

## VITA

Velinda F. Stubbs

### EDUCATION

- 2013          Indiana State University  
Ph.D. in Educational Leadership
- 2009          Indiana State University  
Ed.S. in Educational Administration
- 2000          Indiana State University  
School Administration and Supervision Certification
- 1985          University of Evansville  
M.A. in Elementary Education
- 1979          University of Evansville  
B.A. in Elementary Education

### PROFESSIONAL EXPERIENCE

- 2012          Evansville Vanderburgh School Corporation  
Director of English Language Arts and Literacy K-12
- 2011          Evansville Vanderburgh School Corporation  
Director of Elementary Schools
- 2009          Evansville Vanderburgh School Corporation  
Director of Title I Programs and Schools
- 2007          Evansville Vanderburgh School Corporation  
EVSC Intervention Team
- 2000          Evansville Vanderburgh School Corporation  
Principal
- 1982          Evansville Vanderburgh School Corporation  
Teacher, Instructional Coach

RETENTION AS A STATE POLICY MANDATE:

IREAD IN INDIANA

---

A Dissertation

Presented to

The College of Graduate and Professional Studies

Department of Educational Leadership

Indiana State University

Terre Haute, Indiana

---

In Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Velinda F. Stubbs

March 4, 2013

Keywords: Retention, social promotion, competency tests, reading assessment

COMMITTEE MEMBERS

Committee Chair: Terry McDaniel, Ph.D.

Assistant Professor, Department of Educational Leadership

Indiana State University

Committee Member: Steve Gruenert, Ph.D.

Chair and Associate Professor, Department of Educational Leadership

Indiana State University

Committee Member: Susan McDowell-Riley, Ph.D.

Deputy Superintendent, Office of Academic Affairs and Accountability

Evansville Vanderburgh School Corporation

## ABSTRACT

The interpretation of Indiana Public Law 109 and subsequent policy adopted by the Indiana Department of Education resulted in the Indiana State Board of Education mandating circumstances implemented during the 2011-2012 school year regarding grade level retention of Grade 3 students. IREAD-3, a standardized, gateway assessment, was administered to all Grade 3 students to determine eligibility to be promoted to Grade 4. Three quantitative studies analyzed the results from the initial year of assessment data for 1,712 students from one school district in Indiana to determine if there were factors that are predictive of performance on IREAD-3 and to better understand if there were effects on Grade 3 ISTEP+ performance based on the implementation of IREAD-3. Variables including chronological age, ethnicity, socioeconomic status (SES), gender, type of school the student attended (Title I versus non-Title I), and attendance were analyzed to determine if they were predictive of performance on IREAD-3. A logistic regression model identified three variables (low-SES, non-White, and poor attendance) that significantly increased the odds of not passing IREAD-3. The second study examined kindergarten, first grade, and second grade performance on DIBELS and TRC to determine if these assessments predicted passing IREAD-3. Based on the logistic regression model, below grade performance on both DIBELS and TRC (independently) significantly increased the odds of not passing IREAD-3. The statistically significant odds of not passing IREAD-3 were noted as early as the beginning of the kindergarten year but were noted to be more significant in later years, the middle and end of Grade 1 and beginning and middle of

Grade 2. The final study examined whether there was a difference in ISTEP+ performance for Grade 3 students who also took the IREAD-3 assessment as compared to performance of Grade 3 students during the previous three years of ISTEP+ administration when those students did not take IREAD-3—2009, 2010, 2011. The results suggested that although there was a statistically significant difference in scores over the four years, the effect size was insignificant. Practically, the difference appears to represent an upward trend of scores and the statistically significant differences were not necessarily associated with implementation of IREAD-3 in 2012.

## ACKNOWLEDGMENTS

At times our own light goes out and is rekindled from a spark from another person. Each of us has cause to think with deep gratitude of those who have lighted the flame within us.

Albert Schweitzer

During the long process of realizing this life-long goal, I agree with Albert Schweitzer that the fire, frequently, “goes out.” I would be remiss if I did not thank those who helped rekindle my inner spirit. My husband, Michael, has been a constant support. He was always there, quietly managing the distractions of life and keeping the home fires burning to make sure that I was able to focus on this work. My children, Noah and Katie, who bring joy and light to my life, were there to keep me “real” and make me laugh.

My parents, Frank and Vera Blackman, both amazing educators and role-models, instilled the love of learning, pursuit of excellence, work ethic, and mentality of service that have been my mainstays in life. They have always given me the support I needed to keep the fire burning.

I was fortunate to have many EVSC friends and mentors along the way. I learned so much from so many and would certainly miss naming all those who were influential. Dr. Jack Humphrey, Pat Taylor-Denham, and Dr. Jacqueline Neal helped light the passion for teaching. Friends and colleagues from Delaware and Cedar Hall Schools shared in hard work and laughter and kept the flame bright for so many children in Evansville.

Dr. Terry McDaniel, Dr. Steve Gruenert, Dr. Susan McDowell-Riley, and Dr. Dan Diehl are friends, mentors, teachers, and colleagues who answered my every question, pushed my thinking, and did what was needed to pass the torch and keep the fire burning.

EVSC leadership, Dr. Vincent Bertram and Dr. David Smith, were both visionary leaders who supported the idea of a school district, doctoral cohort in Evansville allowing many to realize their dreams and pursue the light of learning.

The EVSC Cohort, a group of amazing, dedicated individuals, worked together to pursue individual goals and created a supportive, collaborative professional learning community. The EVSC was made brighter by each of their contributions.

## TABLE OF CONTENTS

ABSTRACT.....	iii
ACKNOWLEDGMENTS .....	v
LIST OF TABLES .....	x
INTRODUCTION TO THE STUDY.....	1
Statement of the Problem.....	4
Purpose of the Study .....	5
Research Questions.....	6
Limitations, Assumptions, and Design Controls .....	7
Key Terms.....	7
Summary.....	8
REVIEW OF RELATED LITERATURE .....	10
Education Accountability and Social Promotion.....	11
Prevalence of Retention .....	14
Retention versus Social Promotion.....	15
Retention and Social Promotion (Meta-Analysis) .....	16
Retention and Dropping Out of School .....	20
High Stakes Assessment, Minimum Competency, and Retention.....	21
State Policy Regarding Promotion and Retention .....	23
Retention and Related Factors .....	25

Race and SES Factors .....	26
Gender and Retention .....	27
Reading Performance and Retention .....	28
Screening for Early Reading Problems.....	29
The Use of Standardized Assessments to Determine Grade Promotion.....	31
Implications Regarding Retention .....	33
Summary.....	35
RESEARCH DESIGN AND METHODOLOGY .....	38
Research Questions.....	38
Research Hypotheses (Null) .....	39
Population and Sample .....	39
Data Collection and Instrumentation .....	40
Data Analysis .....	42
Summary.....	42
RESULTS .....	44
Selection and Student Samples .....	45
Student Demographic Information .....	45
Student Age.....	46
School and SES.....	47
Data Analysis and Findings .....	48
Analysis of Questions 1 and 2 .....	48
Analysis of Question 3.....	51
Analysis of Question 4.....	54

Summary .....	56
SUMMARY .....	57
Conclusions.....	57
Research Questions 1 and 2 .....	57
Research Hypotheses (Null) for Questions 1 and 2 .....	58
Research Question 3 .....	60
Research Hypothesis (Null) for Question 3 .....	60
Research Question 4 .....	61
Research Hypothesis (Null) for Question 4 .....	62
Limitations of the Study.....	62
Implications.....	63
Implications for Practice .....	63
Implications for Public Policy .....	65
Implications for Further or Related Research.....	67
Conclusion .....	69
REFERENCES .....	72

## LIST OF TABLES

Table 1. Demographic Information for Grade 3 IREAD Students .....	46
Table 2. Student Age.....	47
Table 3. School Title I and Student SES Status.....	48
Table 4. Descriptives of Independent Variables Predicting Not Passing IREAD-3.....	49
Table 5. Coefficients for the Logistic Regression Analysis Predicting Not Passing IREAD-3 .....	50
Table 6. Coefficients for the Logistic Regression Analysis Predicting Not Passing IREAD-3 Based on DIBELS .....	52
Table 7. Coefficients for the Logistic Regression Analysis Predicting Not Passing IREAD-3 Based on TRC .....	53
Table 8. Descriptive Statistics for ISTEP+ Scores (2009, 2010, and 2011, combined) and ISTEP+ Scores 2012.....	54
Table 9. Descriptive Statistics for the Third Grade ISTEP+ Scores 2009-2012 .....	55

## CHAPTER 1

### INTRODUCTION TO THE STUDY

The past 20 years of education policy, resulting in increased accountability combined with a national focus on economic development, resulted in increased pressure on K-12 schools to generate students who are college and career ready. Policymakers identified the elimination of social promotion as one means to ensure improved student performance. House Enrolled Act 1367, also known as Public Law 109 (2010) required the Indiana State Superintendent of Public Instruction in conjunction with the Indiana State Board of Education to develop a plan to improve reading skills of students to be enacted by the spring of 2012. The bill has been interpreted to ensure that students can read before accessing Grade 4 curriculum by administering a test of reading, Indiana Reading Evaluation and Determination (IREAD-3), in the spring of Grade 3. For students who do not pass the reading assessment, remediation was to be provided by the local school, and students must have the opportunity to retake the test in mid-summer. Students who do not successfully pass the IREAD-3 test must repeat the Grade 3 reading curriculum, retake the IREAD-3 and the Grade 3 ISTEP+ test the following spring. Essentially, the students who do not pass IREAD-3 must be retained in Grade 3. Exemptions to retention are exclusively for students with disabilities who have an Individual Education Plan (IEP), when their case conference committees determine promotion was appropriate; students who have been retained twice; or English-learner students whose Individualized Learning Plan

(ILP) committees determine promotion was appropriate (Indiana Department of Education, 2010a).

Minimum competency requirements for promotion to the next grade are not novel to Indiana. The 1983 report, *A Nation at Risk: The Imperative for Educational Reform*, proved to be the catalyst for a reform movement in education in the United States (National Commission on Excellence in Education, 1983). This reform was followed by subsequent movements including *America: 2000* (U.S. Department of Education, 1991), *Goals 2000: Educate America Act* (1993) and the far reaching *No Child Left Behind Act* (NCLB, 2002). These efforts focused the nation on higher expectations for student performance and stronger accountability measures (Caples, 2005). In 2002, 11 states reported requiring minimum competency tests used to determine student promotion (Council of Chief State School Officers, 2003).

Based on information from the Common Core State Standard Initiative (2011), 46 states have adopted the Common Core State Standards coordinated by the National Governors Association Center for Best Practice and the Council of Chief State School Officers. The Common Core State Standards focus on college and career readiness and literacy skills across content areas. These national standards require students in primary grades to read fluently and demonstrate comprehension through application at higher levels in order to be successful (Common Core State Standard Initiative, 2011). The national discourse regarding education reform has focused on workforce development with the expectations for students in the early grades becoming increasingly rigorous. Legislators and the general public agree that setting higher expectations and holding schools, teachers, and students accountable will lead to the elimination of educational inequities and ensure a well-educated workforce.

Ending social promotion and retaining students who are not ready to progress appears to be a natural result of the past 30 years of educational reform. However, a debate continues regarding the effectiveness of the practice. The inconsistency of results appeared to stem from issues with methodology (Alexander, Entwisle, & Dauber, 2003; Jackson, 1975; Jimerson, 2001; Roderick & Nagaoka, 2005; Shepard & Smith, 1990, Wu, Hughes, & West, 2010). Many have viewed retention as having long-term negative effects on a child's self-esteem and point to the limited research available that shows long-term positive results. Others have supported retention as giving the child *the gift of time* to mature and master academic requirements (National Education Commission on Time and Learning, 1994). Researchers have identified factors associated with an increased probability of retention as well as other factors related to retention including gender, socio-economic status, and ethnicity (Entwisle, Alexander, & Olson, 2007; Hauser, Frederick, & Andrew, 2007; Hauser, Pager, & Simmons, 2004; McCollum, Cortez, Maroney, & Montes, 1999).

Most proponents of high stakes testing with minimum competencies for promotion make the claim that by identifying and providing failing students with additional time, resources, and scientifically based teaching methods, students at risk are offered the best hope for success (Heubert & Hauser, 1999; Shepard, 2004). Most states and districts that preceded Indiana in the adoption of high-stakes competency promotion policies provided additional supports to students who were at risk of being retained or have been retained, such as additional time for reading instruction, small group tutoring, summer school, before- and after-school tutoring. Findings from these specific states and districts showed little difference or slightly favored the students who were retained when comparing academic results; however, the positive results diminished over time (Alexander et al., 2003; Christensen, 2010; Wu et al., 2010; Xia & Kirby, 2009). The

cost of providing an additional year of school for students who may not realize long-term benefits has an economic impact that should be considered. In addition, a factor noted in the literature was the relationship between students who have been retained with an increased risk of dropping out of school (Roderick, 1995).

### **Statement of the Problem**

The interpretation of Indiana Public Law 109 and subsequent policy adopted by the Indiana Department of Education resulted in mandating circumstances regarding when to retain a student. The policy was enacted with conflicting empirical evidence of the effects of such policy. Indiana, following many other states and school districts in implementing a high-stakes, standardized competency test of reading for students in Grade 3, made the assumption that the threat of retention or retention for those students who do not pass the assessment will lead to improved academic achievement. Decades of research regarding retention has been inconclusive regarding improved achievement. Although the policy of minimum competency reading assessment in Grade 3 has been implemented in Texas, Florida, Ohio, and Chicago Public Schools, the IREAD-3 assessment was a unique assessment and new to Indiana. No research had been conducted to determine the impact of this particular assessment, making it difficult for Indiana school personnel to predict which students were most at risk of failure and how the assessment may affect other high-stakes assessments. Indiana schools were confronted with the responsibility of responding to the IREAD-3 policy in order to best meet the needs of their students and ensure successful outcomes. Indiana educators faced the challenge of determining the

- Impact of IREAD-3 on the educational process, including the impact on the results of other accountability assessments and student achievement;

- Student factors that correlated with the results of the assessment; and
- Factors correlated with improved performance on the summer administration of the assessment.

Retention represented an expensive intervention for struggling students. Based on Indiana figures from 2008, adding an additional year of school for students represents a minimum of an additional average per pupil expenditure of over \$11,000 (Indiana Youth Institute, 2011). If 10% of the current Grade 3 population in Indiana were to repeat a year of K-12 education, approximately an additional \$47,000,000 of state funding would be required for additional year of education for those students (Indiana Department of Education, 2011a, 2011b).

In drafting and enacting Public Law 109, Indiana legislators realized the critical importance of reading as the foundation skill to all academic success. In order for the requirements of Public Law 109 to realize the goal of ensuring students who were able to read, comprehend, and progress to college and career readiness successfully, schools must be aware of what practices have proven successful, how to identify students at risk of failure early, and how to minimize potentially harmful results of retention.

### **Purpose of the Study**

This study examined the results of the policy implementation of Public Law 109 and IREAD-3 as well as the initial results of IREAD-3 in Indiana. In addition, the study adds to the body of knowledge regarding the minimum competency promotion policies and informs educators and others interested in the debate regarding social promotion and retention.

The purpose of this experimental study was to compare demographic variables including gender, socioeconomic status (SES), age, ethnicity, school attendance, and type of school attended to performance on IREAD-3 for Grade 3 students in the Evansville Vanderburgh

School Corporation. In addition, a comparison was made of performance on early indicators of reading performance including Dynamic Indicators of Basic Early Literacy Skills (DIBELS) and Text Reading Levels (TRC) using Wireless Generation to performance on IREAD-3 for Grade 3 students in the Evansville Vanderburgh School Corporation (Wireless Generation, 2012).

Finally, the overall performance on ISTEP+ in Grade 3 during the 2011-12 school year, when Grade 3 students also took IREAD-3, was compared to overall performance on ISTEP+ in Grade 3 for students who did not take IREAD-3 during the previous year's test administrations.

### **Research Questions**

The following questions were examined in this study of IREAD-3.

1. Can passing IREAD-3 be predicted by student chronological age (age compared to peer group by quartiles), ethnicity, SES of the family, gender, type of school attending (Title I versus non-Title I) or attendance?
2. If passing IREAD-3 can be predicted, what variables when considering student chronological age (age compared to peer group by quartiles), ethnicity, socio-economic status of the family, gender, type of school attending (Title I versus non-Title I), and attendance are most predictive?
3. Can passing IREAD-3 be predicted by performance on DIBELS and TRC level in kindergarten, first grade, and second grade?
4. Is there a significant difference between average ISTEP scores in Grade 3 prior to the implementation of IREAD-3 and average ISTEP scores in Grade 3 after implementation of IREAD-3?

### **Limitations, Assumptions, and Design Controls**

This study reviewed data from a sample of 1,720 students, all attending the same countywide school district in southwestern Indiana during the same school year. The students represented a wide range of SES and were primarily White, Black, Hispanic, or bi-racial. The sample included students who attended schools within city limits, in suburban areas, and in some rural areas. Due to the nature of the subjects, the study may not be readily generalized to other areas in the United States without some caution. In addition, the study was specific to the assessment and policy in Indiana; and results may not be generalized to other states. The study assumed that the IREAD-3 assessment was a valid and reliable assessment. Further, it assumed that students received similar educational experiences. An additional limitation was that the study only reviewed one year of data, during the first year of administration of the IREAD-3 assessment.

### **Key Terms**

*Common Core State Standards* are academic standards adopted by a majority of states and were coordinated by the National Governors Association Center for Best Practices (NGA Center) and the Council of Chief State School Officers (CCSSO). These standards are consistent for K-12 and focus on the preparation of children for college and the workforce.

*Individual Education Plan (IEP)* is a written statement for children with disabilities, which includes the child's level of academic functioning, a statement of academic goals, how the child's disability affects the child's involvement and progress in the general education setting, and a description of services.

*Indiana Reading Evaluation and Determination (IREAD-3)* is used to measure foundational reading standards through Grade 3. Based on the Indiana Academic Standards,

IREAD-3 is a summative assessment that was developed in accordance with Public Law 109, which requires the evaluation of reading skills for students who are in Grade 3 beginning in the Spring of 2012 to ensure that all students can read proficiently before moving on to Grade 4 (Indiana Department of Education, 2010a).

*Indiana Statewide Testing for Educational Progress Plus (ISTEP+)* is used to measure student achievement in the subject areas of English/language arts, mathematics, science, and social studies. In particular, ISTEP+ reports student achievement levels according to the Indiana Academic Standards that were adopted in November 2000 by the Indiana State Board of Education (Indiana Department of Education, 2011c).

*Minimum competency test* is assessments meant to determine which students have mastered minimal expectations for promotion to the next grade.

*Oral reading fluency* is the ability to read accurately and quickly that is typically measured as words read per minute.

*Retention* is the practice of requiring a student to repeat a grade level, most often due to lack of mastery of grade level skills, social immaturity, or poor attendance. Similar terms are failure, grade retention, retention in a grade, repeating a grade, and over-aged for grade level.

*Social promotion* is the practice of sending a student who has not demonstrated mastery of grade level expectations on to the next grade level of school allowing them to progress with age-appropriate peers.

### **Summary**

Public Law 109 and the resulting Indiana Department of Education policy required all Grade 3 students to pass a minimum competency assessment (IREAD-3) in the spring prior to progressing to Grade 4 placed Indiana educators in a position of attempting to successfully

implement new policy and determine potential implications. IREAD-3 was anchored in educational theory that concluded that increased accountability would lead to increased academic achievement and that students must master basic reading skills prior to accessing more demanding content area curriculum in order to be successful. This study attempted to compare various student-level demographic variables and other formative reading assessment results to performance on IREAD-3 for Grade 3 students in the Evansville Vanderburgh School Corporation. In addition, a comparison was made of performance on state assessment (ISTEP+) prior to IREAD-3 and after IREAD-3.

## CHAPTER 2

### REVIEW OF RELATED LITERATURE

Beginning in the spring of 2012, Indiana Grade 3 students were required to demonstrate reading proficiency as measured by a standardized assessment (IREAD-3) prior to progressing to Grade 4. This policy was legislated in response to increased accountability and efforts to ensure that every student graduated from K-12 education, college and career ready. The state of Indiana followed several states and districts with similar policies mandating standardized requirements for promotion to the next grade.

Evidence existed that both supported and refuted positive student outcomes resulting from retention. The use of minimum competency assessments as the single determining factor for the decision to promote a student also has resulted in inconclusive results. The IREAD-3 assessment was unique to Indiana and no research had been conducted to determine the impact of this particular assessment. Indiana schools were confronted with the responsibility of responding to the IREAD-3 policy in order to best meet the needs of their students and ensure successful outcomes. Educators faced the challenge of determining

- The impact of IREAD-3 on the educational process, including the impact on the results of other accountability assessments;
- If there were student factors that correlated with the results of the assessment; and

- What, if any, factors correlated with improved performance on the summer administration of the assessment.

This chapter provides an overview of factors related to the current policy emphasis on ending social promotion in schools, with a specific focus on factors related to the implementation of IREAD-3 in Indiana. A review of related research regarding grade retention dating back to 1900 is included as well as a review of research on retention when triggered by minimum competency assessments. The incidence of retention is discussed as well as the policies of various states regarding social promotion. Research regarding utilization of standardized assessments for high stakes decision-making including promotion is presented. Literature that considered factors related to retention are reviewed including socio-demographic, gender, age, the correlation of reading proficiency in the primary grades, and performance on standardized assessments.

### **Education Accountability and Social Promotion**

During the past 30 years, accountability for education has received unprecedented national prominence. From Capitol Hill to the White House, the call has been issued for efforts to improve achievement, leave no child behind, and prepare all students for college and career success and, thus, have a lasting positive impact on our nation's economic future. The critical importance of reading has been acknowledged and supported by government leaders who wanted to ensure that students were ready to read in order to learn more complex content. Because education is in reality a function of individual states, many states have responded with a variety of policies to address the issue of reading readiness and social promotion.

*A Nation at Risk: The Imperative for Educational Reform* raised the question of whether the United States was producing students who were academically superior and able to compete

with students on a global scale (National Commission on Excellence in Education, 1983). *A Nation at Risk* described our national educational system as “a rising tide of mediocrity” (National Commission on Excellence in Education, 1983, p. 1) threatening our way of life and national security. Not since the era of Sputnik had legislators been so focused on public education. *A Nation at Risk* preceded other national education reform movements including *America: 2000* (1991), *Goals 2000: Educate America Act* (1993) and the far-reaching *No Child Left Behind Act (NCLB, 2002)*. As the 21st century unfolded, the nation was focused on public education with demands for higher expectations for student performance, stronger accountability measures, and improved performance for all students (Alexander et al., 2003; Bowman, 2005; Caples, 2005; Hauser et al., 2007; Jimerson, 2004; Shepard & Smith, 1990).

The quest for academic accountability resulted in efforts to ensure that all students were ready to progress and be successful in the next grade level. Students who were sent on to the next grade without the basic skills to be successful were seen to be doomed to failure. The practice of social promotion was viewed by many as counterproductive to the success of the educational process. An effort to end social promotion was clearly part of the national agenda for improved accountability in public education. President Clinton called specifically for an end to social promotion as one factor to help propel our nation into the 21st century. He noted that children should not proceed to Grade 4 without a firm foundation of reading skills (Clinton, 1998). Some states and individual school districts were early adopters and instituted standards-based assessments including minimum competency assessments for identified gateway years. In 1996, the Chicago Public School System instituted minimum competency tests in Grades 3, 5, and 8 based on student performance on the Iowa Test of Basic Concepts in reading and mathematics. Students not passing were required to attend summer school and then repeat the

grade if they did not pass minimum cutoff scores for their grade level (Jacob, Stone, & Roderick, 2004). President Bush, at the National Education Summit in 1999, praised the Chicago Public School System for instituting minimum competency tests and ending social promotion (Bush, 1999). Bush (1999) acknowledged that this policy was not without *painful consequences*. The U. S. Department of Education (1999) guidance, *Taking Responsibility for Ending Social Promotion*, advised educators and state and local leaders that

taking responsibility for ending social promotion requires tough decisions and strong actions by states, districts, and schools. It means requiring students to meet performance standards in key subject areas, at key transition points, in order to advance to the next grade level. It involves setting clear expectations for students and explicit policies for promotion, and adopting measures to hold all stakeholders accountable for student performance. (p. 5)

President Obama continued the call for increased education accountability; however, his message focused on the need to ensure more Americans were college and post K-12 career ready in order to secure the economic future of the nation (Obama, 2011). The call from the White House was heard by many states and local school districts where legislation and policy mandated the end of social promotion and instituted assurances that children demonstrate the ability to read successfully prior to progressing to Grade 4.

As with many other states including Florida, Texas, and Oregon, Public Law 109 required the Indiana State Superintendent of Public Instruction in conjunction with the Indiana State Board of Education to develop a plan to improve reading skills. The Indiana plan was to be enacted by the spring of 2012. As implemented by the Indiana Department of Education, in order to demonstrate reading competency, all Indiana Grade 3 students must pass the IREAD-3

assessment or be retained in Grade 3. Students who do not pass IREAD-3 are required to repeat the Grade 3 reading curriculum, officially be coded as a Grade 3 student, receive additional interventions in reading and retake both IREAD-3 and the Grade 3 ISTEP+ examination (English language arts and mathematics) the following spring. Exemptions to retention are limited to students with disabilities who have IEPs and whose case conference committees determined promotion is appropriate, students who have previously been retained twice, or English-learner students whose ILP committees determine promotion is appropriate (Indiana Department of Education, 2010a).

### **Prevalence of Retention**

When reviewing grade-level retention, failure, or repeating a grade, empirical data regarding grade retention has not been collected on a national scale. Much has to be inferred from other survey information (Hauser et al., 2004; Hauser et al., 2007; Shepard, 2004). A national survey indicated that more than 89% of students in kindergarten through Grade 12 reported that they had never been retained in a grade, indicating that approximately 10% of students had been retained (U.S. Department of Education, 2007).

In a study by Frederick and Hauser (2008), data from the current population survey from 1972-2005 were analyzed to determine the prevalence of retention. The data do not directly report students who were retained or repeated a grade; however, Frederick and Hauser conducted a complex analysis of the ages of individuals in the survey data and compared age to grade to determine those individuals below modal (expected) grade. The analysis revealed that the percentage of students below modal grade after school entry was fairly consistent at 13% through the cohort born in 1976. The percentage appeared to decrease to between 5% and 10% after the 1983 cohort but rebounded to the 13% level for the cohorts born in the 1990s. Frederick and

Hauser also found that the data from 1990 to present indicated that younger students were more likely than older students to be retained. Based on their analysis of the data, Frederick and Hauser concluded that retention had not significantly increased from the 1980s; however, younger students appeared to be retained more frequently than older students.

Some states reported grade retention rates, and those reported varied significantly from state-to-state. Indiana reported only 1.5% of students retained and North Carolina reported 4.5% of students retained per year (Shepard, 2004). Texas reported that during the 2009-2010 school year, 3.6% of students were retained (Shepard, 2004). However, Shepard (2004) cautioned not to confuse annual percentage of students retained with the cumulative percentage for a particular cohort of students. In other words, as a group (cohort) of students of the same age progressed through school, each year a percentage of students from the cohort were retained. Based on data from North Carolina, which followed a cohort of students from 1987-1997 using the cumulative model, by the end of Grade 9, 41% of the cohort of students had been retained. Retention varied widely by year from .5% in Grade 5 to 7.4% in kindergarten, 7.2% in Grade 1 and 15.8% in Grade 9 (Shepard, 2004).

### **Retention versus Social Promotion**

The current U.S. system of public education makes the assumption that all students enter school on equal footing with their age-appropriate peers and that they will progress at similar rates through the grade levels. As higher accountability measures were enacted, including scrutiny of high school graduates and student performance in higher education, educators looked to the preceding grade levels to offer explanations as to why so many students were struggling regarding academic performance. The notion that students were being passed on to the next grade without adequate skills became a reasonable and defensible explanation for poor

performance. A renewed focus on acquisition of reading skills and back to basics education surfaced as a strategy to combat less than adequate academic achievement.

Social promotion was defined by the U.S. Department of Education (1999) in *Taking Responsibility for Ending Social Promotion* as “a practice where students are allowed to continue to pass through school with their peers without satisfying academic requirements or meeting performance indicators at key grades” (p. 4). A review of relevant literature revealed several related terms including being held back (in a grade), retained, retention, flunking a grade, failing a grade, and failed (Jackson, 1975; Jimerson & Kaufman, 2003; Shepard & Smith, 1990). These terms refer to a practice that began with the advent of graded schools where students who were not socially or academically ready to progress were *held back* in a grade. Many viewed the idea that children who were not ready to progress needed the *gift of time* in order to mature and have additional time to master materials before progressing to the next grade and possible academic failure.

### **Retention and Social Promotion (Meta-Analysis)**

Several authors have conducted meta-analyses of the literature regarding grade retention, failure in a grade, or ending of social promotion, which yielded mixed results. Although grade retention or failure has been a common practice dating back to the turn of the 20th century, there remains a debate regarding the effectiveness of the practice. Many authors attributed the inconsistent results to issues with methodology stemming from when achievement was measured and compared, lack of comparability of test scores across grades, and difficulty in constructing matched comparability groups (Alexander et al., 2003; Jackson, 1975; Jimerson, 2001; Roderick & Nagaoka, 2005)

An early study of the topic, frequently cited in the literature, was conducted by Jackson (1975), who noted the longstanding history of retention in the United States, revealing five published reports on the topic prior to 1930. Jackson's thorough literature review, which reviewed reports dating from 1911-1975 regarding grade retention, resulted in 44 studies. Jackson (1975) categorized these studies by three prevailing analytical designs. His analysis revealed the following:

- Studies that attempted to match students retained under normal school policy with students not retained under normal school policy created a bias toward the achievement of retained students due to the lack of common metrics for comparison.
- Studies that compared the same students before and after retention Jackson found to be biased toward improved achievement of retained students due to the inability to control for other factors contributing to improvement.
- Three of the 44 studies reviewed used a random assignment of students. None of the three indicated significant differences in academic achievement of similar students who were retained and promoted based on random assignments.

Based on his analysis, Jackson (1975) concluded that there was no reliable evidence to support either retention or social promotion for students who were struggling. Jackson attributed the lack of evidence to the inadequate methodology of the studies. As noted by Johnson (1984), given the timeframe and nature of schools that Jackson reported, the information may hold little relevance to today's students and schools.

More recently, Jimerson (2001) reviewed 20 studies from 1990 until 1999 that met the criteria of publication in professional publications, journals, or books, using identifiable comparison groups of promoted students, and addressing academic or social-emotional

achievement. Utilizing an alpha level of  $p < .05$  as the criterion for significance, Jimerson (2001) found

- Studies comparing academic achievement favored the matched comparison group over the retained students in 47% of the studies; 5% of the studies favored the retained group and 48% indicated no significant differences between groups.
- Two-thirds of the analysis that favored the retained students reviewed data from the year after retention but did not review long-term data.
- Authors of four studies reported favorable results for retention; 16 reported no difference or results that did not support retention.

In a meta-analysis conducted by Shepard (2004), the majority of studies reviewed showed no gain or loss in academic achievement of those students who were retained when achievement scores were weighted equally. A study completed on first-grade students in Texas by Wu et al. (2010) attempted to match students predicted to be retained with those students who were actually retained. This study, unlike earlier studies, found improved social-emotional indicators for retained students. Students who were retained were shown to have decreased teacher-rated hyperactivity, decreased peer-rated sad/withdrawn behaviors, increased teacher-rated behavioral engagement in both the short and long term and an increase in acceptance by peers maintained in the long term. Regression model analysis also suggested that a positive sense of school-belonging improved in the short term but did not persist over time. Additional results indicated positive results for reading and math achievement, but those achievement results did not persist beyond two years (Wu et al., 2010). Xia and Kirby (2009) reviewed 91 studies from 1980 to 2008 and found that retained students did not appear to benefit

academically from retention, and those gains that were documented were brief and tended to fade with time.

Allen, Chen, Willson, and Hughes (2009) conducted a meta-analysis of studies conducted between 1990 and 2007, reviewing only those studies that (a) were published and peer reviewed, (b) reported outcomes at the student level, (c) used a quasi-experimental or experimental design, and (d) reported achievement outcomes that did not duplicate previously reported studies. Out of 199 original studies considered, only 22 met the established criteria. Studies were rated as low-, medium-, and high-quality based on design quality of the study. The medium- and high-quality studies had less negative student-level effects with the effect size on average .34 higher than studies with the low-quality. The effect size of low-quality studies was found to be -.30 and the effect size of medium- to high-quality studies was .04. The design quality of the study regarding retention accounted for a statistically significant amount of variance in effect size.

Most of the research regarding retention has compared students who were recommended for retention by their teachers with students who were not recommended for retention. A survey of teachers revealed that a majority of teachers believed that retention was an acceptable practice, that it acts to prevent failure, and is a source of motivation for students to work harder (Tomchin & Impara, 1992). Most teachers agreed that the practice was not harmful for students in primary grades, but beyond Grade 3 there was less agreement among teachers as to potential benefits and possible harm (Tomchin & Impara, 1992).

Alexander et al. (2003) summarized a review of research by stating that the most frequent finding is that students who have been retained and those who have not differ little in academic achievement during the year of retention or after the retention. Alexander et al. also found that retention in the early grades appeared to have positive results, particularly for students who were

retained at the end of first grade; however, the positive results diminished over time. Johnson (1984) found that academic achievement was improved when the decision to retain students was based on a lack of maturity and not solely on academic achievement.

### **Retention and Dropping Out of School**

A recurring factor in the literature was the relationship between students who have been retained and an increased risk of dropping out of school (Roderick, 1995). Alexander, Entwisle, Dauber, and Kabbani (2004) analyzed data from the Beginning School Study, a database of students who attended Grade 1 in the Baltimore City Public Schools in the fall of 1982, where students were tracked for eight years. A high percentage of these students were retained with 17% of the cohort retained in Grade 1 and over 40% retained by the end of elementary school. Factors associated with a high risk of dropping out of school were the same factors associated with a high risk of being retained in a grade, including low-SES, minority, high-poverty school, and parents with little formal education. The students in the study from the Baltimore Public Schools represented a high-poverty, highly at-risk of academic failure demographic and were likely to exhibit factors associated with dropping out as well as being retained in a grade (Alexander et al., 2004; Entwisle et al., 2007).

The Indiana Department of Education has used retention in school as one factor related to predicting students at risk of dropping out of high school in a matrix to predict dropouts for Grade 9 students. When reviewing the data regarding the age when students were retained, Alexander et al. (2004) found that students who were retained in later grades (4-7) were at significantly greater risk of dropping out of school than if retained in primary grades.

### **High Stakes Assessment, Minimum Competency, and Retention**

Whether to retain or promote a student was traditionally a decision based on multiple factors including academic achievement in reading and mathematics, grades, attendance, social-emotional maturity, speech and language development, physical development, physical size, and behavior. The decision to retain or promote a student based solely on minimum competency standards and high-stakes academic assessments appears to be the result of recent efforts to make the decision less subjective and more aligned with accountability. The research was limited regarding retention as the sole result of minimum competency and high-stakes testing.

From 1996, when the Chicago Public School System ended social promotion, to 2004, nearly one in five third graders and one in 10 sixth and eighth graders were retained (Roderick & Nagaoka, 2005). More recent studies that reviewed students retained due to minimum competency assessments appeared to favor minimal academic gains for students who were retained (Alexander et al., 2003). Jacob and Lefgren (2002) examined the records of 148,000 students in Chicago's third and sixth grades for 1997 until 1999. They compared students who had scored just above passing the retention test with those who scored just below passing. The students who failed the test were required to attend summer school and faced retention if they did not pass a second assessment. There was substantial academic improvement among the third graders who faced retention; however, the effect was insignificant for sixth graders who faced retention. Hauser et al. (2007) speculated that Jacob and Lefgren's findings may reflect incentives to improve test performance as well as programs to enhance student achievement in order to avoid retention. Shepard (2004) reviewed the studies conducted by Roderick on Chicago Public School students and concluded that those students who were at-risk of retention but passed minimum competency tests showed improved academic performance that appeared to

persist. Students who did not pass the minimum competency test and were retained did not show improvement in academic performance and were often retained a second time.

A study by Lorence, Dworkin, Toenjes, and Hill (2002) of students in Texas revealed that after repeating Grade 3, reading scores of students who had been retained surpassed their initial Grade 3 scores and remained comparable to those students who were promoted after controlling for initial differences between the students who were not promoted and socially promoted students. The advantage of the students who were retained was even more significant. The researchers found no evidence of adverse effects of retention on the students. Christensen (2010) reviewed data from three school districts in Texas where students were required to pass minimum competency tests in Grade 3 for reading and Grades 5 and 8 for reading and mathematics. Christensen found that students retained in Grade 3 did perform better in the following year, but their improvement was not sustained through later assessments in Grades 5 and 8. Students retained in Grade 5 for math and Grade 8 for math and reading did not show significant improvement; however, Grade 5 students retained for reading did show significant improvement in the following year and again in Grade 8.

Alexander et al. (2004) reviewed data from the Beginning School Study of Baltimore Public School students' reading performance and found that in Grades 1 and 3, during the repeat year, student growth was near that of the yearly growth of non-repeaters based on the results of the California Achievement Test of Reading, which suggested some lasting benefits for Grade 3 repeaters.

A study by the Florida Legislature, Office of Program Policy Analysis and Government Accountability (2006) found that the percent of Grade 3 students repeating after the minimum reading competency promotion policy was instituted rose from 3% in 2001-02 (prior to the

policy) to 13% the following year. They also found that 65% of the students who scored at a Level 1 on the reading assessment scored at a Level 2 or above in Grade 4 after repeating Grade 3. Of the students who were promoted and scored a Level 1 in Grade 3, only 44% scored at a Level 1 or above in Grade 4. The findings were not analyzed for statistical significance.

### **State Policy Regarding Promotion and Retention**

Many states have addressed the issue of academic accountability by mandating gateway competency standards in particular grade levels. Framers of policy appear to assume that retention and the threat of academic failure leads teachers, students, and their parents to work harder or that the extra year of instruction will give students the time to improve their skills (Roderick & Nagaoka, 2005). Policy regarding grade retention implementation varied widely from state to state. In a review of state educational policy, Zinth (2005) found that 18 states had indicated an assessment was to be considered in the decision regarding retention and 12 states had instituted promotion gates (grades levels with correlating thresholds of achievement to be demonstrated prior to promotion). In a more recent review, Xia and Kirby (2009) found that states were frequently adjusting existing policies and many additional states were adding language regarding retention and promotion.

The Chicago Public Schools System (2011) considered several factors regarding Grade 3 retention including scoring at a minimum of the 24th percentile on a district-wide assessment of reading, a C or higher report card grades, no more than nine unexcused absences per year and attendance at a summer school program. In Texas, beginning in 2002-03 to 2008-09, students in Grade 3 were required to pass the state reading test to advance to Grade 4. Students in Grade 5 were required to pass the reading and mathematics tests beginning in 2004-05. Starting in 2007-08, students in Grade 8 were required to pass the reading and mathematics tests. Students were

given three opportunities to pass the tests and schools were required to offer accelerated instruction in the subject areas failed. The Texas assembly recently passed legislation to retain minimum competency in Grades 5 and 8 but to drop the requirement for Grade 3 (Christensen, 2010; Texas Education Agency, 2011).

In California, each school district was required to develop a written policy regarding retention and promotion approved by the local board as per state code. The policy was required to address promotion to Grades 3, 4, and 5 and the beginning of middle and high school, and it was also required to include grades and measures of academic performance (California Department of Education, 2011).

Florida instituted an end to social promotion during the 2002-2003 school year. Students were required to meet performance standards based on the Sunshine State Standards as defined by the Commissioner of Education (Florida Department of Education, 2009). A priority was placed on students with a reading deficiency as measured by a score of less than two on the Florida Comprehensive Assessment Test at the end of Grade 3. Those students were required to be retained. Retentions in Florida in Grade 3 ranged from a low of 13,666 students in 2007-08 to a high of 27,713 in 2002-03 when the policy was instituted (Florida Department of Education, 2009).

New York City public schools instituted a policy for promotion in Grade 3 in 2003-04 (New York City Department of Education, 2009). The policy was extended to include Grade 5 in 2004-2005, Grade 7 in 2005-2006, and Grade 8 in 2008-2009. To progress from Grade 3 to 4, students must be at or above a proficiency level of 2 on the English language arts and mathematics assessment. The process was modified to authorize principals to determine whether to promote based on a student work portfolio, teacher observations, and grades in academic

subjects. The data on individual students was reviewed in June to determine need for summer school and again in August to make a final determination regarding grade placement (New York City Department of Education, 2009; Xia & Kirby, 2009).

The Colorado legislature recently passed the Colorado Early Literacy Act (2012) to be enacted during the 2013-14 school year, which requires the use of one of four screening instruments to determine if students are reading at grade level for kindergarten through Grade 3. The bill requires retention for those students who at the end of the school year demonstrate a significant reading deficiency (Colorado Early Literacy Act, 2012).

Beginning in 2012, Indiana students in Grade 3 must demonstrate reading proficiency by passing the IREAD-3 assessment in the spring of the Grade 3 year. Students who do not pass are offered the opportunity to retake the test in the summer. A student who was not successful at passing the spring or summer IREAD-3 must repeat Grade 3 reading curriculum, be officially coded as a third grader, and retake IREAD-3 and the English language arts and mathematics Grade 3 ISTEP+ assessments. Exemptions to retention are exclusively for students with disabilities who have an IEP and their case conference committees determine promotion is appropriate, students who have been retained twice, or English-learner students whose ILP committees determine promotion is appropriate (Indiana Department of Education, 2010a).

### **Retention and Related Factors**

A further review of literature regarding retention indicated several factors associated with the likelihood of being retained. The variables were those that children possessed prior to the decision to retain a student and may or may not be related to retention. They included demographic and descriptive factors including race, gender, SES, and age, as well as academic

factors. Attainment of proficiency in reading was the academic factor most frequently associated with decisions regarding retention.

Hauser et al. (2007) reviewed Current Population Survey data and determined that from 1996-2005 retention of Grade 1 students was considerably higher than in any other grade level with more than 7% of students retained in Grade 1. Silbergitt, Jimerson, Burns, and Appleton (2006) compared students who were retained early in their schooling to later in their school career and found that students who were retained early had a more consistent progress and rate of academic growth. Students retained later had a rapid deceleration of growth. It appeared that there were more negative effects from later retentions based on the socioemotional impact of retention.

### **Race and SES Factors**

Students who were retained appeared to be at the crux of the achievement gap between students of poverty and affluence and between Black, Hispanic, and White students (Alexander et al., 2003; Hauser et al., 2004; Hauser et al., 2007; Lorence et al., 2002; Planty et al., 2009; Shepard, 2004). The National Center for Educational Statistics reported that 16% of Black, 11% of Hispanic, and 8% of White students reported having been retained. Black male students were more than two times as likely to be retained as compared to White or Hispanic male students, and students of poverty were over four times as likely to be retained as students from non-poverty families (U.S. Department of Education, 2007).

Hauser et al. (2004) found that Black students were less likely than Whites to be enrolled in Grade 1 early (below age six). By age nine, Blacks and Hispanics were almost 50% more likely to be overage for grade level than Whites, and by age 17, Blacks and Hispanics were almost 150% more likely to be overage for grade level than Whites. However, Hauser et al.

(2004) found that when controlling for all other factors, race-ethnic differentials in age and grade placement were almost entirely due to differences in socioeconomic and family background. For students at age 17, minority status increased the likelihood of being overage for grade level when all other factors are controlled. In addition, region and metropolitan area did not appear to be a determining factor in students who were retained. Hauser et al. (2007) found that based on review of data from the Current Population Survey from 1996-2005, the gap between Blacks and Whites who were retained was consistent but the retention of Black students began a slight increase in 2000.

Russo (2005) reported that the Chicago Public Schools, when implementing minimum competency testing for promotion, saw retention rates of limited-English-proficient students raise from 6.8% to over 20% with a peak of nearly 15,000 students retained in 1998. Lorence et al. (2002) reviewed the entire population of Texas students over a nine-year period and found that being African American increased the odds of Grade 3 retention based on performance on standardized testing by 62% when compared to non-Hispanic White students. The odds of retaining Hispanic students were increased by 26% when compared with academically challenged non-Hispanic White students. In a Rand report analyzing the implementation of New York City School implementation of retention policy from 2004-2007, McCombs, Kirby, and Mariano (2009) found that from the Grade 3 cohort of 2004-05, 15% of Hispanic students and 21% of Black students were at risk of repeating the grade as compared to 3% of Asian or White students.

### **Gender and Retention**

There were significant differences found in the gender of students who were retained. Boys as compared to girls were more likely to be retained among all grade levels (Hauser et al.,

2004; Lorence et al., 2002; Planty et al., 2009; Shepard, 2004; Xia & Kirby, 2009). Hauser et al. (2004) compared student age to grade-level enrollment and found that boys were more likely to be overage for grade level, which increased with age from a 5% difference from girls between the ages six to eight to a 10% difference for ages 15-17. Based on data from 1972-1998, the likelihood of boys being retained as compared to girls increased with age. At age six, boys as compared to girls were 40% more likely to be retained, at age nine, 50% more likely to be retained, and at age 17, 70% more likely to be retained (Hauser et al., 2004).

An analysis of data from the Baltimore Beginning School Study by Entwisle et al. (2007) revealed that boys on free and reduced lunch were more likely to be retained in elementary grades. Lower reading marks and lower conduct marks at the end of Grade 1 were associated with retention. Boys were most likely to struggle in reading and conduct and, thus, be retained in a grade. Entwisle et al. (2007) summarized that the same root cause that resulted in a gender gap for retention also resulted in a gender gap for test scores. McCollum et al. (1999) found that 40% of students who repeat were from the lowest quartile of socioeconomic with only 8.5% from the highest quartile.

### **Reading Performance and Retention**

The majority of promotion policies consider reading attainment as a critical factor (Xia & Kirby, 2009). The acquisition of reading is undoubtedly a skill critical to future success both in K-12 education and in later life. Students who struggle with reading in elementary school are typically not successful in school. Difficulties in reading and other academic areas in primary grades have been found to persist into later school years (Entwisle, Alexander, & Olsen, 2005). In a survey conducted in 1999 to determine the scope of work of school psychologists, it was revealed that the majority (57%) of referrals to school psychologists were for problems related to

reading (Bramlett, Murphy, Johnson, Wallingsford, & Hall, 2002). The five areas of reading critical to successful readers as identified by the National Reading Panel were phonemic awareness, phonics, vocabulary, comprehension, and fluency (National Institute of Child Health and Human Development, 2000). The IREAD-3 assessment for Indiana assessed students on three of the five areas: phonics, vocabulary, and comprehension (Indiana Department of Education, 2010a).

### **Screening for Early Reading Problems**

The Indiana Department of Education (2010b) identified appropriate assessments used to screen students in kindergarten through Grade 2. DIBELS was identified as a research-based screening instrument that has correlation with success on standardized assessments in later grades.

In order to determine those students most at risk of reading difficulties, plan for appropriate interventions and instruction, and monitor student progress a number of oral reading fluency assessments have been utilized. The ability to read fluently was found to be closely related to reading comprehension (National Institute of Child Health and Human Development, 2000). Oral reading fluency has been utilized as an indicator of children's ability to read accurately and quickly. The DIBELS was based on the program of research and development of curriculum-based measurement of reading by Stan Deno (Shinn, Good, Knutson, Tilly, & Collins, 1992). It consisted of a standardized, individually administered test. Children were asked to read a short passage of text that had been calibrated for difficulty. After students read aloud for one minute, student performance was measured by the number of words correctly read from the passage per minute. If a student omitted words, substituted words, or hesitated for more than three seconds during the reading, deductions were taken. When a student self-corrected a

word within three seconds, the word counted as correct. The number of correct words per minute from the passage was the oral reading fluency score (Good, Kaminski, & Dill, 2007; Good, Simmons, & Kame'enui, 2001).

Several studies have documented the correlation of oral reading fluency measures and standardized assessments used for accountability purposes. Reschly, Busch, Betts, Deno, and Long (2009) conducted a meta-analysis to determine evidence of correlation between curriculum-based measurements of oral reading and standardized reading assessments. The study found a significant estimate of the  $z$ -transformed correlation on the state-specific tests,  $N = 70$ ,  $\gamma_0 = 0.77$ ,  $t(139) = 46.92$ ,  $p < .001$ . When compared to national tests of reading, the expected correlation coefficient with national tests of .74 was found to be significantly higher than the correlation coefficient of .65 with state-specific tests. Individually administered tests had a correlation of .83, which was significantly higher than the correlation coefficient of .71 with group-administered tests.

In a study by Roehrig, Yaacov, Nettles, Hudson, and Torgesen (2008), DIBELS Oral Reading Fluency (ORF) assessment was found to be significantly correlated with the Florida Comprehensive Assessment Test-Sunshine State Standards (FCAT-SSS) and the Stanford Achievement Test (SAT-10) for Grade 3 students in Florida Reading First schools when both assessments were administered during the same school year ( $r = .70$ ). There was also a statistically significant correlation with reading skills when the oral reading assessment was administered within the same school year as the standardized reading assessment. The correlation coefficient dropped from .70 to .63 when the standardized assessment was a year after the oral reading assessment administration. When comparing reading subtests with oral

reading assessments, the correlation coefficient for comprehension, vocabulary, and decoding subtests were all found to be .66 and found to be .82 for word identification subtests.

A similar study by Baker et al. (2008) correlated the ORF scores as measured by DIBELS and performance on the SAT-10 the following school year. The ORF assessments administered throughout the year in Grade 1 correlated with spring Grade 2 SAT-10 scores ranging from .63 to .80. ORF assessments administered in Grade 2 were correlated with spring Grade 3 SAT-10 ranging from .58 to .68. On the Grade 3 reading test as measured by SAT-10, the Grade 2 ORF level and slope accounted for 51% of the variance; however, Grade 1 ORF was found to be a better predictor of Grade 2 performance. The researchers offered the explanation that fluency alone may not be as critical to reading proficiency as children matriculate to Grade 3.

In addition to DIBELS, the Indiana Department of Education also identified Wireless Generation DIBELS 3-D, which included text reading comprehension (TRC) assessment. The TRC assessment included an oral running record of student reading accuracy as well as an oral and written comprehension assessment on texts of graduated difficulty (Wireless Generation, 2012). A measure of fluency alone (accuracy and speed of reading) was found to be not as predictive of comprehension especially as students developed as readers (Hosp & Fuchs, 2005; Samuels, 2006). Comprehension was found to be more accurately assessed when a comprehension component was added to an assessment of accuracy. Valencia et al. (2010) found that by adding an assessment of comprehension to rate, accuracy, and prosody they could account for 41% of the variance at Grade 2 and 51% at Grade 4.

### **The Use of Standardized Assessments to Determine Grade Promotion**

Linn (2000) reviewed the use of standardized assessments and indicated that the use of tests for school accountability purposes has not produced dramatic improvements in outcomes

and that the assessments have not been adequate for the demands placed on them. The position statement on grade retention by the National Association of School Psychologists (2003) was strongly opposed to the use of retention as a policy to help achieve higher academic standards.

Penfield (2010) noted that central to the issue of high-stakes tests being used as a basis for retention decisions, the following must be considered.

- Did the students have the opportunity to learn what was on the test?
- Did the tests measure the constructs that it intended to measure?
- Did the test lead to the intended educational goals?
- Were scores reflective of high-quality instruction?
- Did the test give students the opportunity to demonstrate the knowledge and skills they have mastered?

Penfield (2010) drew these guidelines from ones proposed by the National Research Council and the 1999 Standards for Educational and Psychological Testing to determine if the use of assessments for retention decisions was a fair and appropriate use. Penfield found that two areas were of particular concern and in his opinion violated fair and appropriate use. He suggested that the test results should not be attributed to poor instruction or linguistic or cultural content and that because marginalized groups have consistently performed poorly on such assessments there is a potential violation. In addition, Penfield questioned whether the testing led to consequences that are educationally beneficial, based on the conflicting body of research regarding the effects of retention. Several cases regarding test-based retention have been heard in the courts; however, the judicial system has been reluctant to interfere with educational policy as it relates to retention (Penfield, 2010).

### **Implications Regarding Retention**

Shepard (2004) noted that retention appears to generally be accepted as a one-time *intervention*, intended to give students an opportunity to catch up. In reality, for retention to be successful, students must learn more than the average student in one year in order to catch up and maintain an average rate of learning to keep up. The National Education Commission on Time and Learning (1994) noted that

Decades of school improvement efforts have foundered on a fundamental design flaw, the assumption that learning can be doled out by the clock and defined by the calendar. Research confirms common sense. Some students take three to six times longer than others to learn the same thing. Yet students are caught in a time trap-processed on an assembly line scheduled to the minute. Our usage of time virtually assures the failure of many students. (p. 7)

Retaining students in a program or class that was not successful for them the first time will not necessarily lead to success by repeating the year (Alexander et al., 2003; Allen et al., 2009; Jackson, 1975; Jimerson, 2001; Shepard, 2004; Roderick, 1995). Entwisle et al. (2007) found that of students who were retained and had been on free and reduced lunch, by the age of 22, 44% of the men and 52% of the women were neither gainfully employed nor in school.

After reviewing all meta-analyses published between 1990 and 1999, Jimerson (2004) concluded the evidence regarding retention does not support retention as an effective intervention for students and may even be considered educational malpractice. Students most likely to be retained are those from the most disadvantaged backgrounds, and those same students are the most likely to experience negative outcomes including dropping out of school (Jimerson, 2004; Xia & Kirby, 2009). After reviewing and critiquing several studies on

retention, Shepard (2004) called for a more rigorous evaluation that would hold up under the scientific scrutiny of the Federal Drug Administration for safe and effective treatment. Shepard purported that retention should be approached as an intervention for struggling students much as special education placement whereby the burden of proof is on determining that the benefits of placement outweighs potentially negative effects.

In a National Research Council report, Heubert and Hauser (1999) suggested early identification of students who are at risk of failure and provision of effective remedial education as an alternative to retention. Interventions including summer school, tutoring, one-on-one reading instruction, multi-aged and cross-grade classrooms, and before- and after-school programs need further investigation to determine which interventions are most likely to prevent retention and which in combination with retention lead to long-term academic success. Early identification and remediation supports that have been identified as effective are recommended to improve student achievement as well as prevent the need for retention (Heubert & Hauser, 1999; McCollum et al., 1999; Xia & Kirby, 2009). Shepard (2004) suggested that educators consider

Other treatments for poor achievement have a much greater chance of success. After-school programs, tutoring, summer school, and one-on-one reading instruction are more effective in raising achievement than repeating a grade, as shown by large positive results in their respective research literatures. (p. 201)

Retention as a policy mandate is an expensive intervention for struggling students. Based on Indiana figures from 2008, with an average per-pupil expenditure of over \$11,000, adding a year of schooling for each student retained is an expensive venture for the state (Indiana Youth Institute, 2011). If 10% of the current Grade 3 population in Indiana were to repeat a year of K-

12 education, approximately an additional \$47,000,000 of state funding would be required for additional year of education for those students (Indiana Department of Education, 2011a, 2011b). The state of Texas spent an estimated \$1.7 billion for the extra year of schooling for the children retained in the 2000-2001 school year as they instituted a policy to end social promotion (Texas Education Agency, 2011).

### **Summary**

The current U.S. system of public education makes the assumption that all students enter school on equal footing with their age-appropriate peers and that they will progress at similar rates through the grade levels. However, in the era of increased accountability, this notion has become a target of scrutiny. Through several presidential administrations, both Republicans and Democrats have called for improved student achievement in order to prepare a globally competitive workforce. From *A Nation at Risk* through *No Child Left Behind*, state and national policy responded with setting higher standards of achievement, requiring additional standardized testing and, in many states, enacting minimum competency or gateway assessments.

It has been difficult to determine the prevalence of grade-level retention, failure, or repeating a grade. Data has not been consistently collected on a national scale. From a U.S. Department of Education national survey, respondents self-reported that just over 10% had ever repeated a grade (U.S. Department of Education, 2007). Frederick and Hauser (2008) analyzed data from the Current Population Survey from 1972-2005 and found that based on the age of pupils, for students born in the 1990s, 13% appeared to have been retained. There was wide variance in what individual states reported regarding the number of students retained.

Grade retention or failure has been a common practice dating back to the turn of the 20th century. The review of literature does not clearly establish the effectiveness of the practice.

Many authors attributed the inconsistent results to issues with methodology stemming from when achievement was measured and compared, lack of comparability of test scores across grades, and difficulty in constructing matched comparability groups (Alexander et al., 2003; Jackson, 1975; Jimerson, 2001; Roderick & Nagaoka, 2005; Shepard & Smith, 1990, Wu et al., 2010). Findings generally favored the students who were retained when comparing academic results; however, the positive results diminished over time. A factor noted in the literature was the relationship between students who have been retained with an increased risk of dropping out of school (Roderick, 1995).

Several studies analyzed factors associated with the likelihood of being retained in a grade. Black male students were more than two times as likely to be retained as compared to White or Hispanic male students, and students of poverty were over four times as likely to be retained as students from non-poverty families (U.S. Department of Education, 2007). Boys when compared to girls were more likely to be retained among all grade levels (Hauser et al., 2004; Lorence et al., 2002; Planty et al., 2009; Shepard, 2004; Xia & Kirby, 2009). Students in the early grades who struggle with reading were most frequently identified for retention, and most promotion policies considered reading attainment as a critical factor (Xia & Kirby, 2009). Students most likely to be retained were those from the most disadvantaged backgrounds, and those same students were the most likely to experience negative outcomes including dropping out of school (Jimerson, 2004; Xia & Kirby, 2009).

The use of standardized assessments to determine retention has been questioned as whether it represented a fair and ethical use of standardized assessments. Penfield (2010) noted that due to the conflicting body of research regarding effects of retention, it was not clear that the testing led to consequences that were educationally beneficial. Retention is an expensive

intervention, having inconclusive evidence to support long-term effectiveness for students. If 10% of the current third-grade population in Indiana were to repeat a year of K-12 education, approximately an additional \$47,000,000 of state funding would be required for additional year of education for those students (Indiana Department of Education, 2011a).

## CHAPTER 3

### RESEARCH DESIGN AND METHODOLOGY

This study examined the results of the policy implementation Public Law 109 and IREAD-3 as well as the initial results of IREAD-3 in Indiana. The study will add to the body of knowledge regarding minimum competency promotion policies and inform educators and others interested in the debate regarding social promotion and retention. Public Law 109 and the subsequent policy adopted by the Indiana Department of Education resulted in mandating circumstances regarding when to retain a student in Grade 3. This policy was enacted in spite of decades of conflicting empirical evidence of the effects of retention and no relevant research regarding the implementation of this specific assessment and policy in Indiana.

Retention represents an expensive intervention for struggling students representing a minimum of an additional year of per pupil expenditure of over \$11,000 ((Indiana Youth Institute, 2011). To make prudent decisions, it is both fiscally responsible and in the best interest of students that educators have relevant information regarding IREAD-3. Educators need to be aware of what practices have proven successful, how to identify students at risk of failure early, and how to minimize any potentially harmful results of retention.

#### **Research Questions**

The following questions were examined in this study of IREAD-3.

1. Can passing IREAD-3 be predicted by student chronological age (age compared to peer group by quartiles), ethnicity, SES status of the family, gender, type of school attending (Title I versus non-Title I) or attendance?
2. If passing IREAD-3 can be predicted, what variables when considering student chronological age (age compared to peer group by quartiles), ethnicity, SES status of the family, gender, type of school attending (Title I versus non-Title I), and attendance are most predictive?
3. Can passing IREAD-3 be predicted by performance on DIBELS and TRC level in kindergarten, first grade, and second grade?
4. Is there a significant difference between average ISTEP scores in Grade 3 prior to the implementation of IREAD-3 and average ISTEP scores in Grade 3 after implementation of IREAD-3?

#### **Research Hypotheses (Null)**

1. Passing IREAD-3 cannot be predicted by student chronological age (age compared to peer group by quartiles), ethnicity, socio-economic status of the family, gender, type of school attending (Title I versus non-Title I), or attendance.
2. Passing IREAD-3 cannot be predicted by performance on DIBELS and TRC level in kindergarten, first, and second grade.
3. There is no significant difference between average ISTEP scores in Grade 3 before implementation of IREAD-3 and after implementation of IREAD-3.

#### **Population and Sample**

The sample for the study consisted of 1,719 Grade 3 students in attendance in the Evansville Vanderburgh School Corporation during the 2011-12 school year. Only those

students assessed using the IREAD-3 assessment in March of 2012 were included in the sample. Students who were English language learners as well as those students receiving special education services with the exception of students with severe cognitive or adaptive delays who were unable to attempt the test were included in the sample.

For Research Question 1 and 2, all students who were assessed by IREAD-3 were included in the sample. Research Question 3 included only those students who had been administered the end of the year kindergarten, first-, and/or second-grade DIBELS assessment, the TRC level assessment, and the IREAD-3 assessment. Research Question 4 included students who took both IREAD-3 and ISTEP+ in Grade 3 during the 2011-12 school year as well as students from the Evansville Vanderburgh School Corporation who took ISTEP+ in Grade 3 during the 2008-09 (1,444 students), 2009-10 (1,542 students), and 2010-11 (1,440 students) school years.

### **Data Collection and Instrumentation**

Grade 3 students in the state of Indiana were assessed on the IREAD-3 assessment in March 2012. The data file of pass/fail for students in the Evansville Vanderburgh School Corporation provided by the Indiana Department of Education was used in this study. Student data included date of birth, ethnicity, gender, lunch status, and school attending. All other identifying information was removed from the database. For Research Question 4, aggregate ISTEP data provided by the Indiana Department of Education were used from archived files for the 2010, 2011, and 2012 files.

The demographic data included in Research Questions 1 and 2 were collected from archival files as indicated by enrollment information. Students were coded based on the information as provided by parents or caregivers upon enrollment regarding chronological age,

ethnicity, and gender. The information was entered into the school system data base by school staff and then uploaded to the test data file. Schools were instructed to check for accuracy prior to uploading data to the Indiana Department of Education. To compare chronological age, student birthdates were used to recode the data. Students in Indiana typically by state policy must be 5 years old on or before August 1 in order to enroll in kindergarten (Indiana Department of Education, 2010b). In order to create a dichotomous variable, students were coded as being youngest in the class if they were in the youngest quartile or born after May 1, 2003. All other students who were born on or before May 1, 2003, were coded as oldest in the class. Data regarding socioeconomic status of the family of the student were collected from the indication of free lunch, reduced lunch, or paid lunch status from school data files. Parents and/or caregivers self-reported by completing lunch applications. Applications were processed by school district staff, following federal guidelines, and then uploaded to the test data file. Type of school attending was collected from the district Title I application and enrollment records. Student attendance was collected for the Grade 3 year from the first day of school or first day of attendance through the end of the IREAD-3 assessment window. Attendance was taken two times per day at the school and entered by homeroom teachers and school staff. A student who was not in attendance for any reason was considered an absence in this study. A student who was in attendance during any part of the day, up to and including one-half was considered in attendance for one-half day. A student in attendance for any time beyond the half day but not for the full day was considered attending a full day.

For Research Question 3, DIBELS benchmark assessment data and TRC benchmark assessment data were collected from archival data files. Both assessments were administered by classroom teachers and other school staff during a three-week assessment window in August,

January, and May. DIBELS is a formative assessment with multiple subtests specifically designed to assess the core components of reading, including phonological awareness, alphabetic principal, accuracy and fluency, vocabulary, and comprehension. DIBELS assessments were based on curriculum-based measurements and are brief (one-minute for each subtest) indicators that have been found to be closely related to general student outcomes (University of Oregon Center on Teaching and Learning, 2008). The TRC assessment includes an oral running record of student accuracy of reading as well as an oral and written comprehension assessment on texts of graduated difficulty (Wireless Generation, 2012).

### **Data Analysis**

All data analyses were conducted using SPSS for Windows. Logistic regression was selected as the statistical analysis for Research Questions 1, 2, and 3. Logistic regression was selected in order to overcome the violation of assumption of linearity (Berry & Feldman, 1985; Field, 2009). Logistic regression is an appropriate analysis when the outcome variable is categorical (pass/fail on IREAD-3 assessment) and predictor variables are either continuous or categorical (Field, 2009). Assumptions including independence of errors and multicollinearity were tested as a part of the analysis. For Research Question 4, an independent-means *t*-test was utilized to compare the means of performance on the Grade 3 ISTEP+ assessment during years prior to administration of the IREAD-3 assessment and in 2012, when IREAD-3 was administered in addition to the Grade 3 ISTEP+ assessment. Assumptions including homogeneity of variance and independent scores were tested as a part of the analysis. Due to the sample size, the assumption of normality was assumed to be met.

### **Summary**

The sample for the study consisted of 1,719 Grade 3 students in attendance in the

Evansville Vanderburgh School Corporation during the 2011-12 school year who were assessed using the IREAD-3 in March of 2012. For Research Question 4, students who took ISTEP+ during the 2008-09, 2009-10, and 2010-11 school years were also included in the study.

IREAD-3 was a standardized achievement test administered to all students in the state of Indiana meant to determine those students who were able to read and comprehend 4th grade material.

For Research Question 3, DIBELS and TRC archival data were used. These formative assessments, administered by classroom teachers, measured performance on the core components of reading and has been found to closely correlate to general student outcomes. For Research Questions 1, 2, and 3 multivariate logistical regression was used to analyze the model. For Research Question 4, an independent-means *t*-test was used to analyze the means for each year of data collected.

## CHAPTER 4

### RESULTS

The purpose of this quantitative study was to determine if there are factors that are predictive of performance on IREAD-3 and to better understand if there were effects on Grade 3 ISTEP+ performance based on the implementation of IREAD-3. This chapter describes the data collected for this study, including an overview of the criteria used to select student samples, and the demographic characteristics of the students included. The statistical models used in the study are described in detail as well as the results for each of the four research hypothesis.

In this study, the first and second questions examined factors that that may be predictive of passing the IREAD-3 assessment. Independent variables including chronological age, ethnicity, SES, gender, type of school the student attended (Title I versus non-Title I), and attendance were analyzed to determine if they are predictive of performance on IREAD-3. The third question examined kindergarten, first-grade, and second-grade performance on DIBELS and TRC for each of the three assessment windows as independent variables to determine if either of these assessments predicted passing IREAD-3, and if so, at what point performance on IREAD-3 could be predicted. The fourth question examined whether there was a difference in ISTEP+ performance for Grade 3 students who also took the IREAD-3 assessment as compared to performance of Grade 3 students during the previous three years of ISTEP+ administration when those students did not take IREAD-3.

### **Selection and Student Samples**

The target population for this study focused on the Grade 3 students included in the initial year of Indiana's IREAD-3 assessment. The sample for the study was drawn from 1,719 Grade 3 students who were assessed with IREAD-3 during the 2011-2012 school year, all attending the same school district in southwestern Indiana. Students were representative of the community and surrounding county. Students who were English language learners as well as those students receiving special education services with the exception of students with severe cognitive or adaptive delays who were unable to attempt the test were included in the sample.

For Questions 1 and 2, students with less than 162 days enrolled in the school district were not considered in the data set. However, students who transferred schools within the district and whose days enrolled were equal to or greater than 162 days were included. Of the original 1,719 students, 1,615 students had complete data sets including at least 162 days in attendance. For Question 3, only students from the original 1,719 students who were assessed with IREAD-3 and who had DIBELS and TRC benchmark assessment data available for kindergarten, first and/or second grade were included in each of the data sets. For Question 4, students who were assessed with Grade 3 ISTEP+ during the 2008-2009, 2009-2010, or 2010-2011 school years from the same school district were also included. Student data and identified variables were imported into SPSS Version 17 and Version 21 for Windows. The SPSS statistical program was used to manipulate the data.

### **Student Demographic Information**

Demographic information for students included in the IREAD-3 sample is provided in Table 1. As the table illustrates, gender and race were categorical variables. The table shows

that boys outnumbered girls and that the majority of the students in the sample were identified as White, with Black being the next most frequently identified ethnicity.

Table 1

*Demographic Information for Grade 3 IREAD Students*

Variable	Frequency (N=1,719)	%
Gender		
Boys	894	52.07
Girls	821	47.93
Race		
American Indian	3	.17
Asian	20	1.16
Black	228	13.27
Hispanic	66	3.84
Multi-racial	184	10.71
White	1209	70.37
Hawaiian	1	.06
Missing Information	7	.41

Due to the limitations of the sample size and for the purposes of this study, ethnicity data were recoded to represent a dichotomous variable. Students were either coded as White or non-White (all other ethnicities).

**Student Age**

Student chronological age was coded based on the student's birthdate. Students in Indiana typically by state policy must be five years old on or before August 1 in order to enroll in kindergarten (Indiana Department of Education, 2010b). In order to create a dichotomous variable, students were coded as being youngest in the class if they were in the youngest quartile

or born after May 1, 2003. All other students who were born on or before May 1, 2003 were coded as oldest in the class. Table 2 contains descriptive data.

Table 2

*Student Age*

Variable	Frequency (N=1719)	%
Average age in class	1324	77.0
Youngest in class	391	22.8
Missing information	4	.2

**School and SES**

The school the student attended when the IREAD-3 assessment was administered was coded as either Title I served or non-Title I served. According to federal non-regulatory guidance, each local education agency (school district) is given the authority to determine schools within the district that are designated as Title I served based on rank order of percentage of poverty. Federal guidance directs that schools with 75% or more poverty must be served as Title I and schools between 35% and 74.99% poverty are served at the discretion of the school district (U.S. Department of Education, 2003). For this study, based on practice of the school district sample, schools served as Title I had a minimum of 75% poverty and were coded as Title I.

Students were coded as low-SES or average SES based on free, reduced, or paid lunch status. Students who were designated as qualifying for either free or reduced lunch status based

on federal guidelines were coded as low SES, with all others coded as average. Table 3 reflects descriptive data regarding school and SES status.

Table 3

*School Title I and Student SES Status*

	Variable	Frequency	%
School Title I Status	Title I Served	444	25.8
	Non-Title I Served	1275	74.2
SES	Low SES	639	37.2
	Average SES	1076	62.6
	Missing Information	4	.2

### **Data Analysis and Findings**

#### **Analysis of Questions 1 and 2**

Question One investigated whether passing IREAD-3 could be predicted by student chronological age (youngest students by quartile compared to peer group), ethnicity, SES of the family, gender, type of school attending (Title I versus non-Title I), or attendance. Question 2 explored the notion that if passing IREAD-3 can be predicted, what variables when considering student chronological age (youngest students by quartile compared to peer group), ethnicity, SES of the family, gender, type of school attending (Title I versus non-Title I), and attendance are most predictive. The null hypothesis for Question 1 and 2 assumed that passing IREAD-3 cannot be predicted by student chronological age (youngest students by quartile compared to peer group), ethnicity, SES of the family, gender, type of school attending (Title I versus non-Title I), or attendance. Logistic regression was selected as the statistical analysis for Research

Questions 1 and 2 in order to overcome the violation of assumption of linearity (Berry & Feldman, 1985; Field, 2009). Logistic regression is an appropriate analysis when the outcome variable is categorical (pass/fail on IREAD-3 assessment) and predictor variables are either continuous or categorical (Field, 2009). Table 4 summarizes the descriptive statistics and Table 5 summarizes the analysis results.

Table 4

*Descriptives of Independent Variables Predicting Not Passing IREAD-3*

Name	Mean	SD	Description of Categorical Variables
School Status (Title I)	.26	.44	0 = Title I 1 = Non-Title I
Gender	.52	.50	0 = Girls 1 = Boys
Age (Youngest in Class)	.23	.42	0 = youngest 1 = oldest
Ethnicity	.29	.45	0 = Minority 1 = White
SES Status	.62	.49	0 = Low SES 1 = Average SES
Attendance	3.01	4.58	NA

*Note.*  $N = 1,703$ ; IREAD-3 has categories: 0 = Not passing; 1 = Passing (80%)

Table 5

*Coefficients for the Logistic Regression Analysis Predicting Not Passing IREAD-3*

Name	<i>B</i> (Std. Error)	Wald	Exp( <i>B</i> )
School Status (Title I)	.24(.14)	2.86	1.27
Gender	.07(.13)	.30	1.07
Age (Youngest in Class)	.03(.16)	.04	1.03
Ethnicity	.62(1.3)	21.56**	1.86
SES Status	1.35(.18)	56.81**	3.87
Attendance	-.03(.01)	6.39*	.97

*Note.*  $N = 1703$ ;  $B$  is the unstandardized coefficient;  $\text{Exp}(B)$  is the factor change in odds for a unit increase in the Independent Variable; 0 = Not passing; 1 = Passing; \* $p < .05$ , \*\*  $p < .001$ , (two-tailed)

Low SES, non-White and attendance significantly predicted not passing IREAD-3,  $\beta = 1.35$ , Wald  $\chi^2(1) = 56.8$ ,  $p < .001$ ;  $\beta = .62$ , Wald  $\chi^2(1) = 21.56$ ,  $p < .001$ ; and  $\beta = -.03$ , Wald  $\chi^2(1) = 6.39$ ,  $p < .01$ , respectively. The model was significant because the model Chi square was significant,  $\chi^2(6, N = 1703) = 155.39$ ,  $p < .001$ . The Hosmer and Lemshew's test was not significant,  $\chi^2(8, N = 1703) = 6.08$ ,  $p = .638$ , indicating adequate model fit. Nagelkerke  $R^2 = .14$ . The percentage of correct classifications for the full model was 80%.

The null hypotheses for Questions 1 and 2 were rejected because the model, using the various predictors, was an adequate fit. Based on the logistic regression model, three identified variables (low SES, non-White, and poor attendance) significantly increased the odds of not passing IREAD-3.

### **Analysis of Question 3**

Question 3 investigated if passing IREAD-3 could be predicted by performance on DIBELS and TRC level in kindergarten, first grade, and second grade. The null hypothesis contended that passing IREAD-3 cannot be predicted by performance on DIBELS (Dynamic Indicators or Early Literacy Skills) and TRC (Text Reading Comprehension) level in kindergarten, first grade, and second grade. Logistic regression was selected as the statistical analysis for Research Question 3 in order to overcome the violation of assumption of linearity (Berry & Feldman, 1985; Field, 2009). Logistic regression is an appropriate analysis when the outcome variable is categorical (pass/fail on IREAD-3 assessment) and predictor variables are either continuous or categorical (Field, 2009).

Each variable was entered separately into the logistic regression model for predicting failure on IREAD-3. DIBELS data were recoded as below grade for students identified for support recommendations as strategic (needs additional intervention) or intensive (needs substantial intervention) and at grade for those identified as at grade. For TRC data, students identified as far below or below proficient were recoded as below grade and students identified as proficient or above proficient were recoded as at or above grade.

Table 6 summarizes the analysis results of the logistic regression models for DIBELS performance indicators below grade and at grade kindergarten through Grade 2.

Table 6

*Coefficients for the Logistic Regression Analysis Predicting Not Passing IREAD-3 Based on DIBELS*

Name		N	B (Std. Error)	Wald	Exp(B)
Grade 2	EOY (End of the Year)	1542	-3.01(.19)	240.47*	.05
	MOY (Middle of the Year)	1558	2.79(.16)	318.15*	16.23
	BOY (Beginning of the Year)	1524	2.81(.18)	239.84*	16.63
Grade 1	EOY (End of the Year)	1495	2.94(.16)	333.42*	19.00
	MOY (Middle of the Year)	1489	2.60(.15)	294.95*	13.35
	BOY (Beginning of the Year)	1467	1.97(.15)	186.27*	7.20
Kindergarten	EOY (End of the Year)	1390	1.79(.16)	120.73*	5.98
	MOY (Middle of the Year)	1322	1.71(.16)	121.22*	5.54
	BOY (Beginning of the Year)	1329	1.66(.17)	100.63*	5.27

*Note.* *B* is the unstandardized coefficient; *Exp(B)* is the factor change in odds for a unit increase in the Independent Variable; For IREAD-3, 0 = Not passing; 1 = Passing; For DIBELS, 1 = Below grade, 0 = Grade level, \*  $p < .001$ , (two-tailed).

A below grade performance indicator on DIBELS significantly predicted not passing IREAD-3; however, the most significant odds ratios were for MOY and BOY in Grade 2 and Grade 1 EOY,  $\beta = 2.79$ , Wald  $\chi^2(1) = 318.15$ ;  $\beta = 2.6$ , Wald  $\chi^2(1) = 294.95$ , and  $\beta = -2.94$ , Wald  $\chi^2(1) = 333.42$ , all at  $p < .001$ , respectively. The model was significant for all variables based on Chi square analysis. For Grade 2 MOY and BOY and Grade 1 EOY and the Chi square and Nagelkerke  $R^2$  were respectively,  $\chi^2(1, N = 1558) = 399.17$ , Nagelkerke  $R^2 = .36$ ;  $\chi^2(1, N = 1489) = 336.95$ , Nagelkerke  $R^2 = .32$ ; and  $\chi^2(1, N = 1495) = 419.08$ ,  $p < .001$ , Nagelkerke  $R^2 =$

.39,  $p < .001$ . The percentages of correct classifications for Grade 2 MOY and BOY and Grade 1 EOY were 80.5%, 80.1%, and 80.2%, respectively. Table 7 summarizes the analysis results of the logistic regression models for TRC performance indicators below grade and at Grades 1 and 2.

Table 7

*Coefficients for the Logistic Regression Analysis Predicting Not Passing IREAD-3 Based on TRC*

Name	<i>N</i>	<i>B</i> (Std. Error)	Wald	Exp( <i>B</i> )
Grade 2 EOY (End of the Year)	1537	-2.55(.15)	298.25*	.08
MOY (Middle of the Year)	1542	-2.24(.15)	224.56*	.11
BOY (Beginning of the Year)	1514	-2.56(.19)	189.38*	.08
Grade 1 EOY (End of the Year)	1496	-2.79(.19)	209.74*	.06
MOY (Middle of the Year)	1488	-2.40(.15)	251.59*	.09
BOY (Beginning of the Year)	1241	2.06(.23)	82.74*	.13

*Note.* *B* is the unstandardized coefficient; Exp(*B*) is the factor change in odds for a unit increase in the Independent Variable; For IREAD-3, 0 = Not passing; 1 = Passing; For TRC, 1 = Below grade, 0 = Grade level, \*  $p < .001$ , (two-tailed).

A below grade performance indicator on TRC significantly predicted not passing IREAD-3; however, the most significant odds ratios were for EOY and MOY in Grade 2 and EOY and MOY in Grade 1,  $\beta = -2.55$ , Wald  $\chi^2(1) = 298.25$ ;  $\beta = -2.24$ , Wald  $\chi^2(1) = 224.56$ ;  $\beta = -2.79$ , Wald  $\chi^2(1) = 209.74$ ;  $\beta = -2.4$ , Wald  $\chi^2(1) = 251.59$ , all at  $p < .001$ , respectively. The model was significant for all variables based on Chi square analysis. For Grade 2 EOY and MOY and Grade 1 EOY and MOY the Chi square and Nagelkerke  $R^2$  were respectively,  $\chi^2(1, N$

= 1537) = 343.13, Nagelkerke  $R^2 = .32$ ;  $\chi^2(1, N = 1542) = 269.05$ , Nagelkerke  $R^2 = .25$ ;  $\chi^2(1, N = 1496) = 328.62$ , Nagelkerke  $R^2 = .31$ ; and  $\chi^2(1, N = 1488) = 299.88$ ,  $p < .001$ , Nagelkerke  $R^2 = .29$ . The percentages of correct classifications for Grade 2 EOY and MOY and Grade 1 EOY and MOY were 79.8%, 80.2%, 80.1%, and 80.2%, respectively.

The null hypothesis for Question 3 was rejected because the model, using the various predictors was an adequate fit. Based on the logistic regression model, below grade performance on both DIBELS and TRC significantly increased the odds of not passing IREAD-3.

#### **Analysis of Question 4**

Question 4 examined if there was a significant difference between average ISTEP scores in Grade 3 prior to the implementation of IREAD-3 and average ISTEP scores in Grade 3 after implementation of IREAD-3. Conversely, the null hypothesis for Question 4 maintained that there is no significant difference between average ISTEP scores in Grade 3 before implementation of IREAD-3 and after implementation of IREAD-3. Initially, an independent-means  $t$ -test was conducted to compare the ISTEP+ performance of Grade 3 students prior to implementation of IREAD-3 (2009, 2010, and 2011 combined) and following implementation of IREAD-3 (2012). Descriptive statistics for the  $t$ -test analysis is presented in Table 8.

Table 8

*Descriptive Statistics for ISTEP+ Scores (2009, 2010, and 2011, Combined) and ISTEP+ Scores 2012*

ISTEP+ Years	$N$	$M$	$SD$
2009, 2010, 2011	5024	449.96	64.57
2012	1624	456.96	59.64

There was a significant difference in the scores for 2009, 2010, and 2011 combined ( $M = 449.96$ ,  $SD = 64.57$ ) and 2012 ( $M = 456.96$ ,  $SD = 59.64$ ),  $t(2948.91) = -4.03$ ,  $p < .001$ . The null hypothesis was rejected, and in the  $t$ -test analysis, there was a significant difference between average ISTEP scores in Grade 3 before implementation of IREAD-3 and after implementation of IREAD-3.

Because there were three years of test data combined (2009, 2010 and 2011), a year-to-year analysis was needed to determine the extent of difference in scores. In order to conduct a pairwise comparison, an ANOVA test was selected for more detailed analysis of multiple years of data. Four years of ISTEP+ scores were analyzed including the students who took both IREAD-3 and ISTEP+ in Grade 3 during the 2011-12 school year (1,624 students) as well as students who took ISTEP+ in Grade 3 during the 2008-09 (1,612 students), 2009-10 (1,748 students), and 2010-11(1,664 students) school years. Descriptive statistics for ISTEP+ performance for Grade 3 for these four years are presented in Table 9. The mean score increased for each of the four years from 2009 until 2012.

Table 9

*Descriptive Statistics for the Third Grade ISTEP+ Scores 2009-2012*

ISTEP+ Years	<i>N</i>	<i>M</i>	<i>SD</i>
2009	1612	445.37	62.645
2010	1748	449.55	68.855
2011	1664	454.83	61.362
2012	1624	456.67	59.641
Total	6648	451.67	63.466

A one-way ANOVA test was conducted to compare the effect of IREAD-3 on Grade 3 ISTEP+ scores. The ISTEP+ scores for Grade 3 differed significantly across the four years,  $F(3, 6644) = 11.13, p < .001, \omega^2 = .000$ . An Omega squared analysis indicated no measurable effect size.

Because the null was rejected, a post-hoc comparison using the Tukey honestly significantly difference test was conducted to determine exactly which mean differences were significant. The Tukey honestly significantly difference test indicated that there were no significant differences between the means of 2009 ( $M = 445.37$ ) and 2010 ( $M = 449.55$ ),  $p = .222$ ; 2010 ( $M = 449.55$ ) and 2011 ( $M = 454.83$ ),  $p = .071$ ; or 2011 ( $M = 454.83$ ) and 2012 ( $M = 456.96$ ),  $p = .770$ . However, there were significant differences between the means of 2009 ( $M = 445.37$ ) and 2011 ( $M = 454.83$ ) and 2009 ( $M = 445.37$ ) and 2012 ( $M = 456.96$ ), both at  $p < .001$ , and a significant difference between the means of 2010 ( $M = 449.55$ ) and 2012 ( $M = 456.67$ ),  $p = .004$ . These results suggest that although there was a statistically significant difference in scores over the four years, the effect size was insignificant.

### Summary

This chapter described the data collected for this study, including an overview of the criteria used to select student samples and the demographic characteristics of the students included. The statistical models used in the study are described in detail as well as the results for each of the four research hypotheses. Chapter 5 discusses implications of the results of the statistical analysis of data, including causes that may be related to the findings, and offers recommendations for future research and practice.

## CHAPTER FIVE

### SUMMARY

This study examined the implementation of Public Law 109 and IREAD-3 as well as initial results of IREAD-3 in Indiana for one school district in Indiana. The interpretation of Indiana Public Law 109 and subsequent policy adopted by the Indiana Department of Education resulted in the Indiana State Department of Education mandating circumstances implemented during the 2011-2012 school year regarding grade level retention of Grade 3 students. Chapter 4 presented the results of the statistical analysis used in the study. This chapter presents a brief review of the pertinent literature, summarizes the findings of the study, and discusses the results of the study. The chapter concludes with implications for future research, practice, and policy makers.

### **Conclusions**

#### **Research Questions 1 and 2**

1. Can passing IREAD-3 be predicted by student chronological age (age compared to peer group by quartiles), ethnicity, SES of the family, gender, type of school attending (Title I versus non-Title I) or attendance?
2. If passing IREAD-3 can be predicted, what variables when considering student chronological age (age compared to peer group by quartiles), ethnicity, SES of the

family, gender, type of school attending (Title I versus non-Title I), and attendance are most predictive?

### **Research Hypotheses (Null) for Questions 1 and 2**

Passing IREAD-3 cannot be predicted by student chronological age (age compared to peer group by quartiles), ethnicity, socio-economic status of the family, gender, type of school attending (Title I versus non-Title I), or attendance.

The null hypotheses for Question 1 and 2 were rejected because the model, using the various predictors, was an adequate fit. Based on analysis of the Wald statistic, low SES and non-White were found to significantly increase the odds of not passing IREAD-3,  $p < .001$ . Poor attendance also significantly increased the odds of not passing IREAD-3,  $p < .05$ . This finding was consistent with the literature, which identified an achievement gap that existed between students of poverty and middle class and among Black, Hispanic, and White students on the basis of those who were retained in a grade (Alexander et al., 2003; Hauser et al., 2004; Hauser et al., 2007; Lorence et al., 2002; Planty et al., 2009; Shepard, 2004). According to the U.S. Department of Education (2007), 16% of Black, 11% of Hispanic, and 8% of White students reported having been retained. Black male students were more than two times as likely to be retained as compared to White or Hispanic male students, and students of poverty were over four times as likely to be retained as students from non-poverty families (U.S. Department of Education, 2007).

Lorence et al. (2002) reviewed the entire population of Texas students over a nine-year period and found that being African American increased the odds of Grade 3 retention based on performance on standardized testing by 62% when compared to non-Hispanic White students. The odds of retaining Hispanic students were increased by 26% when compared with

academically challenged non-Hispanic White students. In a Rand report analyzing the implementation of New York City School implementation of retention policy from 2004-2007, McCombs et al. (2009) found that from the Grade 3 cohort of 2004-05, 15% of Hispanic students and 21% of Black students were at risk of repeating the grade as compared to 3% of Asian or White students.

Attendance was identified as being significant within the model, after poverty and ethnicity status. Attendance was not a factor strongly associated with retention in the literature, although it was identified as a consideration in several retention policies reviewed including the Chicago Public Schools policy which considered several factors regarding Grade 3 promotion including scoring at a minimum of the 24th percentile on a district-wide assessment of reading, a C or higher report card grades, no more than nine unexcused absences per year and attendance at a summer school program (Chicago Public School System, 2011).

As noted, students of poverty, as identified by free or reduced lunch status, were less likely to pass IREAD-3; however, students who attended a school with a high percentage of students on free or reduced lunch was not a significant factor in the model. Students attending schools identified by the school district as Title I (a poverty concentration of 75% or more) did not appear to perform significantly different on IREAD-3 than students attending non-Title I schools.

The gender of students was not noted in the model as a significant factor but was identified in the literature as being related to retention. Boys as compared to girls were more likely to be retained among all grade levels (Hauser et al., 2004; Lorence et al., 2002; Planty et al., 2009; Shepard, 2004; Xia & Kirby, 2009). As noted by Entwisle et al. (2007), traditional models of retention rely on teacher recommendation, including grades, and conduct marks which

are more subjective in nature. Since the IREAD-3 assessment is less subjective in nature, the results may be a more balanced identification of genders than found in the literature regarding retention.

### **Research Question 3**

Can passing IREAD-3 be predicted by performance on DIBELS and TRC levels in kindergarten, first grade, and second grade?

### **Research Hypothesis (Null) for Question 3**

Passing IREAD-3 cannot be predicted by performance on DIBELS (Dynamic Indicators or Early Literacy Skills) and TRC (Text Reading Comprehension) level in kindergarten, first grade, and second grade.

The logistic regression analysis revealed that a below grade performance indicator on DIBELS significantly predicted not passing IREAD-3; an analysis of the Wald statistic revealed that the most significant results were for EOY in Grade 2 and Grade 1 MOY and EOY,  $p < .001$ . Analysis of TRC scores revealed that a below grade performance indicator on TRC significantly predicted not passing IREAD-3; however, the most significant odds ratios were for EOY and MOY in Grade 2 EOY and MOY in Grade 1. All results were significant at  $p < .001$ .

The results of the study rejected the null hypothesis and indicated that passing IREAD-3 can be predicted by performance on DIBELS and TRC in kindergarten, first grade, and second grade. Based on the logistic regression model, below-grade performance on both DIBELS and TRC significantly increased the odds of not passing IREAD-3. These results provide preliminary evidence that early reading assessments, specifically DIBELS and TRC, are predictive of later performance in reading on standardized assessments, specifically IREAD-3.

This finding is in alignment with the literature. The majority of promotion policies consider reading attainment as a critical factor (Xia & Kirby, 2009). The IREAD-3 assessment for Indiana assessed students on three of five areas identified as critical to reading by the National Reading Panel (National Institute of Child Health and Human Development, 2000): phonics, vocabulary, and comprehension (Indiana Department of Education, 2010a). Wireless Generation DIBELS 3-D, which included DIBELS as well as TRC assessments, was identified as a research-based screening instrument that has correlation with success on standardized assessments in later grades. In a study by Roehrig et al. (2008), DIBELS PRF assessment was found to be significantly correlated with the Florida Comprehensive Assessment Test-Sunshine State Standards and the AT-10 for Grade 3 students in Florida Reading First schools when both assessments were administered during the same school year ( $r = .70$ ). There was also a statistically significant correlation with reading skills when the oral reading assessment was administered within the same school year as the standardized reading assessment. The TRC assessment included an oral running record of student reading accuracy as well as an oral and written comprehension assessment on texts of graduated difficulty (Wireless Generation, 2012). A measure of fluency alone (accuracy and speed of reading) was found to be not as predictive of comprehension especially as students developed as readers (Hosp & Fuchs, 2005; Samuels, 2006).

#### **Research Question 4**

Is there a significant difference between average ISTEP scores in Grade 3 prior to the implementation of IREAD-3 and average ISTEP scores in Grade 3 after implementation of IREAD-3?

**Research Hypothesis (Null) for Question 4**

There is no significant difference between average ISTEP scores in Grade 3 before implementation of IREAD-3 and after implementation of IREAD-3.

The initial analysis using an independent-means *t*-test rejected the null hypothesis and found that there was a significant difference in the scores for 2009, 2010, and 2011 combined,  $p < .001$ . Using a one-way ANOVA test revealed that ISTEP+ scores for Grade 3 differed significantly across the four years; however, a Tukey honestly significantly difference test indicated that there was no significant difference between the means of 2009 and 2011 as well as between 2009 and 2012, both at  $p < .001$ , and a significant difference existed between the means of 2010 and 2012,  $p = .004$ . These results suggest that although there was a statistically significant difference in scores over the four years, the effect size was insignificant. Practically, the difference appeared to represent an upward trend of scores and the statistically significant differences were not necessarily associated with implementation of IREAD-3 in 2012.

**Limitations of the Study**

This study reviewed data from a sample of 1,720 students, all attending the same countywide school district in southwestern Indiana during the same school year. The students represented a wide range of SES, were primarily White, Black, Hispanic, or bi-racial. Due to the limited number of students in the sample representing various ethnicities, all non-White students were combined in the study to represent one category. Findings may not be readily generalizable to all races/ethnicities. The sample included students who attended schools within city limits, suburban areas, and some rural areas. Due to the nature of the subjects, the study may not be readily generalized to other demographic areas.

In addition, the study was specific to the assessment and policy in Indiana and results may not be generalized to other states. The study assumed that the IREAD-3 assessment was a valid and reliable assessment. Further, it assumed that students received similar educational experiences. An additional limitation is that the study only reviewed one year of data, during the first year of administration of the IREAD-3 assessment.

### **Implications**

This study adds to the body of knowledge regarding minimum competency promotion policies and informs educators and others interested in the debate regarding social promotion and retention. Based on research and findings of this study, there are several implications appropriate for researchers, policy makers, and educators to consider.

#### **Implications for Practice**

One goal of this study was to provide relevant information to Indiana educators faced with implementing IREAD-3 policy. School leaders have been confronted with the responsibility of responding to the IREAD-3 policy in order to best meet the needs of their students and ensure successful outcomes. These findings, supported by literature, suggested that students of poverty continue to be at risk of not performing as well on standardized assessments as peers of average SES. In addition, these findings supported that minority students have a greater odds of not passing IREAD-3. A study by Hauser, Pager, and Simmons (2004) analyzed data from the current population survey of the U.S. Bureau of the Census and found that the differential in the higher percentage of minority students retained in a grade was not significant when other social and economic factors were controlled; race differentials were almost entirely due to group differences in SES. Decades of research have supported an association between

low academic achievement as measured by standardized assessments and children raised in circumstances associated with low SES (Bradley & Corwyn, 2002).

School and civic leaders, when reviewing these findings, may consider how to overcome disadvantages faced by children of poverty and focus attention and resources on research-based strategies, including high-quality, academically-oriented, full-day preschool and kindergarten, family literacy programs, tutoring and summer reading programs, reading specialists, and programs to improve attendance. Hecht, Burgess, Torgensen, Wagner, and Rashotte (2000) found that beginning kindergarten knowledge of print accounted for most SES-related variance in growth of reading skills. Based on their findings, they recommended that students at risk of reading failure receive intensive interventions in pre-kindergarten and kindergarten focused on training in print knowledge, phonological awareness, and rate of access. This research suggests schools may consider a pre-kindergarten, kindergarten, and early primary grade model that ensures screening students early for possible delays in acquisition of basic reading skills and offers intensive, early interventions for those students at risk.

Although many studies have indicated an achievement gap for students representing minority groups, the results of this study make it difficult to determine the effect of low-SES versus the effect of minority status. Based on these findings and the body of research regarding the academic achievement of minority students, educational leaders in Indiana may consider that minority students, particularly Black and Hispanic students, may need different instructional styles, may experience differences in teacher expectations, and may be more negatively influenced by peer groups as well as by social and cultural expectations. Students of minority groups often are negatively influenced by a culturally biased curriculum and pedagogy not responsive to cultural differences (Hallinan, 2004). Based on the results of this study,

educational leaders should be mindful of these factors particularly as they relate to reading instruction for young children. A further implication of these findings was that the IREAD-3 assessment may contain content considered culturally biased for some student groups.

Given the predictive nature of early reading screening assessments (DIBELS and TRC) found in this study, couple with the research from the literature review, educators may be considerate of when in a student's educational career retention may be most appropriate. Alexander et al. (2003) found that retention in the early grades appeared to have positive results, particularly for students who were retained at the end of first grade; however, the positive results diminished over time. Students who are identified early, who are provided research-based, intensive interventions, and who continued to exhibit below grade-level reading skills, may benefit from retention at the end of Grade 1 rather than waiting until the end of Grade 3. If the retention is to have any benefit, from the assertion that a child needs more time to master academic content, the earlier the retention in the child's educational career the more likely it may have some positive effects (at least in the short term) and may minimize any negative effects.

### **Implications for Public Policy**

Although the policy of minimum competency reading assessment in Grade 3 has been implemented in Texas, Florida, Ohio, and Chicago Public Schools, the IREAD-3 assessment was a unique assessment and new to Indiana. No research was conducted to determine the impact of this particular assessment. Based on the findings of this study, Indiana policy makers may be concerned about the fair and appropriate use of standardized tests as it pertains to IREAD-3 implementation. Penfield (2010) found that there was particular concern regarding use of standardized tests as a tool for determining retention and, in his opinion, violated fair and

appropriate use of assessments. Penfield suggested that the test results should not be attributed to poor instruction or linguistic or cultural content and that because marginalized groups have consistently performed poorly on such assessments, there is a potential violation. The results of this study confirm these findings. Given the findings regarding the performance of minority students on IREAD-3, there may be reason to suspect some items may be culturally biased.

In addition, Penfield (2010) also brought to question whether testing led to consequences that are educationally beneficial. Based on the conflicting body of research regarding the effects of retention, the practice of retaining students in a grade solely on the results of one standardized test may be questionable use of testing. The IREAD-3 provided the only data utilized to determine whether a child is retained in a grade, other than the three situations for approved waivers. This may violate the use of standardized tests and could be considered a questionable practice.

Some states and districts including Chicago Public Schools, Texas, and Florida changed practice regarding the use of gateway assessments to include additional factors considered in order to determine promotion. Additional data included report card grades, attendance, attendance at a summer school reading program, student work portfolios, and teacher observations (Chicago Public School System, 2011; New York City Department of Education, 2009; Xia & Kirby, 2009). The Grade 3 gateway assessment in Texas was recently retracted by the Texas General Assembly; however, the assessment-based promotion policy for Grades 5 and 8 was continued (Texas Education Agency, 2011).

Considering the positive relationship between students who have been retained with an increased risk of dropping out of school (Alexander et al., 2004; Entwisle et al., 2007; Roderick, 1995), retention based on IREAD-3 may put student groups already at risk of dropping out of

school in further jeopardy. Alexander et al. (2004) found that students who were retained in later grades (Grades 4 to 7) were at significantly greater risk of dropping out of school than if retained in primary grades. Students can be waived from retention based on failing IREAD-3 if they have already been retained two times as per Indiana Department of Education policy. In *Taking Responsibility for Ending Social Promotion* (U.S. Department of Education, 1999), it was noted that being “held back” twice made it a virtual certainty that a student will drop out of school. Policy makers may consider the long-term educational and economic consequences of students being retained in Grade 3 and later dropping out of school.

If the intent of IREAD-3 implementation was to ensure students are reading proficiently prior to entering the middle grades, policy-makers may consider funding several strategies to improve reading achievement of students of poverty including high quality, academically oriented pre-kindergarten programs, family literacy programs, reading intervention programs, summer school and tutoring programs, and reading specialists positions. In order to truly overcome persistent gaps in achievement, additional funding and support for exploration of ungraded school experiences, offering students multi-aged, ungraded classes with additional, time and support in the absence of stigma, may provide a long-term solution (Alexander et al., 2004).

### **Implications for Further or Related Research**

The results of this study suggest several opportunities for related research. Further research analyzing the effects of gender within ethnic groups would help to determine if gender is more significant within groups rather than as a variable applied to the entire sample and help confirm or refute earlier research claims that boys are more frequently retained.

This study of IREAD-3 revealed the type of school attended (Title I served versus non-Title I) did not have a significant impact on the odds of passing IREAD-3. Although low SES was found to be a significant variable, subgroups could be further defined by comparing students of lower SES who qualify for the federal free lunch program with students who qualify for reduced lunch status, to determine if there is a difference in passing IREAD-3, based on status of poverty. Given the findings regarding the performance of minority students on IREAD-3, a thorough review of the IREAD-3 assessment should be conducted to determine if there are items that would be suspect of being culturally biased.

Two additional recommendations for further research address limitations of sample size and replication of results and practical application of data. This study was limited in that it analyzed a sample from one school district in Indiana. A more thorough analysis of students across the state of Indiana would confirm if results can be replicated with a larger, more representative sample. In order to practically utilize the finding of poor attendance significantly increasing the odds of not passing IREAD-3, a further study is recommended that determines specific ranges of days absent. For this study, the attendance variable was analyzed as a continuous variable and attendance data would need to be coded in ranges for practical purposes.

An additional study adding to the debate regarding retention could compare those Indiana students who scored just above the passing scaled score for IREAD-3 and who were promoted with their age appropriate peers with those students who scored just below the scaled score for passing and were retained. This further analysis would provide insight to the effects of retention. In addition, a similar long-term analysis comparing multiple years of data would be of interest and add to the body of literature regarding retention. Comparing the number of students who

drop out of school from the cohort of students retained because of IREAD-3 with students of similar academic performance would also be of interest.

### **Conclusion**

This study examined the implementation of Public Law 109 and IREAD-3 as well as initial results of IREAD-3 in Indiana for one school district in Indiana. The interpretation of Public Law 109 and subsequent policy adopted by the Indiana Department of Education resulted in the Indiana State Board of Education mandating circumstances implemented during the 2011-2012 school year regarding grade level retention of third grade students. Retention represents an expensive intervention for struggling students representing a minimum of an additional year of per pupil expenditure of over \$11,000 (Indiana Youth Institute, 2011). To make prudent decisions, it is both fiscally responsible and in the best interest of students that educators have relevant information regarding IREAD-3. Educators need to be aware of what practices have proven successful, how to identify students at risk of failure early, and how to minimize any potentially harmful results of retention. The purpose of this quantitative study was to determine if there were factors that are predictive of performance on IREAD-3 and to better understand if there were effects on Grade 3 ISTEP+ performance based on the implementation of IREAD-3.

In this study, the first and second questions examined factors that that may be predictive of passing the IREAD-3 assessment. Independent variables including chronological age, ethnicity, socioeconomic status, gender, type of school the student attended (Title I versus non-Title I), and attendance were analyzed to determine if they were predictive of performance on IREAD-3. A logistic regression model was constructed to analyze the variables. Based on the model, three identified variables (low SES, non-White, and poor attendance) significantly increased the odds of not passing IREAD-3.

The third question examined kindergarten, first-grade, and second-grade performance on DIBELS and TRC for each of the three assessment windows as independent variables to determine if either of these assessments predicted passing IREAD-3, and if so, at what point could performance on IREAD-3 be predicted. Based on the logistic regression model, below-grade performance on both DIBELS and TRC (independently) significantly increased the odds of not passing IREAD-3. The statistically significant odds of not passing IREAD-3 was noted as early as beginning of the kindergarten year but were noted to be more significant in later years, the middle and end of Grade 1 and beginning and middle of Grade 2.

The fourth question examined whether there was a difference in ISTEP+ performance for Grade 3 students who also took the IREAD-3 assessment as compared to performance of Grade 3 students during the previous three years of ISTEP+ administration when those students did not take IREAD-3. The initial analysis used an independent-means *t*-test and found there was a significant difference in the ISTEP+ scores for 2009, 2010, and 2011 combined. These results suggest that although there was a statistically significant difference in scores over the four years, the effect size was insignificant. Practically, the difference appears to represent an upward trend of scores and the statistically significant differences were not necessarily associated with implementation of IREAD-3 in 2012.

These findings were consistent with the body of literature regarding retention and the use of high-stakes assessment as a gateway to promotion. Based on the findings of this study, Indiana policymakers may be concerned about the fair and appropriate use of standardized tests as it pertains to IREAD-3 implementation. In addition, based on a review of the literature, students who are retained are at a much higher risk of dropping out of school. Policymakers may

consider the long-term educational and economic consequences of students being retained in Grade 3 and later dropping out of school.

If the intent of IREAD-3 implementation was to ensure students read proficiently prior to entering Grade 4, policymakers and educators may consider leveraging resources to provide strategies that will improve the reading achievement of students of poverty. These strategies include providing high quality, academically oriented pre-kindergarten programs, providing family literacy programs, providing reading intervention programs, providing summer school opportunities, providing tutoring programs, and hiring reading specialist positions. Ungraded school experiences, offering students multi-aged, ungraded classes with additional, time and support in the absence of stigma may also provide a long term solution to retention. Jimerson (2004) contended that educators, politicians, and parents must view providing support to children who come to school unprepared as their responsibility. As Jimerson (2004) stated, “When a student fails, adults have failed these children” (p. 87).

## REFERENCES

- Alexander, K. L., Entwisle, D. R., & Dauber, S. L. (2003). *On the success of failure: A reassessment of the effects of retention in the primary school grades*. Cambridge, MA: Cambridge University Press.
- Alexander, K. L., Entwisle, D. R., Dauber, S. L., & Kabbani, N. (2004). Dropout in relation to grade retention: An accounting from the beginning school study. In H. J. Walberg, A. J. Reynolds, & M. C. Wang (Eds.), *Can unlike students learn together? Grade retention, tracking, and grouping* (pp. 183-202). Greenwich, CT: Information Age Publishing.
- Allen, C. S., Chen, Q., Willson, V. L., & Hughes, J. N. (2009). Quality of research design moderates effects of grade retention on achievement: A meta-analytic, multilevel analysis. *Educational Evaluation and Policy Analysis, 31*, 480-499. Retrieved from ERIC database. (EJ866929)
- 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 ERIC database. (EJ817303)
- Berry, W. D., & Feldman, S. (1985). *Multiple regression in practice*. Thousand Oaks, CA: Sage.
- Bowman, L. J. (2005). Grade retention: Is it a help or hindrance to student academic success? *Preventing School Failure, (49)*3, 42-46. Retrieved from ERIC database. (EJ744733)

- Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53*, 371-99.
- Bramlett, R. K., Murphy, J. J., Johnson, J., Wallingsford, L., & Hall, J. D. (2002). Contemporary practices in school psychology: A national survey of roles and referral problems. *Psychology in the Schools, 39*, 327-335. Retrieved from ERIC database. (EJ646839)
- Bush, G. W. (1999). Remarks by the president at the National Education Summit. *White House Education Press Releases and Statements*. Washington, DC: Government Printing Office.
- California Department of Education. (2011). *FAQ pupil promotion and retention*. Sacramento, CA: Author. Retrieved from <http://www.cde.ca.gov/re/lr/pr/faqppr.asp>
- Caples, B. L. (2005). *Grade level retention: A handbook for educators*. Fresno, CA: Educational Leadership Institute.
- Chicago Public School System (2011). *School/parent guide to the elementary school promotion policy benchmark grades 3 & 6* (Board Policy 09-1028-PO2). Retrieved from [http://www.cps.edu/SiteCollectionDocuments/PromotionPolicy/ParentGuide/Grades3\\_6\\_English.pdf](http://www.cps.edu/SiteCollectionDocuments/PromotionPolicy/ParentGuide/Grades3_6_English.pdf)
- Christensen, B. L. (2010). *The effect of grade-level retention on student success as defined by the Student Success Initiative of Texas* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3417493)
- Clinton, W. J. (1998). Memorandum on helping schools end social promotions. *Weekly Compilation of Presidential Documents, 34*, 310-313. Retrieved from EBSCOhost database.
- Colorado Early Literacy Bill. S. 12-0004, H.R. 12-1238, 68th General Assembly of the State of Colorado, (2012).

- Common Core State Standard Initiative. (2011). *About the standards*. Retrieved from <http://www.corestandards.org/about-the-standards>
- Council of Chief State School Officers. (2003). *Annual survey of state student assessment programs*. Retrieved from [http://nces.ed.gov/programs/digest/d10/tables/dt10\\_178.asp](http://nces.ed.gov/programs/digest/d10/tables/dt10_178.asp)
- Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2005). First grade and educational attainment by age 22: A new story. *American Journal of Sociology, 110*, 1458-1502.
- Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2007). Early schooling: The handicap of being poor and male. *Sociology of Education, 80*(2), 114-138. Retrieved from ERIC database. (EJ771521)
- Field, A. (2009). *Discovering statistics using SPSS*. London, England: Sage.
- Florida Department of Education. (2009). *Non-promotion in Florida's public schools, 2007-08* (Data Report Series 2009-11D). Tallahassee, FL: Education Information & Accountability Services. Retrieved from <http://www.fldoe.org>
- Florida Legislature, Office of Program Policy Analysis and Government Accountability. (2006). *Third-grade retention policy leading to better student performance statewide* (Report No. 06-66). Tallahassee, FL: Author. Retrieved from <http://www.oppaga.state.fl.us/reports/pdf/0666rpt.pdf>
- Frederick, C. B., & Hauser, R. M. (2008). Have we put an end to social promotion? Changes in school progress among children aged 6 to 17 from 1972 to 2005. *Demography, 45*, 719-740. doi: 10.1353/dem.0.0015
- Goals 2000: Educate America Act. (1993). H.R. 1804--103rd Congress. Retrieved from <http://www.govtrack.us/congress/bills/103/hr1804>

- Good, R.A., Kaminski, R. A., & Dill, S. (2007). *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: University of Oregon.
- Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high stakes outcomes. *Scientific Studies of Reading, 5*, 257-288. Retrieved from ERIC database. (EJ640703)
- Hallinon, M. T. (2004). Race-effects on ability group outcomes. In H. J. Walberg, A. J. Reynolds, & M. C. Wang (Eds.), *Can unlike students learn together? Grade retention, tracing, and grouping* (pp. 97-114), Greenwich, CT: Information Age.
- Hauser, R. M., Frederick, C. B., & Andrew, M. (2007). Grade retention in the age of standards-based reform. In A. Gamoran (Ed.), *Standards-based reform and the poverty gap: Lessons for "No Child Left Behind"* (pp. 120-153). Washington, DC: Brookings Institute Press.
- Hauser, R. M., Pager, D. I., & Simmons, S. J. (2004). Race-ethnicity, social background, and grade retention. In H. J. Walberg, A. J. Reynolds, & M. C. Wang (Eds.), *Can unlike students learn together? Grade retention, tracking, and grouping* (pp. 97-114). Greenwich, Ct: Information Age.
- Hecht, S. A., Burgess, S. R., Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2000). Explaining social Class differences in growth of reading skills form beginning kindergarten through fourth-grade: The role of phonological awareness, rate of access, and print knowledge. *Reading and Writing: An Interdisciplinary Journal, 12*, 99-127.

- Heubert, J. P., & Hauser, R. M. (1999). *High stakes: Testing for tracking, promotion, and graduation* (Report of the Committee on Appropriate Test Use). Washington, DC: National Academy Press. Retrieved from ERIC database. (ED439151)
- Hosp, M. K., & Fuchs, L. S. (2005). Using CBM as an indicator of decoding, word reading, and comprehension: Do the relations change with grade? *School Psychology Review*, 34(1), 9-26. Retrieved from ERIC database. (EJ683517)
- House Enrolled Act 1367, Ind. Code 5-10-8-6.7. §§ 2-3 (2010).
- Indiana Department of Education. (2010a). *Common questions about the Indiana reading evaluation and determination assessment*. Retrieved from [http://www.doe.in.gov/assessment/docs/IREAD\\_Common\\_Questions.pdf](http://www.doe.in.gov/assessment/docs/IREAD_Common_Questions.pdf)
- Indiana Department of Education. (2010b). *Frequently asked questions about kindergarten programs*. Retrieved from <http://www.doe.in.gov/sites/default/files/curriculum/faqfdk1.pdf>
- Indiana Department of Education. (2011a). *2012 & 2013 School formula simulation*. Retrieved from [http://www.in.gov/legislative/house\\_republicans/pdfs/sf\\_26.pdf](http://www.in.gov/legislative/house_republicans/pdfs/sf_26.pdf)
- Indiana Department of Education. (2011b). *2011 State of Indiana, enrollment by grade*. Retrieved from <http://compass.doe.in.gov/dashboard/enrollment.aspx?type=state>
- Indiana Department of Education. (2011c). *2011-2012 ISTEP+ program manual*. Retrieved from <http://www.doe.in.gov/sites/default/files/assessment/2011-12-istep-program-manual2-23-12.pdf>
- Indiana Youth Institute. (2011). *Kids count data center: Total per pupil expenditure (currency) - 2008*. Retrieved from <http://datacenter.kidscount.org/data/bystate/Rankings.aspx?state=IN&ind=1161>

- Jackson, G. B. (1975). The research evidence on the effects of grade retention. *Review of Educational Research, 45*, 613-635. Retrieved from ERIC database. (EJ135378)
- Jacob, B. A., & Lefgren, L. (2002). *Remediation education and student achievement: A regression-discontinuity analysis*. (NBER Working Paper No. 8918). Cambridge, MA: National Bureau of Economic Research. Retrieved from [http://sitemaker.umich.edu/bajacob/files/restat\\_2003.pdf](http://sitemaker.umich.edu/bajacob/files/restat_2003.pdf)
- Jacob, R. T., Stone, S., & Roderick, M. (2004). *Ending social promotion: The response of teachers and students*. Chicago, IL: Consortium on Chicago School Research.
- Jimerson, S. R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review, 30*, 420-437. Retrieved from ERIC database. (EJ667518)
- Jimerson, S. R. (2004). Is grade retention educational malpractice? In H. J. Walberg, A. J. Reynolds, & M. C. Wang (Eds.), *Can unlike students learn together? Grade retention, tracking, and grouping* (pp. 71-95). Greenwich, CT: Information Age Publishing.
- Jimerson, S. R., & Kaufman, A. M. (2003). Reading, writing, and retention: A primer on grade level retention. *Reading Teacher, 56*, 622- 635. Retrieved from ERIC database. (EJ664327)
- Johnson, J. R. (1984). Synthesis of research on grade retention and social promotion. *Educational Leadership, 41*(8), 66-68. Retrieved from ERIC database. (EJ299543)
- Linn, R. L. (2000). Assessments and accountability. *Educational Researcher, 29*(2), 4-16. Retrieved from ERIC database. (EJ602768)

- Lorence, J., Dworkin, A. G., Toenjes, L. A., & Hill, A. N. (2002). Grade retention and social promotion in Texas, 1994-99: Academic achievement among elementary school students. *Brookings Papers on Educational Policy*, 13-67. Retrieved from ERIC database. (EJ898070)
- McCollum, P., Cortez, A., Maroney, O. H., & Montes, F. (1999). *Failing our children: Finding alternatives to in-grade retention*. San Antonio, TX: Intercultural Development Research Association Institute for Policy and Leadership.
- McCombs, J. S., Kirby, S. N., & Mariano, L. T. (2009). *Ending social promotion without leaving children behind*. Santa Monica, CA: Rand Corporation. Retrieved from [http://www.rand.org/pubs/monographs/2009/RAND\\_MG894.pdf](http://www.rand.org/pubs/monographs/2009/RAND_MG894.pdf)
- National Association of School Psychologists. (2003). *Position statement on student grade retention and social promotion*. Retrieved from [http://www.cdl.org/resource-library/articles/nasp\\_position\\_stmt.php?type=subject&id=17](http://www.cdl.org/resource-library/articles/nasp_position_stmt.php?type=subject&id=17)
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: Author
- National Education Commission on Time and Learning. (1994). *Prisoners of time*. Washington, DC: Author. Retrieved from <http://www2.ed.gov/pubs/PrisonersOfTime/Prisoners.html>
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

- New York City Department of Education. (2009). *Promotion standards: Regulation of the Chancellor, A-501*. Retrieved from <http://docs.nycenet.edu/docushare/dsweb/Get/Document-24/A-501.pdf>
- No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002).
- Obama, B. H. (2011). Remarks by the president on education in Arlington, Virginia. *White House Education Press Releases and Statements*. Washington, DC: Government Printing Office.
- Penfield, R. D. (2010). Test-based grade retention: Does it stand up to professional standards for fair and appropriate test use? *Educational Researcher*, 39, 110-119. Doi: 10.3102/0013189x10363007
- Planty, M., Hussar, W., Snyder, T., Kena, G., KewalRamani, A., Kemp, J., & Dinkes, R. (2009). *The condition of education 2009: Indicator 3—Knowledge and skills of young children*. (NCES 2009-081). Washington, DC: National Center for Educational Statistics, Institute of Educational Sciences. Retrieved from ERIC database. (ED505406)
- 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, 427-469.
- Roderick, M. (1995). Grade retention and school dropout: Policy debate and research questions. *Phi Delta Kappan Research Bulletin*, 15, 1-6. Retrieved from ERIC database. (ED397213)
- Roderick, M., & Nagaoka, J. (2005). Retention under Chicago's high-stakes testing program: Helpful, harmful, or harmless? *Educational Evaluation and Policy Analysis*, 27, 309-340. Retrieved from ERIC database. (EJ737746)

- Roehrig, A. D., Yaacov, P., Nettles, S. M., Hudson, R. F., & Torgesen, J. K. (2008). Accuracy of the DIBELS oral reading fluency measure for predicting third-grade reading comprehension outcomes. *Journal of School Psychology, 46*, 343-366. Retrieved from ERIC database. (EJ789802)
- Russo, A. (2005). Retaining retention. *Education Next, 5*(1), 42-48. Retrieved from ERIC database. (EJ763275)
- Samuels, S. J. (2007). The DIBELS test: Is speed of barking at print what we mean by reading fluency? *Reading Research Quarterly, 42*, 563-566. Retrieved from EBSCOhost database.
- Shepard, L. A. (2004). Understanding research on the consequences of retention. In H. J. Walberg, A. J. Reynolds, & M. C. Wang (Eds.), *Can unlike students learn together? Grade retention, tracking, and grouping* (pp. 183-202). Greenwich, CT: Information Age Publishing.
- Shepard, L. A., & Smith, M. L. (1990). Synthesis of research on grade retention. *Educational Leadership, 47*(8), 84-88. Retrieved from ERIC database. (EJ410221)
- Shinn, M. H., Good, R. H., Knutson, N., Tilly, D., & Collins, V. (1992). Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review, 21*, 459-479. Retrieved from ERIC database. (EJ480688)
- Silberglitt, B., Jimerson, S. R., Burns, M. K., & Appleton, J. J. (2006). Does the timing of grade retention make a difference? Examining the effects of early versus later retention. *Psychology Review, 35*(1), 134-141. Retrieved from ERIC database. (EJ788236)

Texas Education Agency. (2011). *Grade level retention in Texas public schools, 2009-2010*

(Report of the Division of Research and Analysis Department of Assessment and Accountability). Austin, TX: Author. Retrieved from [www.tea.state.tx.us/acctres/Retention\\_2009-10.pdf](http://www.tea.state.tx.us/acctres/Retention_2009-10.pdf)

Tomchin, E. M., & Impara, J. C. (1992). Unraveling teachers' beliefs about grade retention.

*American Educational Research Journal*, 29, 199-223. Retrieved from ERIC database. (EJ446642)

University of Oregon Center for Teaching and Learning. (2008). *DIBELS data system: General*

*information about DIBELS measures*. Retrieved from <https://dibels.uoregon.edu/docs/dibelsinfo.pdf>

U.S. Department of Education. (1991). *America 2000: An education strategy. Sourcebook*.

Washington, DC: Author.

U.S. Department of Education. (1999). *Taking responsibility for ending social promotion: A*

*guide for educators and state and local leaders*. Washington, DC: Author.

U.S. Department of Education. (2003). *Non-regularity guidance: Local educational agency*

*identification and selection of school attendance areas and schools and allocation of Title I funds to those areas and schools*. Retrieved from:

<http://www.doe.in.gov/improvement/federal-legislation-regulations-and-guidance>

U.S. Department of Education. (2007). *Parent and family involvement in education survey of the*

*national household education surveys program*. Retrieved from

[http://nces.ed.gov/pubs2010/2010015/tables/table\\_17a.asp](http://nces.ed.gov/pubs2010/2010015/tables/table_17a.asp)

- Valencia, S. W., Smith, A. T., Reece, A. M., Li, M., Wixson, K. K., & Hewman, H. (2010). Oral reading fluency assessment: Issues of construct, criterion, and consequential validity. *Reading Research Quarterly, 45*, 270-291. Retrieved from ERIC database. (EJ890799)
- Wireless Generation. (2012). *The mClass: Reading three D system*. Retrieved from <https://www.mclasshome.com/wgenhelp/DN3DS/index.htm>
- Wu, W., Hughes, J. N., & West, S. G. (2010). Effect of grade retention in first grade on psychosocial outcomes. *Journal of Educational Psychology, 102*(1), 135-152. Retrieved from ERIC database. (EJ876296)
- Xia, N., & Kirby, S. N. (2009). *Retaining students in grade: A literature review of the effects of retention on students' academic and nonacademic outcomes*. Santa Monica, CA: Rand Corporation.
- Zinth, K. (2005). *Student promotion/retention policies*. Denver, CO: Education Commission of the States.