below average growth for at least three consecutive years	Academic Emergency
								Otherwise no effect on rating	Academic Watch
0% - 30.9% (0-8)	and	0 to 69.9	and	Not Met	Academic Emergency		and	Above average growth for at least 2 consecutive years	Academic Watch
								Otherwise no effect on rating	Academic Emergency

* - Ohio law states that:
1) A school or district that MEETS AYP can be rated **NO LOWER** than Continuous Improvement.
2) A school or district that FAILS TO MEET AYP for three consecutive years AND in the current year fails to meet AYP in more than one subgroup can be rated **NO HIGHER** than Continuous Improvement.

** - Administrative rule states that if a school's or district's preliminary rating is increased or decreased to Continuous Improvement because of its AYP status, value-added will not make additional changes to the rating.

Updated 10/15/2009

As Tables 3 and 5 demonstrate, accountability begins with the Ohio Third Grade Reading Achievement Assessment. Thus it is critical for educational leaders and educators alike to engage in data based decision making prior to the administration of this literacy assessment and ensure students are successful when achievement assessment begins. In doing so it will allow for the evaluation of curriculum as well as teaching methods providing a basis for early intervention to take place ensuring students are able to acquire the skills necessary to be successful. When students are able to demonstrate proficiency on the assessments, schools and districts will be able

to meet or exceed the indictors, performance index, AYP and value added standards set forth by Ohio.

Summary

Literacy instruction is a complicated yet essential, time consuming activity that professional educators engage in each school day. Since the passage of No Child Left Behind, individual states have generated their own assessments to measure proficiency level as well as growth of individual student's literacy skills. Consequently, studies have been performed measuring various aspects of literacy development in addition to the effectiveness of literacy strategies. These studies have also attempted to predict performance on state assessments based on performance on early literacy measurements (Grant, Savage & Carless, 2001). Although certain aspects of each assessment were found to be predictive for some groups of students, samples often lacked diversity or longevity as predictive measures were only tested two years following the administration of the initial identification instrument. Others have examined the relationship between instruments (Good, Simmons & Kame'enui, 2001), finding a strong relationship between oral reading fluency and the Oregon reading achievement assessment. This finding is of interest to many school leaders since oral reading fluency is uncomplicated to administer as well as interpret.

While assessment tools are important, it is also of importance to understand the working knowledge of educators working directly with students, in order for building and district leaders to make informed decisions regarding professional development. As Baily and Drummond (2006) found, although teachers can often identify struggling readers, it is also critical for them to be able to provide effective intervention. For this to take place, educators must have a proficient understanding of literacy development, which according to Baily and Drummond

(2006) is often lacking. Restraints such as time and financial constraints are the leading reasons given in respect to why a lack of quality professional development exists. Educational institutions need to be cognizant of the needs of staff and ensure they understand literacy development and can identify student needs for all students, not just those who are not reaching benchmark standards.

Consequently, further research is needed in several areas to bridge gaps that currently exist. One such gap that will be addressed through this study is the predictive nature of the Developmental Reading Assessment, AIMS Web oral reading fluency probes, Kindergarten Readiness Assessment for Literacy (KRA-L) as well as attainment of state standards in comparison to the Ohio Achievement Assessment. Once the findings are presented, school leaders will be able to better utilize school resources, make instructional decisions, and provide appropriate instruction for all students. School leaders will be able to decipher which assessments are appropriate for use for each child, allowing them to make data based decisions. They will also be able to narrow the focus of staff and concentrate professional development based on the results the assessments being used, once this study clarifies the alignment of each instrument to the third grade OAA. This research will aspire to also provide educational leaders with the information necessary to reduce the number of assessments administered to children, therefore increasing instructional time for all students. In doing so, schools and school districts will be able to better utilize available resources while providing meaningful literacy assessments and instruction for children.

CHAPTER III. METHODOLOGY

This chapter presents a comprehensive summary of the methodology developed and applied to answer the research questions concerning relationships between district diagnostic reading measurement tools and the Ohio Third Grade Reading Achievement Assessment. It explains the research design, study participants, instrumentation used, data collection procedure applied, data analysis and the assumptions and limitations that need to be considered.

Research Design

This study featured a correlational research design and examined the relationships of five district diagnostic reading measurement tools with the Ohio Third Grade Reading Achievement Assessment. The design determined to what degree a predictive relationship exists between the district diagnostic reading measurement tools and the Third Grade Reading Achievement Assessment. All student reading scores and literacy assessment measures were accessed and examined from a combination of a preexisting student data bases as well as student record reviews.

Participants

The participants of this study are students in one school district's third and fourth grade that received a progress report from the study site at the end of second grade and have completed the district diagnostic reading measurement tools: (1) Kindergarten Readiness Assessment-Literacy (KRA-L); (2) second grade AIMS fluency probes; (3) Developmental Reading Assessment (DRA); (4) STAR reading assessment; (5) Second grade standards based report card. Each student had a score on the previously mentioned reading diagnostic instruments, entered into a database or accessible via an online application. The students had either completed the Third Grade Reading Achievement Assessment for the first time in October of 2008 (cohort 1) or

October 2009 (cohort 2). For the purposes of analysis the cohorts were combined, representing two consecutive years of student data.

The participants were students currently or previously enrolled in a suburban public school district that shares a border with one of Northwest Ohio's largest public school systems. The district's average daily membership is 3,690 students (Ohio Department of Education, 2008). Compared to other surrounding districts, the district in this study has a diverse populace as 29.1% of the students belong to an underrepresented population. Black, non-Hispanic students comprise 13.8% of the student population while 8.2 % are multiracial. The remaining minority students are either from an Asian background 4% or Hispanic background 2.8%. According to the ODE 33.3% of the students who attend this district qualify for the Federal Free/Reduced Lunch Program. Also, 16% of the students have been identified as having a disability (ODE, 2008). All students from this district who completed the second semester of second grade and completed the October administration of the Ohio Third Grade Reading Achievement Test in one of their four elementary schools were included in the study.

The diagnostic test scores, report card indicators, and Ohio Third Grade Reading Achievement Assessment scores of students in grade three and four during the 2009 – 2010 school year were utilized in this study (n=341).

Instrumentation

This study utilized the five district diagnostic reading measurement tools designed to provide information to inform instruction and one achievement measure that is administered to measure reading achievement. The diagnostic instruments are the (1) KRA-L; (2) AIMS reading fluency probes second grade administration; (3) DRA literacy assessment; (4) STAR Reading

Assessment from second grade; (5) standards based report card indicators. The achievement measure utilized in this study was the Third Grade Ohio Reading Achievement Assessment.

As discussed earlier, many literacy assessments assess similar essential literacy skills. Table 6 shows the repetitiveness of several reading skills measured by different instruments each school year. The KRA-L assesses multiple pre literacy skills including concepts of print as well as phonological skills and phonemic awareness. The STAR Reader, Report Card Indicators and OAA all contain measures of vocabulary. It should be noted that although they each address vocabulary, each examines this topic in a unique way. The STAR uses a computer based model and asks students multiple choice questions, while the OAA asks a combination of multiple choice and short answer questions.

Table 6

Literacy Skills Assessed by Diagnostic Instruments

Instrument	Literacy Skill(s) Assessed	Timeline of Administration
KRA-L	- oral language - phonological awareness - concepts of print	No sooner than 6 weeks prior to the start of school year; no later than 10/1
AIMS ORF	- fluency with connected text	* Beginning, middle, end of school year
DRA	- reading engagement - oral reading fluency - comprehension	* Fall and Spring of school year
STAR	- context clues - vocabulary knowledge - construct meaning	* Fall and Spring of school year
Report Card Ind.	- strategies to read unknown words - word attack skills - high frequency words - vocabulary - retells story - comprehension	* Four times per school year
OAT	- reading process - reading applications: informational, technical, persuasive text - reading applications: literary text - acquisition of vocabulary	* October and April/May

* Indicates only end of year data will be used for the purpose of this study.

Kindergarten Readiness Assessment – Literacy (KRA-L)

The KRA-L is a diagnostic measurement tool designed to provide feedback to parents and educators alike in understanding a student's readiness to acquire literacy skills. The KRA-L is a standardized instrument that is typically administered by a classroom teacher, counselor or administrator that has taken part in training as outlined by ODE. This assessment is administered one on one and typically takes 10-15 minutes to administer. The skills assessed through the

KRA-L include but are not limited to oral language, phonemic awareness, and concepts of print. Before or shortly after entering school, all kindergarten children are assessed through the use of the KRA-L. Results of this assessment must be reported back to ODE through the use of EMIS, the state database collection system. Students are asked to complete a series of questions such as answering who, what, when, why, and how questions, repeating sentences, identify rhyming words, generating rhyming words, letter names and sounds as well as sounds that are at the beginning of words. The reliability and validity of the KRA-L are aligned with measurement guidelines developed by the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Center for Measurement in Education (NCME) (Hollinger, 2009). This instrument is state-mandated for all first-time kindergarten students attending a public school and results must be reported to ODE. ODE considers this a secure document; hence, no parts of this instrument can be reproduced or used for any other purpose than its original design. This educational measurement tool has been in place for all public schools in the state of Ohio since the 2004 – 2005 academic year. Table 7 presents the score bands and recommendation made by ODE based on the score achieved by each child.

Table 7

KRA-L Student Score Interpretation

Score Band	Recommended Action
0-13	Further assessment needed for intense instruction
14-23	Further assessment needed for targeted instruction
24-29	Further assessment needed for enriched instruction

(ODE, 2009)

AIMS Web Fluency Probes

AIMS Web Fluency Probes are reading passages taken from Edformation's Standard Reading Assessment Passages, which contain rich, connected text. Probes are designed to be a simple, quick measure that provides information around which educators can develop targeted lessons. These probes provide data that is easily understood by both educators and parents and can be monitored in an efficient way, allowing for frequent measures of growth to take place.

AIMS Web Fluency probes are administered three times a year in the district involved in this study. An administration begins with a student having an unnumbered copy of the probe in front of him/her. The test proctor has a numbered prompt in front of them and tells the student when to begin reading. Once the student has begun to read the proctor starts a stopwatch. The proctor then places a slash through any word that is misread and at the end of one minute instructs the student to stop reading. The total number of words is counted, from which the number of errors made is subtracted, creating a total word per minute count (WCPM). Table 8 displays the norms regarding WCPM by grade level according to the fall and spring administration of fluency probes.

Table 8

Words Correct Per Minute Norms Fall and Spring

Grade	Percentile	Fall WCPM	Spring WCPM
1	25	-	28
	50	-	53
	75	-	82
2	25	25	61
	50	51	89
	75	79	117
3	25	44	78
	50	71	107
	75	99	137

(Hasbrouck-Tindal, 2006)

AIMS Web data are reported to a central office administrator who then reports back to teachers and administrators statistical data, which include the mean and median at the classroom, building and district level. Teachers are also provided a line graph documenting individual student scores, including growth for the January and April administrations.

This instrument and subsequent scoring procedures were developed by Pearson Educational Measurement. It is a nationally recognized progress monitoring tool and is one of two commercial tools recommended by the National RTI panel. AIMS web has been the focus of at least 12 published studies, all confirming the validity and reliability of this instrument. Due to its ease of use and quality control features, AIMS web has become the industry standard in regard to measuring oral reading fluency (ORF).

Developmental Reading Assessment

The Developmental Reading Assessment (DRA) enables educators to observe, record, and evaluate changes in students reading performance (Beaver, 2006). This tool has been a district-mandated practice at the study site since the 1997 academic year. It is administered twice a year and results are reported to parents via progress reports.

The administration of this instrument for students involved in this research study begins with selecting a text that is at the independent reading level for the student. Typically, using district and DRA suggested guidelines; a second grade student should be at a level 24 by the end of the school year. This leveling system contains one alphabetical label as it begins with an A then proceeds to a numerical system beginning with level 1. Table 9 provides grade level expectations for the fall and spring administration of the DRA.

Table 9

DRA Expectations

Grade	Fall	Spring
K	A	3 fiction
1	4 fiction	16 non-fiction
2	18 fiction	24 non-fiction
3	28 fiction	34 fiction

(Beaver, 2006)

The assessment administrator then asks the student to talk about one of his or her favorite books. The administrator will also ask the student about his or her reading preference, whether they like to read alone, with a partner or in a group, as well as a rational for their answer. The last question asked is with whom the student reads with at home. The student is told about a few

attributes of the book he or she will be reading. For example, when administering a level 38 reader the assessor would say, "This book is called *Mae Jemison: Shooting for the Stars*. It is a biography about a woman who wanted to be an astronaut. Please read aloud pages 2 through 4" (Beaver, 2006). Following this introduction the assessor readies to begin a timer when the student has begun to read. Once reading begins, the assessor completes a running record and records the results of the orally read pages. If the student is able to read the passage requested within the guidelines provided (oral words per minute) the assessor then begins to ask the student the questions as outlined for each story. If the student does not read the passage within the parameters provided, the assessor regresses back a level of text and begins at the first step of this process with the new text. Assuming the student passes the fluency portion of the test, and he or she has completed answering the orally asked questions, the student is then given an assessment page on which he or she must answer additional comprehension questions through writing his or her own answers. Once the student has completed this assessment measure, it is turned in to the teacher to be analyzed. Each response is analyzed and charted on a DRA continuum, which ends with a total score. This comprehension score is then evaluated to determine if the text that the student engaged with is at his or her independent or instructional level.

Several research studies have been completed including a reliability analysis that demonstrated inter-rater reliability as well as the validity of the DRA (Beaver, 2006). Table 10 highlights the areas of the DRA that are scored, and then provides detailed descriptors to clarify what each factors are taken into consideration for each scoreable area.

Table 10

DRA II. Scorable Areas and Descriptors (Range for Each Descriptor is 1-4)

Reading Engagement	Comprehension
Book Selection	Previewing
Sustained Reading	Retelling – Sequencing
Oral Reading Fluency	Retelling – Character and Detail
Expression	Retelling - Vocabulary
Phrasing	Retelling – Teacher Support
Rate	Reflection
Accuracy	Making Connections

(Beaver, 2006)

STAR Reading Assessment

The STAR Reading Assessment, an online diagnostic instrument produced by Renaissance Learning. Although frequently used in conjunction with Accelerated Reader, the STAR Reading Assessment is designed to measure the reading level of students beginning in first grade and continues through middle school. This web based assessment designed to take approximately ten minutes to administer. According to the National Center for RTI (2008), this instrument has been found to be able to be generalized to a broad population as well as a reliable and valid instrument for minority populations.

Scoring of this instrument is automated and does not require any time by the test administrator. The software calculates a maximum likelihood Rasch ability estimate based on the difficulty of the items that were presented to the student, and the student's right and wrong responses to those items. The ability estimate is then transformed to the equivalent STAR

Reading scaled score. This program offers the opportunity for the user to select the type of score that will be reported (National Center for RTI, 2010).

Before beginning, students are first assigned an instructional grade level, typically the grade level in which they are enrolled. The assessment begins by asking the student to read a short passage then respond to multiple choice questions. Depending on whether the answer given was correct or incorrect the following question will be more complex or uncomplicated. This assessment also utilizes the CLOZE procedure which requires the student to pick the correct word to complete a sentence. Since every seventh word is missing, it assesses comprehension and vocabulary using this technique. This process of question engagement and level of difficulty adjustment continues until the student arrives at their instructional level as determined by their responses. Educators can then access the results which provide a color level which can be used with Accelerated Reader, a sister program as well as a reading level span indicating the results of the assessment. The STAR Reading Assessment is designed to be administered up to three times per academic school year.

According to the National Center for RTI, the STAR Reading Assessment has been found to have reliability greater than .8, thus concluding it is a reliable measure. These same sources also confirmed both the predictive and construct validity to be at least .7, verifying it to be a valid tool as well. Other than statistical information available through the National Center for RTI and Renaissance Learning, there is little evidence to support this product. Consequently, this study will be one of the first to examine this tool as a predictive measure of achievement.

Local Report Card Indicators

During the 1998-1999 school year, select districts in Ohio participated in a study that focused on the implementation of a standards-based report card. In doing so, the district made

the choice to select specific standards from each content area that provided a learning sequence for students as they progress from one grade level to the next. For over ten years, the district of study has continued to use this reporting tool to communicate academic competency to parents. Aligned to the State of Ohio language arts course of study, teachers use this tool as a guide when deciding whether a student is promoted, placed or retained at the end of each school year. The indictors that are currently used on the second grade report card include: (1) Integrates strategies to read unknown words (visual, structural, context, phonics, decoding); (2) Applies word attack skills; (3) Recognizes high frequency words; (4) Identifies vocabulary that is critical to the reading; (5) Retells a story that includes: story elements, characters, setting, problem, sequence of events, and solution in a fiction selection; (6) Identifies topics, main idea, and details of a nonfiction selection; (7) Answers literal, inferential, and evaluative questions to demonstrate comprehension.

Every nine weeks each student receives a report updating parents as to the progress of their student. At the conclusion of the year students earn either an “S” meaning they have satisfactorily mastered the standard or a “U” standing for unsatisfactory, indicating the student has not mastered the standard. If a student earns six or more indicators they are considered to be adequately prepared for the next grade level. If a student earns five or less indicators they are at risk of not being successful in the next grade and encouraged to receive remediation services over the summer. In an effort to control the reliability and validity of these indicators, common grade level assessments and rubrics have been developed and are used consistently across the district. These assessments were developed cooperatively by teachers and administration over a period of two years. Teachers evaluating these measures are provided with professional development time each year to discuss student work samples and ensure evaluation is consistent

across the district. To ensure consistency in evaluating student performance, all staff use common end of year assessments. Table 11 contains the standards second grade teachers are responsible for evaluating each year.

Table 11

Standards Selected to be on a Standards Based Report Card

Standards Based Report Card Indicators	Evaluation Mark
Integrates strategies to read unknown words (visual, structural, context, phonics, decoding)	S,U
Applies word attack skills	S,U
Recognizes high frequency words	S,U
Identifies vocabulary that is critical to the reading	S,U
Retells a story that includes: story elements, characters, setting, problem, sequence of events, and solution in a fiction selection	S,U
Identifies topics, main idea, and details of a nonfiction selection	S,U
Answers literal, inferential, and evaluative questions to demonstrate comprehension	S,U

(Springfield Local Schools, 2008)

Each year, a revision process takes place whereby teachers have the opportunity to submit changes to the report card. This is done in an effort to ensure this tool reflects the skills that are being taught and expected to be mastered by the end of the academic year. The past three years, staff has determined that the current instrument has continued to be accurate and no changes have been made. Consequently, this study will examine whether a relationship exists between the number of indicators a student earns and their ability to earn a proficient score on the Third Grade Achievement Assessment.

Although all teachers have access to the data for their students, the manner in which the data is used varies greatly. Ideally educators will reflect upon the standards met as well as those

students are struggling to progress toward, in turn creating instructional lessons to meet the needs of the students. It would also be beneficial for staff to adjust their teaching strategies and materials based on the educational needs of the students. However the current report card serves more as an informational tool for parents, then a data tool for teachers.

In order to access this data, the researcher will need to review the progress reports of those involved in the study via Progress Book, an electronic grade book program. These data are readily accessible and the researcher will not use student names or any other identifiable data for the purpose of this study.

Ohio Third Grade Reading Achievement Assessment

Since October of 2003 the Ohio Department of Education has administered a Reading Achievement Assessment to all students who participate in third grade reading instruction. This assessment is designed to measure a student's literacy skills specifically in the areas of acquisition of vocabulary, reading process, informational text and literary text. Typically a student receives a score report six to eight weeks following the administration of the assessment. Scores achieved on this measure are then placed into one of the five state categories. Table 12 provides the score needed to attain a given category, the label associated with each score as well as a descriptive explanation as to the skills needed to attain each level.

Table 12

Performance Level Descriptors (score range) Grade 3 Reading Achievement Assessment

Score Category	Score Range	Description of Score
Limited	241 - 384	Students performing at the limited level do not yet have the skills identified at the basic level.
Basic	385 - 399	Students performing at the Basic Level make limited use of reading comprehension strategies, such as inferencing, predicting, comparing and contrasting and summarizing, to build meaning from text. They can usually respond accurately to literal questions but inconsistently answer inferential and evaluative questions. They are able to decode words and can define some unfamiliar words by using context clues in grade-appropriate reading material. A student at the basic level needs teacher support and prompting to comprehend grade-level texts.
Proficient	400 - 440	Students performing at the Proficient Level usually apply reading comprehension strategies to construct meaning. They use their understanding of the elements of literature (e.g., characters, setting and plot) and the author's use of language to develop an accurate understanding of the text. They use text features, such as titles, subtitles and visual aids, to support their comprehension. They often use organizational text features such as chronology, cause and effect and problem/solution to help them organize and recall information. Typically, these students are able to determine the meanings of unfamiliar at- or above-grade level words by using context clues and structural analysis. These students can read and comprehend grade-level texts with little or no teacher support.
Accelerated	441 - 458	Students performing at the Accelerated Level consistently apply comprehension strategies to develop a thorough understanding of what they read. These students demonstrate an ability to use text structures to comprehend and recall what they have read. They can identify and describe various literary elements including plot, theme, character and setting. They respond accurately to inferential and evaluative questions. They consistently determine meanings of unfamiliar at- and above-grade level words by using word structure analysis and context clues. These students can read and comprehend grade-level texts independently and relate what they have read to what they have read and to their own experiences.
Advanced	459 - 530	Students performing at the Advanced Level apply comprehension strategies to develop a thorough and cohesive understanding of what they read. These students demonstrate an

Score Category	Score Range	Description of Score
		ability to use text structures to interpret, evaluate and extend what they read. They consistently respond accurately to questions about what they have read. They can infer and evaluate the ways that authors affect texts. They use their knowledge of word structure and context clues to extend their vocabulary. These students can use critical reasoning to evaluate texts and are able to relate their understanding of textual information to other texts or situations.

(ODE, 2009)

Data Collection Procedures

The assessment data needed to complete this study were housed in hard copy form of a spreadsheet originating at the central office of the participating district. The cooperating district gathered the data from a variety of sources including but not limited to state testing databases, student information systems as well as district-created databases. The spreadsheet generated contained no student identification information, as students are identified by a numerical value ranging from 001 to 341 and received by the researcher after permission was granted by the superintendent. The cooperating district's DRA data as well as report card data were gathered through a review of each student's cumulative file.

The researcher utilized data from 341 students with their individual scores on each of the diagnostic measures as well as the achievement test data and number of grade card indicators. The data collected were used as a basis to provide evidence in answering the research questions provided in the next section.

Research Questions

1. What is the relationship between the score attained on a district-mandated reading measurement tool or progress measure and students' Ohio Third Grade Reading Achievement Assessment score?

2. What is the minimum score needed on each reading measurement tool (DRA, AIMS WEB fluency probes, KRA-L, standards based report card, STAR Reader) to predict passage on the Ohio Third Grade Reading Achievement Assessment?

3. Which reading measurement tool (pass/no pass) best predicts passing the Ohio Third Grade Reading Achievement Assessment?

4. What reading measurement tool (DRA, AIMS WEB fluency probes, standards based report card, STAR Reader) best predicts the passage on the Ohio Third Grade Reading Achievement Assessment for children in each subgroup (African American, Multi-Racial, Hispanic, Low SES, IEP)?

Data Analysis Procedures

The independent variables or predictor variables in this study were the independent scores on each of the following measures: (1) KRA-L; (2) AIMS probes second grade administration; (3) DRA literacy assessment; (4) STAR reading assessment; and (5) indicators earned on second grade standards based report card. The reported scores will be taken from the second grade, spring administration of each instrument with the exception of the KRA-L which is administered when a child enters kindergarten. Each of these independent variables was quantitative. The dependent variable in this study was categorical, representing whether a student earned at least a proficient score on the Ohio Third Grade Reading Achievement Assessment.

This study utilized several methods of data analysis, depending on the research question presented. A summary of the data analysis techniques is included in Table 9. The first research question addressed in this study examined the strength of the relationship between each of the five independent variables and the quantitative scores of the dependent variable, the Ohio Third Grade Reading Achievement Assessment. These relationships were examined through the use of inferential statistics with a correlational analysis.

The first research question examined the strength of the relationships that exist between the score attained on a district mandated reading measurement tool (1) KRA-L; (2) AIMS probes from second grade administration; (3) DRA literacy assessment; (4) STAR reading assessment (5) Report card indicators and students' scores on the Ohio Third Grade Reading Achievement Assessment.

The second research question examined the minimum score needed on each reading measurement tool (1) KRA-L; (2) AIMS probes from second grade administration; (3) DRA literacy assessment; (4) STAR reading assessment (5) Report card indicators, to predict passage on the Ohio Third Grade Reading Achievement Assessment

The third research question examined the predictive relationship of those who passed the diagnostic instruments: (1) KRA-L; (2) AIMS probes from second grade administration; (3) DRA literacy assessment; (4) STAR reading assessment; (5) Report card indicators, and those who were unsuccessful on each, as to whether or not they were successful in earning at least a proficient score on the Ohio Third Grade Reading Achievement Assessment.

The fourth research question examined the predictive ability of each of the five independent variables in regard to their ability to predict passage (proficient, accelerated, advanced) of the categorical dependent variable, the Ohio Third Grade Reading Achievement

Assessment for each subgroup examined: (1) African American; (2) Multi-Racial; (3) Hispanic; (4) Low SES; and (5) IEP.

Table 13

Research Questions, Variables and Data Analysis

Research Question	**Independent Variables**	**Dependent (s)**	**Data Analysis**
1. What is the relationship between the score attained on a district mandated reading measurement tool or progress measure and Ohio Third Grade Reading Achievement Assessment scores?	* 2nd gr. DRA (quant); * AIMSweb (quant); * KRA-L (quant); * STAR (quant); * 2nd gr. Report Card Indicators (quant)	Ohio Third Grade Reading Achievement Assessment (quant)	Pearson Correlation
2. What is the minimal score needed on each reading measurement tool (DRA, AIMS WEB fluency probes, KRA-L, standards based report card, STAR Reader) to predict passage on the Ohio Third Grade Reading Achievement Assessment?	* 2nd gr. DRA (quant); * AIMSweb (quant); * KRA-L (quant); * STAR (quant); * 2nd gr. Report Card Indicators (quant)	Ohio Third Grade Reading Achievement Assessment (quant)	Simple Regression
3. Which reading measurement tool (pass or no pass) best predicts passage on the Ohio Third Grade Reading Achievement Assessment?	2nd gr. DRA (cat); * AIMSweb (cat); * KRA-L (cat); * STAR (cat); * 2nd gr. Report Card Indicators (cat);	Ohio Third Grade Reading Achievement Test (categorical)	Logistic Regression
4. What reading measurement tool	* 2nd gr. DRA (quant); * AIMS web (quant);	Ohio Third Grade Reading Achievement	Logistic Regression

Research Question	Independent Variables	Dependent (s)	Data Analysis
(DRA, AIMS WEB fluency probes, standards based report card, STAR Reader) best predicts the passage on the Ohio Third Grade Reading Achievement Assessment for children in each subgroup (African American, Multi-Racial, Hispanic, Low SES, IEP)?	* KRA-L (quant); * STAR (quant); * 2nd gr. Report Card Indicators (quant);	Assessment (categorical)	

Assumptions

Several assumptions were made when this study was conducted. It was assumed that the instruments involved in this study were administered in a consistent manner by the teaching staff at the study sites. Another assumption was each student put forth their best effort during the administration of each the instruments. Another element not accounted for were interventions students have received prior to the administration of each assessment. Lastly, this study assumes the scores of the participants are representative of the population of the district.

In addition to assumptions, this study also contains limitations. School practitioners will need to use caution in generalizing this study to any student population that is different than the one represented in the study. Also, this study was limited geographically as the sample came from only one district in the state of Ohio and did not contain a state or nationwide sample. The final limitation that needs to be considered is the changing nature of the Ohio Achievement Assessment. It is anticipated that the reading assessment will become a language arts assessment

including additional written responses, therefore, this study needs to be applied with caution as it only examines the third grade reading test data. Also compounding this issue, new English/Language Arts content standards will be implemented across Ohio by 2014 as part of the national Common Core initiative.

CHAPTER IV. RESULTS

This chapter presents the results of the statistical analysis conducted for this study. The purpose of this study was to examine the degree students' performance on five diagnostic reading instruments relate to and predict passage on the Ohio Third Grade Reading Achievement Assessment. The data were analyzed using descriptive statistics, correlations and regressions using the Statistical Package for Social Sciences (SPSS). This chapter was organized by first presenting the descriptive results, which include demographic information about the participants.

Descriptive Results

Demographic Characteristics

The 341 students whose data were used for the purpose of this study were all enrolled in a suburban public school system located in Northwest Ohio. Data were collected from each of the student's individual scores on the (1) Kindergarten Readiness Assessment-Literacy (KRA-L); (2) Second grade AIMS fluency probes; (3) Developmental Reading Assessment (DRA); (4) Second grade district-created standards-based report card; and (5) STAR reading assessment at the second grade level. As shown in Table 14, of the 341 participants, 279 (81.8%) were White while the remaining 62 (18.2%) subjects represented four other ethnic backgrounds (Asian, Black, Hispanic, Multi-Racial).

Table 14

Frequency and Percent of Ethnic Groups

Ethnic Designation	*n*	%
Asian	7	2.1
Black	25	7.3
Hispanic	7	2.1
White	279	81.8
Multi-Racial	23	6.7

Also, 46 (13.5%) had been identified as having a disability, while 295 (86.5%) did not have an identified disability. Additionally, 91 (26.7%) of the participants qualified as low socioeconomic status while 250 (73.3%) were categorized as not qualifying as low socioeconomic status. Table 15 presents a summary of demographic results for the total sample.

Table 15

Demographic Summaries

	Yes		No	
	n	%	*n*	%
White	279	81.8	62	18.2
Disability	46	13.5	295	86.5
Low SES	91	26.7	250	73.3

Summary of Assessment Results

The mean and standard deviation were calculated for each diagnostic reading instrument, related to the participants on each of the instruments and data sources available (see Table 16). Since STAR reading data was available in multiple formats (STAR SS, STAR GE, STAR PR, STAR NCE), each was examined for the purpose of this study. The maximum and minimum scores attained on each instrument are also reported. There was also variation in the number of subjects who participated in each diagnostic reading measure due to the availability of data.

Table 16

Descriptive Statistics on Reading Assessments

	N	*M*	*SD*	Min	Max
STAR SS	269	365.02	141.992	57	736
STAR GE	269	3.25	1.275	0.3	6.6
STAR PR	269	58.83	29.495	1	99
STAR NCE	269	56.70	22.183	1	99
AIMS	329	86.40	38.227	9	216
KRAL	340	21.79	5.614	2	32
Fall 2008 OAA	166	418.86	29.604	332	499
Fall 2009 OAA	163	404.43	23.100	348	460
Fall 2008 & 2009 OAA	330	411.73	27.517	332	499
DRA	333	27.53	6.916	4	50
Report Card Indicators	331	6.06	1.761	0	7

Note. Fall 2008 & 2009 OAA combine the results from Fall 2008 OAA and Fall 2009 OAA into one value.

The number and percentage of subjects who participated in each assessment, including the passage and failure rates were also calculated (see Table 17). STAR PR was selected from the four STAR assessments to be the value used for reporting purposes. This decision was based on practitioner preference after interviewing teachers and parents and considering the most feasible way to communicate the data in an understandable format. In discussing results interpretations with staff, there is an existing understanding of the statistical meaning of percentile ranks, and how to interpret scores as they deviate from the 50^{th} percentile. The same can be said when explaining results to parents. Parents appear to understand percentile ranks after it is explained to them that if you line up 100 children with the first child in line having the weakest skills and the last child in line has the strongest skill, then sharing where their child falls. This method of explanation creates a clear visual for parents and educators alike helping each to understand the meaning of percentile ranks. The diagnostic reading assessments with the lowest passage rates were STAR PR (62.5%) and AIMS (63.8%). Those with the highest passage rates include KRA-L (89.4%) and the second grade DRA (86.8%).

Table 17

Pass/Fail Frequencies and Percent for Reading Assessments

	Total	Pass		Fail	
	n	*n*	%	*n*	%
STAR PR	269	168	62.5	101	37.5
AIMS	329	210	63.8	119	36.2
KRAL	340	304	89.4	36	10.6
Fall 2008 OAA	166	122	73.5	44	26.5
Fall 2009 OAA	163	103	63.2	60	36.8
DRA	333	289	86.8	44	13.2
Report Card Ind.	331	289	84.8	44	13.2

Results by Research Question

Four research questions were asked for this study, with the first addressing the core of the study. In examining the data a statistical significance at the .05 alpha level was applied to all inferential results.

Research Question #1

What is the relationship between the score attained on a district diagnostic reading measurement tool and students' score on the Ohio Third Grade Reading Achievement Assessment?

The relationships between the fall administration of the Ohio Third Grade Reading Achievement Assessment and the five diagnostic reading measurement tools were calculated using Pearson Correlations. As shown in Table 18, each of the reading assessments is

significantly related to the third grade OAA. The STAR SS has the strongest relationship with the third grade reading achievement assessment with an r=.725. The STAR PR, which has a strong practical significance given its ease of use, also showed a strong significant relationship with the third grade reading achievement assessment as r=.722. While not quite as strong, AIMS also had a significant relationship with the OAA (r=.585). The report card indicators were significantly related to the OAA, as well (r=.508). Lastly, the DRA had an r value of .468, while the KRA-L had the weakest relationship with the third grade reading achievement assessment (r=.445).

Table 18

Correlation Coefficients of Reading Assessments with OAA

	r	*p*	*n*
STAR SS	.725	<.0001	265
STAR GE	.713	<.0001	265
STAR PR	.722	<.0001	265
STAR NCE	.710	<.0001	265
AIMS	.585	<.0001	325
KRAL	.445	<.0001	329
DRA	.468	<.0001	327
Report Card Ind.	.508	<.0001	327

Research Question #2

What is the minimum score needed on the five district diagnostic reading measurement tool (DRA, AIMS WEB fluency probes, KRA-L, standards based report card, STAR Reader) to predict passage on the third grade reading achievement assessment?

Regression equations were first generated for each independent variable (AIMS WEB fluency probes, KRA-L, standards based report card, STAR Reader). A passing score of 400 for the Ohio Reading Achievement Assessment was then entered into each regression equation to calculate the minimum score needed on each diagnostic reading measure to equate to passage on the OAA (Table 19).

The STAR PR equated to a 50.11 to earn a passing score on the OAA with a regression equation of y= -257.24 + .77x. AIMS required a score of 76.66 to equate to a passing score given the regression equation of y=-48.67 + .81x. Additional equations as shown in Table 19 were generated for KRA-L, DRA and Report Card Indicators.

The minimum score that signifies a student has passed an assessment according to district guidelines is identified in Table 19. For STAR PR, the district has identified the 50^{th} percentile as the score needed to ensure a student is not in need of remediation. Students scoring below this mark are identified as at risk of not meeting grade level standard and participate in additional assessments to examine whether they qualify for intervention services. The district of study used the work of Hasbrouck and Tindal (2006) to identify 89 words per minute as the benchmark for signifying a passing score on the AIMS assessment. Regarding the KRA-L, the district used guidelines provided by the Ohio Department of Education in setting the benchmark score of 17 for this assessment. Much of the same is also true in reference to the score identified to signify passing on the DRA. According to Beaver (2006), students earning a score of 24 or greater on

the DRA are considered to have reading skills that are at grade level expectation. According to district personnel, the expectation for a child to earn at least seven report card indicators is the outcome of meetings that have been held over a period of more than five years. Staff members came to an agreement that a student would be able to not earn one report card indicator and could still be promoted to the next grade level. They also agreed that if a student earns six or fewer indicators, the child may still go on to the next grade level; however, they are placed rather than promoted, which signifies to the teacher of the next grade level, and the parent that the child may need reading intervention in order to be successful.

In conducting the statistical analysis, students need to earn a score of 400 or greater to be counted as passing the Ohio Third Grade Reading Achievement Assessment. This has been the minimum score at which the Ohio Department of Education considers a student to have proficient reading skills (ODE, 2010).

Table 19

Regression Equations to Calculate Minimum Score for OAA Pass

IV (x)	*n*	Regression Equation	*F*	*p*	Necessary score to pass OAA	Score identified by district as passing
STAR PR	265	Y = -257.24 + .77X	287.12	<.0001	50.11	50.00
AIMS	325	Y = -248.67 + .81X	167.63	<.0001	76.66	89.00
KRA-L	329	Y = -14.32 + .09X	80.80	<.0001	20.91	17.00
DRA	327	Y = -20.57 + .12X	91.12	<.0001	26.23	24.00
Report Card Ind.	327	Y = -7.12 + .03X	113.08	<.0001	5.69	6.00

Research Question #3

Which district diagnostic reading measurement tool (pass/no pass) best predicts passing the third grade reading achievement assessment?

A forward logistic regression was conducted to determine which independent variable(s) (DRA, AIMS, KRA-L, STAR, Report Card Indicators) best predict passage on the third grade Ohio Reading Achievement Assessment. A significant model was generated that includes the STAR PR and AIMS (-2 Log Likelihood = 210.53; $\chi^2(1) = 111.29$, $p<.0001$). Regression coefficients are presented in Table 20. The Wald statistics confirmed that the variables significantly predict OAA passage; however, the odds ratios for these predictors indicate little change in the likelihood of OAA passage.

Table 20

Logistic Regression Results for Total Sample

	B	Wald	*df*	*p*	Odds Ratio
STAR PR	.049	32.737	1	.000	1.050
AIMS	.015	3.809	1	.051	1.015
Constant	-2.917	30.661	1	.000	0.054

Research Question #4

What district diagnostic reading measurement tool (DRA, AIMS WEB fluency probes, standards based report card, STAR Reader) best predicts the passage on the third grade reading achievement assessment for children in each subgroup (African American, Multi-Racial, Hispanic, Low SES, IEP)?

A forward logistic regression was conducted in an effort to reveal which district diagnostic reading measurement tool would best predict passage on the Ohio Third Grade Reading Achievement Assessment for children in each subgroup. Table 21 presents regression coefficients for each subgroup of disability. A significant model was not generated for students with a disability (n=46). However, for those without a disability (n=295), STAR PR and AIMS were two diagnostic assessment tools that best predict passage on the third grade reading OAA. This significant model (-2 Log Likelihood = 174.42; $\chi^2(1) = 91.66$, $p<.0001$) correctly classified 83% of the students with respect to OAA Reading passage.

Table 21

Logistic Regression Results for Disability Groups

	B	Wald	*df*	*p*	Odds Ratio
Disability *n*=46					
	No significant model was generated				
No Disability *n*=295					
STAR PR	.052	28.820	1	.000	1.053
AIMS	.021	6.003	1	.014	1.021
Constant	-3.603	28.052	1	.000	.027

Assessment predictors of the Reading OAA were also examined for SES groups. Results of those who qualified as Low SES were compared against those who did not qualify as Low SES. Table 22 displays the results of the forward logistic regression and reveals that for both groups, Low SES (*n*=91) and Non-Low SES (*n*=250), STAR PR was the assessment which best predicted passage on the OAA. The significant model for the Low SES group (-2 Log Likelihood = 53.30; $\chi^2(1) = 38.84$, $p<.0001$) ; while the model for Non-Low SES was (-2 Log Likelihood = 159.54; $\chi^2(1) = 66.05$, $p<.0001$).

Table 22

Logistic Regression Results for SES Groups

	B	Wald	*df*	*p*	Odds Ratio
Low SES *n*=91					
STAR PR	.077	20.282	1	.000	1.080
Constant	-2.999	.777	1	.000	.050
Non –Low SES *n*=250					
STAR PR	.054	44.041	1	.000	1.055
Constant	-2.015	20.090	1	.000	0.133

Finally, a forward logistic regression was also conducted to examine the comparison of students of color (n=62) to those identified as White (n=279). The statistical analysis results in Table 23 revealed for minority students, STAR PR and the AIMS fluency probes best predicted passage on the OAA Reading; -2 Log Likelihood = 36.92; $\chi^2(1) = 18.70$, $p<.0001$. For White students, STAR PR was found to be the instrument that best predicted passage; -2 Log Likelihood = 171.63; $\chi^2(1) = 94.27$, $p<.0001$.

Table 23

Logistic Regression Results for Ethnic Groups

	B	Wald	*df*	*p*	Odds Ratio
Non-White n=62					
STAR PR	0.033	3.271	1	.070	1.034
AIMS	0.043	4.086	1	.043	1.044
Constant	-4.019	1.614	1	.013	0.018
White n=279					
STAR PR	0.062	56.076	1	.000	1.063
Constant	-2.472	30.852	1	.000	0.084

Summary

This study was conducted using the data from 341 students enrolled in a suburban school district located in Northwest Ohio. The researcher utilized several data sources including existing databases that were both local and web based. In order to be included as a participant in the study, each subject must have completed four of five of the diagnostic assessments [(1) Kindergarten Readiness Assessment- Literacy (KRA-L); (2) second grade AIMS fluency probes; (3) Developmental Reading Assessment (DRA); (4) STAR reading assessment; (5) Second grade standards based report card] examined. Table 24 summarizes the statistical results from each research question. Pearson Correlation coefficients indicate that the STAR SS and STAR PR to have the strongest relationship with the third grade Reading Achievement Assessment. Multiple regression results indicate the STAR PR and the AIMS best predict passage on the third grade Reading Achievement Assessment for the total sample. The final research question examined

the predictive nature of each tool for each subgroup based upon disability, SES and race. The STAR PR was found to be the most accurate for students without disabilities, both SES groups, as well as White students. AIMS was found to be additional predictor for minority students and students without disabilities.

Table 24

Summary of Inferential Results by Research Question

Research Question	Results	
1. What is the relationship between the score attained on a district mandated reading measurement tool or progress measure and students' third grade Reading Achievement Assessment scores?	Reading Tool	*r*
	STAR SS	.725
	STAR GE	.713
	STAR PR	.722
	STAR NCE	.710
	AIMS	.585
	KRA-L	.445
	DRA	.445
	Report Card Indicators	.508
2. What is the minimum score needed on each reading measurement tool (DRA, AIMS WEB fluency probes, KRA-L, standards based report card, STAR Reader) to predict passage on the third grade reading achievement test?	Measurement tool	Score needed on reading measurement to pass OAA
	STAR PR	50.11
	AIMS	76.66
	KRA-L	20.91
	DRA	26.23

Research Question	Results	
	Report Card Indicators	5.69
3. Which reading measurement tool (pass or no pass) best predict passage on the third grade reading achievement test?	Significant predictive model: • STAR PR • AIMS	
4. What reading measurement tool (DRA, AIMS WEB fluency probes, standards based report card, STAR Reader) best predicts the passage on the third grade reading achievement test for children in each subgroup (African American, Multi-Racial, Hispanic, Low SES, IEP)?	Disability predictive model: • No predictive model generated No Disability predictive model: • STAR PR • AIMS Low SES predictive model: • STAR PR Non – low SES predictive model • STAR PR Minority predictive model: • STAR PR • AIMS White predictive model: • STAR PR	

CHAPTER V. DISCUSSION AND RECOMMENDATIONS

Chapter five presents a brief review of the importance of this study as well as the significance of the study to the greater body of empirical knowledge. Organized by research themes, the chapter will also present the research and practical implications of the research conducted, with a critical perspective focusing on data interpretation.

Review of the Study

School districts across the state of Ohio administer multiple assessment instruments every year in an effort to identify students in need of intervention. Grounded in the theoretical framework of Teale and Sulzby (1986), this study was based upon the premise that early identification of children in need of remediation will result in more children developing the five core roots essential to becoming literate. Beginning with the concept of print awareness followed by children understanding the flow of print from left to right, the first two roots are critical to the development process in learning to read (Teale & Sulzby, 1986). When a student reaches root three, the focus begins to be on written language as root four then logically connects the written language to oral language skill development. At the pinnacle of this framework, students begin to apply metacognitive skills and are able to analyze and explain language (Teale & Sulzby, 1986). This theoretical framework, as defined by Teale and Sulzby (1986), is important to this study as students must apply concepts from root five in order to be successful on high stakes assessments such as the Ohio Third Grade Reading Achievement Assessment.

These assessments aim to examine students' progress toward grade level standards as defined by the Ohio Department of Education (ODE). However, in addition to district-selected diagnostic assessments there is also a series of state mandated assessments beginning with the Kindergarten Readiness Assessment of Literacy (KRA-L) in kindergarten. Similar to many

other states, Ohio mandates a series of scheduled assessments on a yearly basis. In third grade, a high stakes assessment is used as a key determinant as to whether a child is permitted to move to the next grade level. Such tests have forced districts to examine not only their education practices, but also the reliability of each district-selected reading diagnostic assessment in predicting passage on the Ohio Third Grade Reading Achievement Assessment. In doing so, districts are able to identify and intervene with those students who are at risk of failing, therefore maximizing the passage ratio. This is important not only to ensure student learning, but also to ensure that a desirable school and district rating on the State Report Card is achieved.

The concept and effectiveness of early intervention is well documented (NAEP, 2009; Sloat, Beswick & Willms, 2007). For Ohio school districts, identifying children who are at risk at an early age, effective intervention strategies can be deployed, increasing the likelihood that a child will be successful on the Ohio Third Grade Reading Achievement Assessment and other assessments. This study aimed to clarify which of the studied diagnostic measures best predicts passage on the on the OAA, providing evidence to support the use of instrumentation that can be used for early identification.

The present correlational study, conducted in an Ohio school district, was designed to examine the relationship of students' reading performance on five different district-mandated diagnostic reading assessment measures and the Ohio Third Grade Reading Achievement Assessment. The five diagnostic instruments that were examined include (1) Kindergarten Readiness Assessment- Literacy (KRA-L); (2) Second grade AIMS fluency probes; (3) Developmental Reading Assessment (DRA); (4) Second grade district created standards-based report card; and (5) STAR reading assessment. A primary factor as to why this study was conducted is to provide assistance for leaders in choosing the diagnostic instrument that best

predicts passage on the Ohio Third Grade Reading Achievement Assessment. Ultimately, the goal is to identify children at risk of not passing this high stakes test so that early intervention can take place. Subgroup variables of SES, IEPs, and race were also examined to determine the relationship between these variables and the Ohio Third Grade Reading Achievement Assessment. Four research questions were synthesized to form the foundation of the information gathered as part of the study.

The participants of the study were 340 students who were enrolled in a suburban public school district located in Northwest Ohio. The students attended four different elementary schools within this district. Each participant had been enrolled continuously in the district from kindergarten through the October administration of the Ohio Third Grade Reading Achievement Assessment.

Discussion

The conclusions of this research study are organized by the four research themes that were extracted from the research questions proposed in this study. Following each theme is a discussion of the research results.

Relationships Between Diagnostic Assessments and the OAA

This study revealed strong relationships between each of the diagnostic assessment measures (Kindergarten Readiness Assessment-Literacy, Second grade AIMS fluency probes, Developmental Reading Assessment, Second grade district-created standards-based report card indicators, STAR reading assessment) and the Ohio Third Grade Reading Achievement Assessment. It should be noted that the instruments that have the strongest correlations with the Ohio Third Grade Reading Achievement Assessment are also instruments that have limited subjectivity due to the manner in which the protocols are administered. The STAR reading

assessment, which had correlations ranging from .725 to .710, depending on the subscale (STAR PR, STAR SS, STAR GE, STAR NCE), is a computer based assessment. The content of this assessment is based on a combination of multiple choice questions students must answer following a passage and the use of the CLOZE procedure by which students must pick the word which best fits each sentence. In taking this assessment, students come into a lab setting, participate in the assessment then teachers are able to access the results via a data base that has been formed. As evidenced in several studies (Topping, Samuels & Paul, D.P., 2007; Topping, Samuels & Paul, T., 2007), computer based assessment, while it may remove the humanistic aspect of assessment, was shown to be highly correlated with nationally published standardized tests.

These research findings indicate that computer based assessments, in this case the STAR, deserve to be valued in the education decision making process, as much and possibly more so than teacher administered assessments such as the DRA. While the research on computer based assessment has been in existence for many years (Knezek & Christensen, 2007; Ross, Nunnery & Goldfeder, 2004; Topping, Samuels & Paul, D.P., 2007; Topping, Samuels & Paul, T., 2006), the general feeling of the educational community (pre k – 12) has been that while beneficial, these assessments are of less value than teacher administered assessments (Clariana & Wallace, 2002; Czubaj, 2004). This sentiment is supported by the indisputable factors and interactions that occur during teacher administered testing that computer based assessment cannot account for. Factors such as reading with expression, the ability of the student to attend to a task as well as the student's ability to understand spoken language cannot be replicated by any one piece of technology at this time. However, when the goal is to identify students in need of intervention as

well as the instruments that are most effective and efficient in doing so, computer based assessment cannot be ignored.

School leaders must also consider the amount of time and resources each assessment requires. In another Midwestern state it was estimated that testing costs, including teacher and administrative time away from normal duties, substitutes and secretarial overtime, the costs of test administration could be estimated as high as seventeen million dollars (Zellmer, Frontier, & Pheifer, 2006). The DRA takes approximately 30 minutes per student to administer to second grade students. In the district of study, the average teacher salary is slightly above $58,000. Assuming there are 25 students in a classroom and the DRA is administered twice a year, about four days per year are spent administering this assessment. The cost of this, assuming the teacher is making an average salary, is more than $625.00 per year in personnel costs alone for each classroom in the district. Amplified across the district, if given twice a year, in each building, given there are three or four sections in each building, the administration costs rise above $20,000 per year. The STAR, which takes 30 minutes to administer, but is a whole class assessment, costs less than $50.00 for personnel to administer. While there is an annual subscription cost associated with this computer based assessment, it is still a fraction of the resources consumed in administering the DRA. This also does not account for the loss of quality instruction while assessments are being administered.

The amount of instructional time lost while assessment is taking place needs to be given consideration when decisions regarding assessment are being made (Zellmer, Frontier & Pheifer, 2006). While it is difficult to attach a dollar value to the instructional time a teacher must sacrifice to administer the DRA to his/her class, it should not be ignored. Over the course of a child's academic career, from the time they enter kindergarten to the end of second grade,

students have been deprived of the equivalent of a week of school due to the administration of the DRA, not taking into account the additional time spent on conducting other similar assessments. This underscores the importance of reviewing the purpose of each assessment a district administers and ensuring the time lost conducting assessment is worth the information gained from conducting the assessment.

Another instrument that had a relationship with the Ohio Third Grade Reading Achievement Assessment was the AIMS fluency probes. These probes, which were highly correlated (r=.585) with the OAA, are also less subjective assessments as the results are reported in the form of a word count. During this assessment children read a given page of print for one minute at the end of which the teacher calculates the amount of words read correctly and records a score. Once again, there is very little subjectivity as the assessment is a word count and does not take into effect for scoring purposes other elements of oral reading such as appropriate voice. Regarding the correlation, this study confirms the research of past research as oral reading fluency has been found to be an accurate predictor of success on reading achievement test scores (e.g., Buck & Torgesen, 2003; Crawford, Tindal & Steiber, 2001; Hintze & Silberglitt, 2005; McGlinchey & Hixson, 2004; Shapiro, Keller, Lutz, Santoro, & Hintze, 2006).

Other than the KRA-L, which is administered in kindergarten, the remaining two reading diagnostics (DRA, Report Card Indicators) are instruments which are designed to be objective; however, each permits an element of subjectivity as results are reported based on teacher interpretation. With respect to the DRA, teachers must listen to students read, then following the oral reading, ask students to respond verbally and through written responses. Teachers must then evaluate the answers given using the rubric which is provided in the teacher's manual. Doing so allows for the human elements of education to interact with the instrumentation. To reduce this

subjectivity, based on the professional development schedule of the district, staff members at the district of study shared that calibration meetings are held on a yearly basis in an effort to standardize the scoring of this diagnostic tool (personal communication, March 18, 2009). Based on the statistically significant relationship between the DRA and the performance on the Third Grade Reading OAA (r=.468), it appears that these yearly meetings are successful in terms of assisting staff to ensure the instrument is given in a manner that follows the integrity of the designed assessment administration.

Much of the same can also be said of the use of Report Card Indicators. While common assessments have been devised to help standardize the assessment and Report Card marking process, teachers still have discretion to make professional judgments as to how content is taught in their classroom and whether or not a student is proficient on a given standard (Archbald & Porter, 1994; Malouff, 2008). Hence there is still room for subjective interpretation. In an effort to minimize subjectivity, similar to the DRA, based on the professional development schedule provided by the Lucas County Educational Service Center, the district of study provides opportunities each year for staff from each of the elementary buildings to collaborate and share student work which exemplifies passing work and that which is below grade level standards.

The correlational results have a significant impact upon the manner in which educators should treat each type of assessment. Prior to conducting this study, professional educators in the district of study shared that they had little confidence in the results of the STAR (personal communication, May 13, 2010). Since teacher referral is the most powerful source of information when determining factors such as qualification for special education (Algozzine,Christenson, Ysseldyke, Underwood & Shriner, 1997) the results of this study, particularly those related to the correlations associated with computer based assessment (STAR)

need to be considered as well. Response to Intervention teams need to be aware of these results to ensure data from diagnostic instruments have a prominent place throughout the process of making decisions regarding the education of children. Educators must insist on the use of district diagnostic instruments as the primary source of data and information, using teacher input as a secondary source of data and information. In doing so it will allow for the correct identification of children needing early intervention.

Predicting Scores

In order to provide high quality early intervention it is necessary to know what minimal score is needed on a diagnostic to predict passage on the high stakes Ohio Third Grade Reading Achievement Assessment. For the purpose of this study, it is important for a diagnostic instrument to be able to identify children at risk in order to ensure they receive the intervention necessary to be successful. The results, which generated a regression equation for each diagnostic instrument, revealed several points that should be noted.

First, the STAR PR (percentile rank) reading assessment revealed that in order to pass the Ohio Third Grade Reading Achievement Assessment, a student needed to earn a percentile rank of 50.11, very similar to the percentile rank of 50.00 as identified by district administration that is used to identify children at risk of not being successful (personal communication, April 6, 2010). Consequently this finding confirmed that the district of study was using the results of this assessment in a manner that is consistent with the results from this study. Also, it was revealed that students in the district studied who were reading at 76.66 words per minute were successful on the Ohio Third Grade Reading Achievement Assessment. This finding was important as it has previously been documented with the AIMS probes, students at the 50th percentile should be reading at 89 words per minute (Hasbrouck-Tindal, 2006). While the 50th percentile is often

seen as the cut score used for identification of students who are at risk of not being successful, the results of the present study demonstrate that educators should use caution when interpreting results and using the 50th percentile in the identification of students who are at risk of not passing the OAA based on the results taken from a fluency passage. Doing so will eliminate the over identification of students in need of intervention by using solely an AIMS fluency probe and allow intervention to be targeted on students in need of remediation.

While the DRA results mirrored national (Beaver, 2006) and district expectations (Springfield Local Schools, 2008) and the results related to report card indicators were in line with district expectations, the results regarding the KRA-L necessitate further examination. According to the Ohio Department of Education (2009), students scoring in the range of 14-23 on the KRA-L should receive grade level targeted instruction. Meanwhile the results of this study indicate that a score of 21 is needed to predict a pass on the OAA. The regression equation generated could be used in determining whether a student is at risk when examining the predicted OAA score. Once again doing so will allow educators to deploy early intervention with the goal of being able to remediate the areas of weakness to ensure success on the OAA.

Predicting Passage

While many districts administer multiple diagnostic assessments, it is important to know which instrument(s) best predict passage on the high stakes assessment given in third grade. In examining which diagnostic reading instrument(s) (DRA, AIMS, KRA-L, STAR, Report Card Indicators) most accurately predicts passage on the Ohio Third Grade Reading Achievement Assessment, the model generated clearly indicated that the combination of the STAR PR and AIMS should be administered if the data sought involves predicting passage on the Ohio Third Grade Reading Achievement Assessment. Once again, these two instruments, which are the

least subjective in nature, were found to be the strongest for the purpose of predicting passage on the Ohio third grade Reading Achievement Assessment. It should also be noted that these assessments consume the least amount of both economic resources as well as human capital to administer. This further amplifies the significance of these findings. These results also reveal that the more stringent the requirement for passage, then more accurate the instrument is in predicting passage on the OAA.

The instruments with the highest rate of failure (STAR, AIMS) proved to be the most predictive of passage on the Ohio Third Grade Reading Achievement Assessment. STAR PR and AIMS, which had failure rates above 35%, were the most predictive while Report Card indicators and DRA which had failure rates of just above 13% were found to be less predictive. Continuing this trend, the KRA-L which was found to have a failure rate of just over 10% was the least predictive, although the fact that this diagnostic assessment is administered in kindergarten, almost three years prior to the other instruments in the study, should be taken into consideration. In addition, this assessment is state mandated as every kindergarten student enrolled in a public school must participate in this assessment no more than six weeks prior to the start of school and no later than October 1st.

Subgroup Performance

This study also examined the predictive ability of each diagnostic tool (DRA, AIMS, KRA-L, STAR, and Report Card Indicators) in relation to demographic subgroups and the Ohio Third Grade Reading Achievement Assessment. In doing so, statistically significant models were generated for every subgroup with the exception of students with disabilities.

For those without a disability, STAR PR and AIMS were once again strong predictors of passage on the OAA. Combining these two tools, 83% of the cases in this study could be

classified. Since this combination of tools utilizes one computer based assessment and one teacher-administered assessment, it is important to note the high predictive rate that results.

STAR PR was also the strongest in predicting success on the OAA for those identified as low SES. While accurately classifying 85.3% of the cases for low SES and 79.6% for Non-Low SES, STAR PR demonstrated its ability to predict passage for both subgroups. These results indicate that, while a strong predictor for both Low SES and Non Low SES, the STAR PR was stronger for with those who qualified as Low SES than for those who did not.

Regarding subgroup performance in reference to race, due to the number of students categorized in each subgroup, students were classified into two groups for this study, Minority and White. Those who qualified as Minority were part of one of the following subgroups: (1) Black, non-Hispanic; (2) Multiracial; (3) Asian; and (4) Hispanic. For these students, STAR PR and AIMS were found to be the strongest in predicting passage on the OAA. For students who were classified as White, the STAR PR was found to the most predictive instrument.

Schools need to continue to examine scores of subgroups in an effort to help identify core instructional issues as well as identify areas which professional development focusing on areas of diversity may be beneficial (Gunning, 2006; Marzano, 2003). Also, while individual buildings may not have the minimum number of tested students (n=30) needed to qualify as a subgroup (ODE, 2010), often when school populations are combined at either the middle school or high school level, the number of students exceeds 30, thus forming an AYP subgroup that is then counted on the state report card (ODE, 2010). Hence it is important at an early stage for schools to begin monitoring the data of all students, including subgroups and ensure these students are making adequate progress toward grade level standards and are on track to pass state assessments.

The topic of subgroup performance is also a well researched field (Gunning, 2006; Hedges & Nowell, 1998; Jenks & Philips, 1998; Marzano, 2006), as historically minority subgroups have been outperformed by their White classmates. This has led subgroup data being incorporated into the Ohio Improvement Process for many urban and suburban districts (OLAC Leadership Development Framework, 2009). Consequently it is the responsibility of the district to ensure proper diagnostic measures are in place to identify students in need of intervention. These data, which must be input into the building and district's decision framework, are a basis for determining the success or failure of a schools and districts to meet their annual goals. With this accountability measure in place, in addition to adequate yearly progress (AYP), which is reported on the state report card, school leaders need to continue to monitor the progress of subgroups to ensure their goals, are met.

Recommendations for Practice

This study examined four key research questions focusing on the predictive abilities of five reading diagnostic instruments as they relate to the Ohio Third Grade Reading Achievement Assessment. While the focus is not on classroom instruction, the importance of high quality classroom instruction cannot be overstated. Effective classroom practices should be at the core of the school experience for every child (Bitter, O'Day, Gubbins, Socias, 2009; Marzano, 2006). Effective classroom strategies need to be employed with students who are not earning grade-level benchmarks and have been identified as in need of intervention.

A recommendation for practice is that schools should develop their own norms and distribution tables based on their scores and the relationships that exist between their reading diagnostic instruments and the Ohio Third Grade Reading Achievement Assessment. While commercial test providers and the Ohio Department of Education will continue to make

recommendations, they often stem from limited body of research and rarely does the student population studied reflect that of an individual district. Districts need to take an in-depth look at their own assessment practices, ensure the practices are meeting their respective desired outcomes and eliminate assessments that are ineffective. Doing so will allow for educators to focus on providing a high quality education for their students. Without local norms, a district will not be able to accurately identify children in need of early intervention.

School leaders also need to engage instructional staff in discussions focusing on the value and use of computer-based assessment and instruction as well as to plan how schools will react when some students don't learn (DuFour, Dufour, Eaker & Many, 2006). It is important for these conversations to take place to allow leaders to understand the mindset of their staff as well as gauge the need for professional development. In doing so, the value of hearing a child read aloud needs to be a central point of focus. While information can be gathered from computer-based assessment, educators still value listening to a child read and being able to conference with children as well as their parents to discuss the instructional needs of their children. In addition, most computer based assessments are unable to account for the humanistic factors that can affect a child's performance. Accommodations for children on an IEP or the ability to sense that a child may not be tracking print correctly are important, yet also unaccounted for when considering computer based assessment. These are all points that should be included as part of this guided conversation as school leaders make decisions regarding computer-based assessment.

In applying the results to daily practice, this study should be taken into consideration when considering instruments to use as part of the response to intervention process. The first step of this process is to identify an instrument that will be considered the district's universal screener. The purpose of the universal screener is to have all children participate in the

administration of the same instrument in an effort to identify the bottom 20% of children who should be targeted for intervention (National Center on Response to Intervention, 2010). In choosing an instrument for this purpose, the diagnostic instrument needs to be accurate, generalizable, reliable, valid and have data based on diverse populations (National Center on Response to Intervention, 2010). From an administration standpoint it must contain benchmarks and norms, and be efficient in the administration and scoring time.

Based on the results of this study, the STAR reader meets the criteria outlined. Following administration to students, using an administrative login, a user can rank order students by grade level which would be beneficial in determining the bottom 20% to target for intervention. Each year The National Center on RtI has a technical committee that meets and reviews research that has been conducted on screening tools. The Center defines screening as follows: Screening involves brief assessments that are valid, reliable, and evidence-based. The assessments are conducted with all students or targeted groups of students to identify students who are at risk of academic failure and, therefore, likely to need additional or alternative forms of instruction to supplement the conventional general education approach (National Center on RtI, 2010). The National Center on RtI found convincing evidence that STAR Reader is both a reliable and valid instrument for screening students (2010). Schools searching for an RtI universal screener should give serious consideration as to the use of the STAR as a screening assessment.

The use of AIMS fluency probes is also a practice that should be considered by school districts seeking a secondary screening measure. Given its one minute administration time, this instrument provides educators with instant data that then be used for progress monitoring. AIMS

probes have been found to be reliable and valid screening tools by the technical committee formed by National Center on RtI (2010).

School leaders also need to consider which diagnostic reading instruments should be used to gain the information needed in order to ensure their districts move forward as part of the Ohio Improvement Process (OIP). This process mandates districts track and report their data through an Implementation Management Module (IMM). In doing so, districts will either meet of not meet their goals based on the results reported. Consequently it is important that districts select tools which measure their desired outcomes. For most districts their desired outcomes focus on attaining the state indicators which include reading tests at grades 3 through 8 and the Ohio Graduation Test. This research should be applied when leaders are gathering information and working through the reading diagnostic instrument selection process. Districts need to consider which instruments best predict passage and provide them the information needed to identify students at risk of failure so that early intervention can take place.

While the results of this study are both statistically powerful and the correlations between the assessments and the Ohio Third Grade Achievement Assessment were high, leaders need to ensure parents and classroom educators that although their child's score puts them at risk of not passing the OAA; it does not predetermine their score. The results of this study should provide general guidance in the decision making process and help provide information that will be beneficial in identifying children in need of early literacy intervention.

Future Research

While this study was able to address its four main research questions, still much work remains to be done. Future studies need to be conducted to replicate this work as districts need to examine which reading diagnostic instruments will assist their educators in correctly

identifying students in need of early literacy remediation. It is also of importance for future research to examine best practices as they relate to the use of computerized assessment.

It would be beneficial for the educational community to continue to evaluate the effectiveness and efficacy of computer based assessment in an effort to form a greater body of research to assist in the decision making process. Additional research needs to be conducted investigating the use of teacher administered assessments versus computer based assessment to add to this body of knowledge and assist educators in making informed decisions regarding assessment practices. Lastly, it would be beneficial for additional studies to be conducted investigating the reasons for any variation that exists concerning teacher administered assessments. These studies could explore the findings of Graney (2008), in that teacher perception is not the most reliable source of data in regard to student achievement. Educators need to learn more about why this phenomenon exists as well as explore possible options which could include professional development is reducing the human element that introduces variation as well as reduced reliability and validity into the assessment picture.

Future studies also need to be conducted to examine the subsections of the KRA-L, to determine the strength of relationship each subsection has to the OAA. This information would allow for the prioritization of reading skills and early identification of students in need of intervention. Fortunately, each subsection of this assessment focuses on a specific skill that may need to be remediated. Progress monitoring tools for kindergarten students could also be identified as part of this work. Following the examination of the subsections, it would be beneficial to also have research conducted in regard to interventions that are deployed and the effectiveness of single interventions as well as different combinations of interventions students

participate in between the KRA-L administration and the Ohio Third Grade Reading Achievement Assessment.

Even as this study accounted for the variables of several different diagnostic measures, there were also variables that could not be accounted for. Although variables such as the amount of time a student is read to at home and the amount of time the student spends reading independently have not been measured for the purpose of this study, student mobility was one variable that data was collected for. The study began with over 700 potential participants and after removing those with incomplete data sets, only 327 of the original data set remained. Consequently the study is based on data taken from students who were continuously enrolled in the district of study from kindergarten through the October administration of the Ohio Third Grade Reading Achievement Assessment. Generalizations cannot be made regarding transient students. The same is also consistent for students on an Individualized Education Plan. The sample size for this population was limited, thus limiting the results that were generated. In studying this population in the future, researchers could develop a model which could be applied to students in this subgroup as well as subcategorizing the data by type of disability.

Another variable which was not accounted for in this study is teacher quality. The district of study did not have qualitative evaluation data on staff members, therefore there was not an objective measure that could be utilized to account for teacher quality. Future studies should explore the relationship that exists between teacher scores on evaluation measures and student scores on achievement assessments. This research could be beneficial in supporting the use of performance data in the evaluation process.

In reviewing the results of this study, school leaders need to consider the role of assessment as it relates to providing teachers the information necessary to make critical decisions

regarding the academic skills of their children. If the goal is to identify students at risk of not passing high stakes tests, leaders should consider the results indicating that the STAR reading assessment has a strong relationship with the Ohio Third Grade Reading Achievement Assessment. While all of the instruments investigated (KRA-L, Report Cards, STAR, DRA, AIMS) had positive relationships with the Ohio Third Grade Reading Achievement Assessment, decisions can now be made regarding the minimal score needed to predict passage on the OAA, because regression equations have been formed, as well as scores generated scores for the population investigated. For those who serve diverse populations, this study also addressed the performance of subgroups in relation to the five assessments. It revealed that for every subgroup, except students with disabilities, STAR reader best predicts passage on the Ohio Third Grade Reading Achievement Assessment. It is suggested that future researchers focus on conducting similar research on other commercial reading diagnostic instruments to allow for the culmination of research to occur and provide educational leaders with the information necessary to make informed decisions regarding assessment.

Conclusion

The goal of this study was to examine the strength of relationships that exist between five reading diagnostic measurement tools (KRA-L, Report Cards, STAR, DRA, AIMS) and the Ohio Third Grade Reading Achievement Assessment. This was accomplished through the collection of data from a district located in a suburban school district in Northwest Ohio. These data were then examined using a series of data analysis techniques. Although due to the transient nature of the district of study, the sample size was limited, the results yielded findings which were powerful and substantial to the field of educational leadership and school improvement.

The results of this study demonstrate the strong relationships that exist between the diagnostic reading instruments studied and the Ohio Third Grade Reading Achievement Assessment. These findings are significant as school districts continue to examine their assessment practices in an effort to ensure they are both efficient and effective. According to the results of this study, the STAR Reading Assessment meets both criteria as it has a strong relationship with the OAA while consuming minimal financial resources and instructional time. This finding adds to the greater body of knowledge as little research exists providing correlations of diagnostic reading assessments to the Ohio Third Grade Reading Achievement Assessment that came about as a result of No Child Left Behind. It is the hope of the researcher that the findings will be shared on a state and national basis, as it provides information school leaders need, to make informed decisions based on current research regarding diagnostic reading assessment instruments.

REFERENCES

Archbald, D. A., & Porter, A. C. (1994). Curriculum control and teachers' perceptions of autonomy and satisfaction. *Educational Evaluation and Policy Analysis, 16*(1), 21-39.

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Urbana-Champaign, IL: University of Illinois, Reading Research and Education Center. 148 pages. 0262011123. Location: Dallas SIL Library 372.4 A215.

Algozzine, B., Christenson, S., & Ysseldyke, J. E. (1982). Probabilities associated with the referral to placement process. *Teacher Education and Special Education, 5*, 19-23.

Anderson, J., & Matthews, R. (1999). Emergent storybook reading revisited. *Journal of Research in Reading, 22,* 293-298.

Bailey, A. L., & Drummond, K. V. (2006). Who Is at Risk and Why? Teachers' Reasons for Concern and Their Understanding and Assessment of Early Literacy. *Educational Assessment, 11,* 149-178.

Bailey, T., Jenkins, D., & Leinbach, T. (2005). *Community college low-income and minority student completion study: Descriptive statistics from the 1992 high school cohort.* New York: Columbia University, Teachers College, Community College Research Center.

Baker, S. K., Smolkowski, K., Katz, R., Fien, H., Seeley, J. R., Kame'enui, E. J., & Beck, C. T. (2008). Reading fluency as a predictor of reading proficiency in low-performing, high-poverty schools. *School Psychology Review, 37*(1), 18-37. Retrieved from http://0-search.ebscohost.com.maurice.bgsu.edu/login.aspx?direct=true&db=a9h&AN=31706982&loginpage=Login.asp&site=ehost-live&scope=site.

Beaver, J. (2008). Developmental Reading Assessment. Pearson Educational Measurement.

Barton, P. E., & Coley, R. J. (2009). Those persistent. *Educational Leadership, 67*(4), 18-23. Retrieved from http://0search.ebscohost.com.maurice.bgsu. edu/login. aspx?direct =true&db =a9h&AN=45463463&loginpage=Login.asp&site=ehost-live&scope=site

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust Vocabulary Instruction.* NY: Guilford Press.

Bishop**,** D. & Adams, C. 1990. A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry, 31,* 1027-1050.

Biemiller, A. (2004). Teaching vocabulary in the primary grades: Vocabulary instruction needed. In J.Baumann & E. Kameenui (Eds.), *Vocabulary instruction: Research to practice* (pp. 28-40). New York: The Guilford Press.

Bitter, C., O'Day, J., Gubbins, J. & Socias, M. (2009). What works to improve student literacy achievement? An Examination of Instructional Practices in a Balanced Literacy Approach. *Journal of Education for Students Placed at Risk, 14,* 17-44.

Buck, J. & Torgesen, J. (2003). *The relationship between performance on a measure of oral reading fluency and performance on the Florida comprehensive assessment test.* (FCRR Technical Report #1) Tallahassee, FL: Florida Center for Reading Research.

Chatterji, M. (2006). Reading achievement gaps, correlates, and moderators of early reading achievement: Evidence from the early childhood longitudinal study (ECLS) kindergarten to first grade sample. *Journal of Educational Psychology, 98*(3), 489-507. doi:10.1037/0012-0663.98.3.489

Clariana R., & Wallace P. (2002). Paper-based versus computer-based assessment: key factors associated with the test mode effect. *British Journal of Educational Technology, 33*(5), 593-602.

Crawford, L., Tindal, G., & Stieber, S. (2001). Using oral reading rate to predict student performance on statewide achievement tests. *Educational Assessment, 7*(4), 303-323.

Crowe, E. C., Connor, C. M., & Petscher, Y. (2009). Examining the core: Relations among reading curricula, poverty, and first through third grade reading achievement. *Journal of School Psychology, 47*(3), 187-214. doi:10.1016/j.jsp.2009.02.002

Czubaj, C. (2004, June 22). Literature review: Reported educator concerns regarding cyberspace curricula. *Education,* 124(4). (ERIC Document Reproduction Service No. EJ705773) Retrieved October 3, 2010, from ERIC database.

Donaldson, M. 1978. *Children's Minds.* Glasgow: William Collins Sons and Co. Ltd.

Draper, L. (1994, March, 31). Ohio Proficiency test challenged: Seniors may get reprieve. *Call and Post,* (Cincinnati).

Entwisle, D. R., & Alexander, K. L. (1993). Entry into school: The beginning school transition and educational stratification in the United States. *Annual Review of Sociology, 19,* 401–423. doi:10.1146/annurev.so.19.080193.002153

Foster, W. A., & Miller, M. (2007). Development of the literacy achievement gap: A longitudinal study of kindergarten through third grade. *Language, Speech, & Hearing Services in Schools, 38*(3), 173-181. Retrieved from http://0search.ebscohost.com.maurice.bgsu.edu/login.aspx?direct=true&db=a9h&AN=25732262&loginpage=Login.asp&site=ehost-live&scope=site

Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal measures of reading comprehension. *Remedial and Special Education, 9*, 20-28.

Gandhi, A. G. (2007). Context matters: Exploring relations between inclusion and reading achievement of students without disabilities. *International Journal of Disability, Development & Education*, *54*(1), 91-112. doi:10.1080/10349120601149797

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

Graney, S. B. (2008). General education teacher judgments of their low-performing students' short-term reading progress. *Psychology in the Schools, 45*(6), 537-549. Retrieved from http://0search.ebscohost.com.maurice.bgsu.edu/login.aspx?direct=true&db=a9h&AN=32625761&loginpage=Login.asp&site=ehost-live&scope=site

Grimm, K. J. (2008). Longitudinal associations between reading and mathematics achievement. *Developmental Neuropsychology, 33*(3), 410-426. doi:10.1080/87565640801982486

Gunning, T. G. (2005). Closing the Literacy Gap. Allyn & Bacon. New York.

Guskey, T. R. (2007). Closing achievement gaps: Revisiting Benjamin S. bloom's "learning for mastery". *Journal of Advanced Academics, 19*(1), 8-31. Retrieved from http://0-search.ebscohost.com.maurice.bgsu.edu/login.aspx?direct=true&db=a9h&AN=30045658&loginpage=Login.asp&site=ehost-live&scope=site

Hart, B., & Risley, R. T. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H. Brookes.

Hedges, L., & Nowell, A. (1998). Black–White test score convergence since 1965. In A. Phillips (Ed.), *The Black–White test score gap* (pp. 149–181). Washington, DC: Brookings Institution.

Hintze, J. M., & Silberglitt, B. (2005). A Longitudinal Examination of the Diagnostic Accuracy and Predictive Validity of R-CBM and High-Stakes Testing. *School Psychology Review, 34*, 372-386.

Hollinger, J. (2009). The relationship between students' reading performance on diagnostic assessments and the third grade reading achievement test in Ohio (Doctoral dissertation). Retrieved from http://www.ohiolink.edu/etd/

Ingle, W. K. (2009). Teacher quality and attrition in a US school district. *Journal of Educational Administration, 47*(5), 557-585.

Jarmulowicz, L., Taran, V., & Hay, S. (2007). *Journal of Speech, Language & Hearing Research, 50*, 1593-1605. doi: 10.1044/1092-4388(2007/107

Jencks, C. & Phillips, M. (1998) The Black-White Test Score Gap. Washington, D.C.: Brookings Institution Press.

Knezek, G. & Christensen, R. (2008). Effect of technology-based programs on first- and second grade reading achievement. *Computers in the Schools, 24*, 23-41.

Kim, D. Y.; Zabel, J. E.; Stiefel, L.; & Schwartz, A. E. (2006). *School efficiency and student subgroups: Is a good school good for everyone?* Portland: Lawrence Erlbaum Associates, Inc.

McGlinchey, M. T., & Hixson, M. D. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review, 33*, 193-203.

McIntyre, Powell, Coots, Jones, Powers, Deeters and Petrosko (2005). Reading Instruction in the NCLB Era: Teachers' Implementation Fidelity of Early Reading Models. *Journal of Educational Research & Policy Studies, 5*(2), 66-102.

Mukherjee, A. *Illiteracy in US: Pot calling the kettle black!*. Retrieved from http://www.merinews.com/article/illiteracy-in-us-pot-calling-the-kettle-black/130727.shtml

McGlinchey, M. T., & Hixson, M. D. (2004). Using curriculum based measurement to predict performance on state assessments in reading. *School Psychology Review, 2,* 193-203.

Molfese, V, Modglin, A, Beswick, J, Neamon, J, Berg, S, Berg, C, and Molnar, A (2006). Letter *"Knowledge, Phonological "*Processing, and Print Knowledge: Skill *"Development in Nonreading Preschool Children*. Journal of Learning Disabilities 39 (4), 296-05.

Nathan, R. G. & Stanovich, K. E. (1991). The causes and consequences of differences in reading fluency. *Theory Into Practice, 30*(3), 176-184.

National Assessment Governing Board. (September, 2008). Reading Framework for the 2009 National Assessment of Educational Progress. Retrieved from http://www.nagb.org/publications/frameworks.htm

Nation, K. & Snowling, M. (2007). Beyond phonological skills: broader language skills contribute to the development of reading. *Journal of Research in Reading, 27,* 342-356. doi:10.1111/j.1467-9817.2004.00238.x

National Center on Response to Intervention. (n.d. Retrieved on September 30, 2010 from http://www.rti4success.org/

Neuman, S.B. & Celano, D. (2001). Access to print in low- and middle-income communities: An ecological study of 4 neighborhoods. *Reading Research Quarterly, 36*, 8-26.

Ohio Department of Education (2010, *Statewide Indicators, Performance by Subject.* Retrieved April 5, 2010 from http://education.ohio.gov/GD/DocumentMangement/DocumentDownload.aspx?DocumentID=85389

Ohio Department of Education (n.d. Retrieved July 30, 2010 from http://www.ode.state.oh.us/reportcardfiles/2009-2010/DIST/048223.pdf

Ohio Department of Education (n.d. Retrieved June 23, 2010 from http://www.education.ohio.gov/GD/Templates/Pages/ODE/ODEDetail.aspx?page=3&TopicRelationID=1366&ContentID=3930&Content=88588

Pinnell, G. S., Lyons, C. A., DeFord, D. E., Bryk, A. S., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high-risk first graders. *Reading Research Quarterly, 29*(1), 9-38.

Purcell-Gates, V. (1996). Stories, Coupons, and the TV Guide: Relationships Between Home Literacy Experiences and Emergent Literacy Knowledge. *Reading Research Quarterly, 31*(4), 406–428. doi: 10.1598/RRQ.31.4.4

Ravitch, D. (2009). Time to Kill No Child Left Behind. *Education Week, 36*, 4-6.

Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology, 47*(6), 427-469. doi:10.1016/j.jsp.2009.07.001

Riedel, B. W. (2007). The Relation Between DIBELS, Reading Comprehension, and Vocabulary in Urban First-Grade Students. *Reading Research Quarterly, 42*(4), 546–567. doi: 10.1598/RRQ.42.4.5

Ross, J.A. (2004). Effects of Running Records Assessment on Early Literacy Achievement, *The Journal of Educational Research, 97,*186-194.

Rouse, H. L., & Fantuzzo, J. W. (2006). Validity of the dynamic indicators for basic early literacy skills as indicators of early literacy skills for urban kindergarten children. *School Psychology Review, 35*(3), 341-355.

Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development, 61*, 1728-1734.

Shethar, A. (1993). "Literacy and 'Empowerment'? A Case Study of Literacy behind Bars." *Anthropology and Education Quarterly, 24*(4), 357-372.

Shapiro, E. S., Keller, M. A., Lutz, J. G., Santoro, L. E., & Hintze, J. M. (2006). Curriculum-based measures and performance on state assessment and standardized tests: Reading and math performance in Pennsylvania. *Journal of Psychoeducational Assessment, 24*(1), 19-35.

Sloat, E. A., Beswick, J. F., & Willms, J. D. (2007). Using early literacy monitoring to prevent reading failure. *Phi Delta Kappan, 88(7),* 523-529.

Snow, C. E., Burns, M .S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Stiefel, L., Schwartz, A. E., & Chellman, C. C. (2007). So many children left behind. *Educational Policy, 21*(3), 527-550. Retrieved from http://0-search.ebscohost.com.maurice.bgsu.edu/login.aspx?direct=true&db=a9h&AN=25386787&loginpage=Login.asp&site=ehost-live&scope=site

Swanborn, M. S. L. & de Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research, 69*(3), 261-285.

Taub, G. E., McGrew, K. S., & Keith, T. Z. (2007). Improvements in interval time tracking and effects on reading achievement. *Psychology in the Schools, 44*(8), 849-863. doi:10.1002/pits.20270

Teale, W., & Sulzby, E. (1986). Emergent literacy: Writing and reading. Norwood, NJ: Ablex Publishing Corporation.

Topping, K. J., Samuels, J., & Paul, T. (2007). Computerized assessment of independent reading: Effects of implementation quality on achievement gain. *School Effectiveness & School Improvement, 18*(2), 191–208.

Topping, K. J., Samuels, J., & Paul, D. P. (2007). Does practice make perfect? Independent reading quality, quantity and student achievement. *Learning & Instruction, 17*, 253-264.

United States Department of Education. (2008). Mapping America's Educational Progress. Retrieved from http://ed.gov/nclb/accountability/results/progress/nation.html

Wang, C., Porfeli, E., & Algozzine, B. (2008). Development of oral reading fluency in young children at risk for failure. *Journal of Education for Students Placed at Risk, 13*(4), 402-425. doi:10.1080/10824660802427702

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69*, 848-872.

Wolf, M. (1999). What time may tell: Towards a new conceptualization of developmental dyslexia. The Norman Geschwind Memorial Lecture. *Annals of Dyslexia, 43*, 1-28.

Yang-Hansen, K. (2008). Ten-year trend in SES effects on reading achievement at school and individual levels: A cross-country comparison. *Educational Research & Evaluation, 14*(6), 521-537. doi:10.1080/13803610802576759

Ysseldyke, J. E., Vanderwood, M. L., & Shriner, J. G. (1997). Changes over the past

decade in special education referral to placement probability. *Diagnostique, 23,* 193-201.

Zellmer, Frontier, & Pheifer (2006). What are NCLB's instructional costs? *Educational Leadership, 64,* 43-46.