

GRADE RETENTION RATES: IMPACT OF STANDARDS-BASED GRADING AS
COMPARED TO A TRADITIONAL GRADING SYSTEM IN 3rd AND 5th GRADES

by

Mishea Griffeth Dean

Liberty University

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University

October 2, 2014

GRADE RETENTION RATES: IMPACT OF STANDARDS-BASED GRADING AS
COMPARED TO A TRADITIONAL GRADING SYSTEM IN 3rd AND 5th GRADES

by Mishea Griffeth Dean

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University, Lynchburg, VA

October 2, 2014

APPROVED BY:

Rollen Fowler, PhD, Committee Chair

Karla Swafford, EdD, Committee Member

Katie Thompson, EdD, Committee Member

Scott Watson, PhD, Associate Dean, Advanced Programs

GRADE RETENTION RATES: IMPACT OF STANDARDS-BASED GRADING AS
COMPARED TO A TRADITIONAL GRADING SYSTEM IN 3rd AND 5th GRADES

ABSTRACT

Student retention is a contentious issue that has spawned considerable debate among school leaders and policymakers. NCLB's accountability requirement has led to the development of a standards-based grading system to ensure students master content and so schools can meet AYP goals. This causal comparative study examined whether a change from a traditional grading system to a standards-based system affected the student retention frequency rate. Convenience sampling was used to collect retention data on third and fifth grade students who attended four Title 1 schools in northeast Georgia during the 2007-08 and 2008-09 school years when the grading system changed. A Fisher's Exact test was used to determine if there was a difference between students' retention rates under a traditional grading system and a standards-based grading system. No statistically significant difference was found between the traditional and standards-based grading systems for reducing the frequency of retention for both third grade students and fifth grade students. Gender differences in retention were examined but no statistically significant difference was found suggesting that the change from traditional grading to a standards-based grading system did not affect gender retention rates among third and fifth grade students. The results from this study suggests that the frequency of grade retention with a traditional grading system was not significantly different from the frequency of grade retention that occurred under a standards-based grading system for third and fifth grade students and did not significantly affect the frequency of retention for females and males.

Descriptors: standards-based, traditional grading, student retention, assessment, high-stakes testing, standards-based grade reporting

Dedication

I dedicated this manuscript to my husband, Russ, who has always been my biggest supporter in following my dreams of earning a doctorate degree. Thank you for your love and support and for the unconditional love that you show me daily by being my partner for life.

This paper is dedicated to my father, Tom Griffeth, who has always supported me in pursuing my dreams. Without you, this degree would never have been possible. I want to also dedicate this paper in memory of my mother, Sandra Griffeth. She may no longer be here on earth, but she has an everlasting impact on my life. She is the reason that I am pursuing this dream. From a very young age, she instilled the importance of education. I will never forget one of her quotes, “People can take money and things from you, but an education is something no one can ever take.”

Finally, this paper is dedicated to my friends, family (biological or not), and to my North Jackson family. Without you guys I would have never made it this far. I appreciate everyone listening to me and supporting me throughout this whole process.

Acknowledgements

First, I want to thank my Lord and Savior, Jesus Christ, who has guided my life and given me the strength and courage to complete this endeavor.

Next, I would like to thank Dr. Rollen Fowler, my committee chair, for your guidance through this process. You have definitely challenged me and helped me to bring my paper to another level. I would like to thank Dr. Karla Swafford for always being so prompt and supportive with your feedback throughout the years. I would also like to thank my other committee member, Dr. Katie Thompson, for your support and guidance from the very beginning. From the first summer, spending two weeks in Virginia together, I would not have made it through without you.

Next, I would like to thank Dr. Dave Gorman. I truly owe this accomplishment to your guidance over the past four years. There were days when I thought I would not ever be able to finish, but you were always there to help me understand which direction I needed to go. I would like to thank Dr. Kim Love-Myers for always answering my questions no matter how crazy they might have been. You were always very prompt and thorough with your replies.

Last, but definitely not least, I would like to thank my friends and family who stood by me through this whole dissertation process. Dr. Steven Hardegree, Dr. Amity Hardegree, and Derelle McMenemy: I would like to thank you guys for the great memories and support throughout this whole dissertation endeavor. To my friends Lisa Ellis, Kelly Crump, Stacy Cantrell, Dr. Kristin Mobbs, and Melinda Thomas: I appreciate you guys listening to me and always supporting me. I also want to thank Aunt B and Tamera for your love, support and prayers over the years.

Table of Contents

ABSTRACT.....	3
Dedication.....	4
Acknowledgements.....	5
List of Tables.....	8
List of Figures.....	9
List of Abbreviations.....	10
CHAPTER ONE: INTRODUCTION.....	12
Introduction.....	12
Background.....	16
Problem Statement.....	17
Purpose Statement.....	20
Significance of the Study.....	23
Research Questions.....	23
Null Hypotheses.....	23
Identification of Variables.....	24
Definitions.....	25
CHAPTER TWO: REVIEW OF THE LITERATURE.....	27
Introduction.....	27
Conceptual or Theoretical Framework.....	28
Prior Studies.....	33
Summary.....	55
CHAPTER THREE: METHODOLOGY.....	58
Introduction.....	58
Design.....	60
Questions and Hypotheses.....	62
Null Hypotheses.....	63
Participants.....	63
Setting.....	64
Instrumentation.....	65

Procedures	70
Data Analysis	72
CHAPTER FOUR: FINDINGS	74
Research Question and Hypotheses	75
Data Summary.....	76
Analysis	91
Summary	98
CHAPTER FIVE: DISCUSSION.....	100
Introduction	100
Discussion	103
Limitations of the Study	105
Implications for Future Research	107
Conclusions	113
REFERENCES	116
APPENDIX A: IRB Letter.....	132

List of Tables

Table 1: Student Demographic Information for the 2007-2008 School Year.....	77
Table 2: Student Demographic Information for the 2008-09 School Year.....	78
Table 3: Overall Retention Rate for the 2007-2008 School Year: Traditional Grading System..	79
Table 4: Overall Retention Rate for the 2008-09 School Year: Standards-Based Grading System.....	80
Table 5: Retention Rate in 2007-08 by Grade and Gender, and Traditional-Based Grading System.....	80
Table 6: Retention Rate in 2008-09 by Grade and Gender, and Standards-Based Grading System.....	81
Table 7: Math Frequency Scores: Traditional and Standards-Based Grading System.....	82
Table 8: Reading Frequency Scores: Traditional and Standards-Based Grading System.....	83
Table 9: CRCT Math Scores for 3rd, 4th, and 5th Grade Students.....	84
Table 10: CRCT Reading Scores for 3rd, 4th, and 5th Grade Students.....	84
Table 11: Pearson Chi-Square Test and Fisher’s Exact Test.....	93
Table 12: Overall Retention Rate for Third Grade Students Based on Grade Level.....	94
Table 13: Pearson Chi-Square Test and Fisher’s Exact Test.....	95
Table 14: Overall Retention Rate for Fifth Grade Students Based on Grade Level.....	95
Table 15: Pearson Chi-Square Test and Fisher’s Exact Test.....	96
Table 16: Overall Retention Rate by Gender and Grading System.....	97
Table 17: Pearson Chi-Square Test and Fisher’s Exact Test.....	98

List of Figures

Figure 1: CRCT score on reading test by retention status by grading system.....	85
Figure 2: CRCT score on math test by retention status by grading system.....	86
Figure 3: CRCT score on reading test by retention status and gender.....	87
Figure 4: CRCT score on math test by retention and gender.....	87
Figure 5: CRCT score on reading test by retention rate and grade.....	88
Figure 6: CRCT score on math test by retention rate and grade.....	89

List of Abbreviations

1. No Child Left Behind (NCLB)
2. Adequate Yearly Progress (AYP)
3. National Academy of Education (NAEd)
4. Criterion Referenced Competency Test (CRCT)
5. American Recovery and Reinvestment Act (ARRA)
6. National Association of School Psychologists (NASP)
7. Georgia Department of Education (GaDOE)
8. Georgia Performance Standards (GPS)
9. Standards-based Report Cards (SBRC)
10. Analysis of Covariance (ANCOVA)
11. Florida's Comprehensive Assessment Test (FCAT)
12. Development Scale Scores (DSS)
13. Student Information Systems (SIS)
14. Iowa Test of Basic Skills (ITBS)
15. Standard Error of Measure (SEM)
16. Technical Advisory Committee (TAC)
17. Does Not Meet (DNM)
18. In Progress (IP)
19. Meets (M)
20. Exceeds (E)
21. Institutional Review Board (IRB)
22. Family Educational Rights and Privacy Act (FERPA)

23. Statistical Packages of the Social Sciences (SPSS)

CHAPTER ONE: INTRODUCTION

Introduction

Davenport, Delgado, Meisels, and Moore (1998) stated that grade retention occurs when a student fails to meet a particular grade level's standards. For students who were unable to learn their requisite grade-level skills, the authors asserted that grade retention provided another opportunity to master grade-level content. Yet, when Davenport et al. compared the standardized test scores of retained and low achieving, non-retained students, they found that retained students had lower test scores and were at a greater risk of dropping out of high school.

Similar to Davenport et al. (1998), Jimerson, Pletcher, and Kerr (2005) found grade retention increased the likelihood that a student would drop out of high school. The authors also believed that irrespective of its initial positive impact on students' academic achievement, grade retention could eventually result in low self-esteem, poor interpersonal skills, and social/emotional difficulties. Jimerson et al. (2005) further underscored the negative impact of grade retention on students' future academic achievement. Leckrone and Griffith (2006) concluded that grade retention does not benefit students—academically, socially, or behaviorally. Jimerson et al. (2005) stressed the importance of adapting instructional strategies and developing appropriate intervention strategies for low-achieving students.

While proponents of grade retention believe that low academic achievement is a rationale for grade retention (Karweit, 1999; Nagaoke & Roderick, 2004), Holmes and Saturday (2000), Holmes (1989), and Bonvin, Bless, and Schuepbach (2008) believed that there were other factors to consider, including the impact of grade retention on students' social and emotional development. Holmes and Saturday (2000) investigated students' perceptions of grade retention and found students believed that grade retention was a punitive measure. Furthermore, the

authors found peers ridiculed retained students.

Holmes (1989) compared the non-retained and retained groups' "measures of personal adjustment" (p. 13) and found retained students had lower self-esteem and personal adjustment scores. Xia and Glennie (2005) concluded that grade retention had a negative impact on students' socio-emotional development. In contrast, Bonvin's et al. (2008) study did not find a negative association between grade retention and students' social and emotional development. Despite its negative impact on students' social and emotional development (Jimerson et al., 2005), Nagaoka and Roderick (2004) concluded that grade retention gave students another opportunity to master grade-level skills before they advance to the next grade.

Like Nagaoka and Roderick, Karweit (1991) found students, who had not mastered grade level skills, benefitted from an additional year at their present grade level. Grade retention advocates, Gesell Institute (1982) and Ames (1966, 1980), believed that grade retention was effective prior to kindergarten and before first grade. Gesell and Ames stressed the importance of achievement testing when determining a child's developmental level. Gesell and Ames further stressed that not only is a student's academic achievement pertinent in a grade retention decision, but also, his or her maturity and behavior should determine grade retention. Greene and Winters (2006) compared the academic achievement of retained and socially promoted students. The authors found the socially promoted students had higher academic achievement scores than their promoted peers. Greene and Winters concluded that test-based grade retention practices increased students' reading achievement.

Nagaoka and Roderick (2004) discussed a potential link between grade retention and students' socioeconomic status, gender, and ethnicity. The authors concluded that grade retention was more prevalent with students from poor socioeconomic backgrounds because there

were low levels of academic support in the home environment and parental involvement in school-related activities. Nagaoka and Roderick also compared the grade retention rates of African American, Latino, and Caucasian students and found African American and Latino students had higher retention rates; in addition, male students had higher retention rates than female students.

Wilson and Hughes (2009) focused on parental involvement as a factor of grade retention. These two authors concluded that parental involvement is a critical component in a grade retention decision. Wilson and Hughes asserted that parents who have a trusting relationship with their children's schools are more likely to agree with a grade retention decision. Additionally, when the authors investigated the cognitive and educational levels of these students' parents, they found their mothers tend to have lower cognitive functioning, and their families have less formal education and less income. McCoy and Reynolds (1999) expanded upon these findings as they examined predictive factors of grade retention. The authors concluded the following were strong predictors for grade retention: (a) if a student has low academic achievement during their early education years, (b) if the student was a male, (c) if there is low parental involvement in the student's life, and (d) if the student's family has a high rate of mobility.

According to Nagaoka and Roderick (2004), Wang and Wang (2007), and Xia and Glennie (2005), grade retention is a costly decision. Nagaoka and Roderick (2004) noted that 7,000 retained students cost one school system \$35 million. Wang and Wang (2007) calculated that the cost of retaining 2.4 million students is around \$10 billion. Xia and Glennie (2005) found that each year five to nine percent of America's K-12 student population is retained, and the retained students are costing taxpayers \$18 billion annually.

Xia and Glennie (2005) not only discussed grade retention's short-term financial impact, but they also discussed its long-term impact. Xia and Glennie stated, "Other fiscal costs associated with retention include decreased lifetime earnings among retained students, foregone earnings due to delayed entry into the workforce, and decreased government tax revenues associated with the decreased earnings of retained individuals" (p. 3). When compared with other preventative measures, the Xia and Glennie concluded grade retention is an ineffective and costly measure.

Xia and Glennie (2005) further asserted grade retention is a controversial issue that has generated considerable debate among researchers, educators, and policymakers. Davenport et al. (1998) stated that student retention is a popular strategy for students whose academic achievement is unsatisfactory. Davenport et al. asserted there is a belief that retention enables students to remain at their current grade level until they catch up to their peers. Once the students learn the requisite material, they can move ahead and succeed.

As a result of accountability mandates, such as the No Child Left Behind (NCLB) and its Adequate Yearly Progress (AYP) measurement, schools are reexamining their grade retention policies (Brown, 2007; NAED, 2009). Grade retention, also referred to as student retention, is a contentious issue that has spawned considerable debate among school leaders and policymakers (David, 2008). Brown (2007) cited NCLB's requirement that all students demonstrate grade-level proficiency in reading and math. This requirement has a profound effect on Title 1 schools that are unable to meet yearly achievement benchmarks as these schools will lose federal funding and face corrective action.

Because all students are required to demonstrate proficiency on grade-level assessments, Brown (2007) and the National Academy of Education (2009) encouraged schools to ensure their

standards and curriculum increase students' academic achievement. Consequently, school systems are reexamining their curriculum, instructional strategies, assessments, and student retention policies. The NAEEd asserted that standards-based education is guiding states' educational policies and reforms.

Johnson (2000) found school systems are implementing state standards to guide student instruction, which has resulted in a standards-based grading system as opposed to a traditional grading system. Johnson underscored the effectiveness of a standards-based grading system, stating that it provided teachers, parents, and students with comprehensive feedback. Guskey (2005) concluded a standards-based grading system guides teachers' instruction and ensures that students master grade-level standards by the conclusion of the school year.

Background

To address students' achievement gaps, public schools have implemented new curriculum standards and utilized high-stakes testing to measure student learning (Guskey, 2005). In a subsequent study conducted by Guskey (2007), he found the NCLB's federally mandated educational reforms required all students to demonstrate proficiency in reading and math by the conclusion of the 2013-2014 school year. According to the U.S. Department of Education (2010), the Race to the Top Assessment Program, which was authorized under the American Recovery and Reinvestment Act of 2009 (ARRA), seeks to accomplish the following:

Develop assessments that are valid, support and inform instruction, provide accurate information about what students know and can do, and measure student achievement against standards designed to ensure that all students gain the knowledge and skills needed to succeed in college and the workplace. These assessments are intended to play a critical role in educational systems; provide administrators, educators, parents, and

students with the data and information needed to continuously improve teaching and learning; and help meet the President's goal of restoring, by 2020, the nation's position as the world leader in college graduates. (U.S. Department of Education, 2010, para. 1)

Brown (2007) provided further evidence that federal educational reforms are forcing schools to reform their pedagogical practices, formative assessment strategies, and content standards. Similar to educational reform, Guskey (2005) stated that standards-based education's goal is to improve teaching and student achievement. Yet, is a standards-based grading system capable of improving instructional pedagogy and student achievement?

Moreover, can a standards-based grading system decrease student retention? Forman and Sanders (2001) associated grade retention with increased behavioral problems and higher absentee rates among retained students. Studies in New York and Chicago revealed that retained students were more likely to drop out when compared to their promoted peers (Roderick, 1995). Prior studies found that grade retention, irrespective of a student's grade level, increased the chance of a student dropping out of school (Hauser, 1999; Holmes, 1989; NASP, 1998; Thompson & Cunningham, 2000). Due to grade retention's overall negative impact on student success throughout their school years (Forman & Sanders, 2001; Jimerson et al., 2005; Roderick, 1995), determining whether a traditional grading system or a standards-based grading system improves students' academic achievement is both relevant and important. The focus needs to be on addressing students' needs and not on following federal educational mandates.

Problem Statement

A review of the literature revealed that grade retention has a negative impact on students' academic and social outcomes (Brown, 2007; Jimerson, 2001; Xia & Glennie, 2005). Even students who had demonstrated quantifiable achievement gains in the year following their grade

retention regressed within three years (Hauser, 1999; Holmes, 1989; Jimerson et al., 2005; Karweit, 1991; Thompson & Cunningham, 2000). Karweit (1991) stated, “The consensus of several extensive reviews of grade retention is that there is not a positive effect for grade retention on academic achievement or on student personal adjustment” (p. 4).

Does a traditional grading system or a standards-based grading system improve grade retention rates? Carifio and Carey (2010) asserted that under a traditional grading system, struggling students had failing grades and were disengaged in the learning process; unfortunately, failing grades and other punitive measures do not correct the problem. In fact, Carifio and Carey concluded these measures perpetuate the problem. Carifio and Carey encouraged schools to focus on comprehensive reforms, including assessment procedures and motivational strategies.

Brookhart (2011) believed a traditional grading system was incompatible with current educational reforms because it does not reflect what students’ mastery of student learning. However, Brookhart concluded that a standards-based grading system, which encourages students to focus on acquiring knowledge and not on earning grades, increased students’ motivation and desire to learn. Yet, how would a transition from a traditional-based to a standards-based grading system impact grade retention rates?

When comparing grade retention rates under a traditional grading system and a standards-based grading system, research found grade retention rates were higher under a traditional grading system (Johnson, 2001). Johnson (2001) highlighted the prevalence of grade retention among America’s students, concluding that schools retained 15% of students each school year. Scriffiny (2008) found that standards-based grading decreased the percent of students retained each year. Scriffiny attributed the lower student retention rates under a standards-based grading

to teachers making instructional decisions based on students' distinct learning needs.

Guskey (2001) also advocated for a standards-based grading system. The author concluded that standards-based grading provides students with multiple opportunities to demonstrate content knowledge through various means and in different contexts. Unlike a traditional grading system, which averages students' grades and does not allow them to recover from a zero, a standards-based grading system gives students multiple opportunities to succeed as their current academic performance replaces zeroes. With this in mind, a standards-based grading system could reduce student retention rates (Guskey, 2001).

Concerning grade retention's effect on students' behavior, the National Association of School Psychologists' (NASP, 2003) found a link between grade retention and behavioral issues. The NASP found retained students are more likely to have behavioral issues when compared to their non-retained peers. Xia and Glennie (2005) concluded there is an association between grade retention and higher incidences of deviant behavior because "old-for-grade" (p. 21) students are more likely to experiment with drugs and engage in risky behavior. Holmes (1989) and Bowman (2005) found that retained students struggle with school attendance.

Jimerson (2001) denounced the notion that student retention is a positive and viable option. In fact, Jimerson stated grade retention negatively affects a student's educational development. Conversely, Johnson (2001) found that grade retention improved students' academic achievement. There are many reasons why grade retention is a successful practice: social/emotional immaturity, a lack of academic growth, and an inability to meet state standards (Bowman, 2005). When comparing retained and promoted students' basic skills, Greene and Winters (2006) concluded an extra year at their current grade level provided students with another opportunity to meet grade-level competencies.

Purpose Statement

The purpose of this study was to determine whether a traditional or a standards-based grading system improved 3rd, 4th, and 5th grade retention rates of Title 1 elementary schools that were part of one North Georgia school system. The target schools' transition from a traditional grading system to a standards-based grading system provided the impetus for this study.

According to the NCLB Act and its AYP mandates, students are required to learn the curriculum standards set forth by their state government. According to the Georgia Department of Education (GaDOE, 2011a), AYP is one of the foundations of the NCLB Act of 2001. AYP measures a school's ability to meet its achievement benchmarks. The state of Georgia uses the students' Criterion Referenced Competency Test (CRCT) scores as one of its benchmarks for assessing a school's effectiveness.

This study utilized a non-experimental, causal comparative, ex post facto research design. This research design was appropriate because the researcher compared two independent variables—standards-based assessment and traditional-based assessment—and their impact on the dependent variable—students' grade retention rates. Unlike experimental research “the groups being compared in causal-comparative studies have already been formed, and any treatment (if there was a treatment) has already been applied” (Johnson, 2000, p. 3). In this study, the independent variables were grouping variables: students who received instruction under a traditional assessment system and students who received instruction under a standards-based grading system. Because this study compared two groups' grade retention rates from the 2007-2008 and 2008-2009 school years, with the 2007-2008 students forming the traditional-based assessment group and the 2008-2009 students forming the standards-based grading, the researcher used pre-existent data. In causal-comparative studies, Johnson (2000) stated the

researcher examines two groups' data and determines "if he or she can offer a reasonable explanation (i.e. what 'caused') [of] the existing differences between the two groups" (p. 3).

In addition to examining pre-existing groups on a dependent variable, researchers do not manipulate the independent variables in causal-comparative studies (Schenker & Rumrill, 2004). With this in mind, this study's categorical independent variables—traditional and standards-based assessments—were "not experimentally manipulated" (Schenker & Rumrill, 2004, p. 118). Conversely, researchers manipulate the independent variable when conducting experimental research (Gay, Mills, & Ariasian, 2006). A causal-comparative design was also appropriate because the researcher conducted a retrospective analysis of the data (Fraenkel & Wallen, 2006).

Since its inception in Georgia's public schools, standards-based grading system has determined whether a student fails to meet, meets, or exceeds the pre-determined assessment criteria. Unlike a traditional grading system, a standard-grading system provides comprehensive feedback on a student's academic progress (Guskey, 2001). Because the traditional-grading system and the standards-based grading system utilized identical promotional criteria, the researcher was able to compare and determine which grading system impacted grade retention rates. The target school system's belief that a standards-based grading system decreased grade retention was worthy of investigation. The ability to determine if a standards-based grading system decreased students' grade retention rates could influence its future implementation in the target school system.

In addition to determining if standards-based grading influenced grade retention rates, the researcher believed this study could lead to supplementary instructional programs that improve grade retention rates. The increased utilization of both standards-based instruction and assessments within America's public schools may decrease student retention and improve

student achievement and success (Forman & Sanders, 2001).

This study occurred in a rural public school district located approximately 60 miles northeast of Atlanta, Georgia. The Georgia Department of Education governs all public schools in the state of Georgia (GaDOE, 2008). The GaDOE has implemented performance standards to guide schools' curriculum. The four elementary schools included in this study were part of the same county school system. All four had Title 1 status as well as similar racial and socioeconomic demographics. Chapter 3 provides an overview of the four schools' demographics.

Since 2008, the target school system has participated in standards-based grading and state-mandated CRCT testing each spring. This study examined the target schools' grade retention rates under two different assessment programs: traditional grading and standards-based grading. This study utilized student data from a two-year period (2007-08 and 2008-09), with the first year examining grade retention rates under a traditional grading system (2007-08) and the second year (2008-09) examining grade retention rates under a standards-based grading system. The goal of this study was to determine if the target schools' grade retention rates decreased as a result of transitioning from a traditional grading system to a standards-based grading system.

Because students' standardized test scores determine a school's effectiveness, standards-based instruction is appropriate and logical (NAEd, 2009). Forman and Sanders (2001) asserted that standards-based instruction needs to be aligned with standards-based assessments. Forman and Sanders believed a standards-based instruction and a standards-based grading system ensures that students meet instructional standards and therefore increase the chance for academic success.

Significance of the Study

This study provided quantitative data regarding the standards-based grading system's effect on grade retention rates at four Title 1 elementary schools. In order to determine the standards-based grading system's effect on grade retention rates, the researcher compared grade retention rates under a standards-based grading system and a traditional grading system. This study's results could persuade or dissuade the future implementation of a standards-based grading system at the four target elementary schools. Grade retention rates have a direct effect on a school's AYP (GaDOE, 2010). Schools that are unable to meet their achievement benchmarks, as measured by AYP, could lose federal educational funding.

Research Questions

The following research questions guided this study:

RQ1: Is there a difference between the third grade students' grade retention rates with a traditional grading system and their grade retention rates with a standards-based grading system?

RQ2: Is there a difference between the fifth grade students' grade retention rates with a traditional grading system and their grade retention rates with a standards-based grading system?

RQ3: Is there a difference among the third and fifth grade female students' grade retention rates as a result of transitioning from a traditional based grading system in 2007-08 to a standards-based grading system in 2008-09?

RQ4: Is there a difference among the third and fifth grade male students' grade retention rates as a result of transitioning from a traditional based grading system in 2007-08 to a standards-based grading system in 2008-09?

Null Hypotheses

H₀₁: There will not be a significant difference between the 2007-08 third grade group's

grade retention rate with a traditional-based grading system and the 2008-09 third grade group's grade retention rate with a standards-based grading.

H₀₂: There will not be a significant difference between the 2007-08 fifth grade group's grade retention rate with a traditional-based grading system and the 2008-09 fifth grade group's grade retention rate with a standards-based grading.

H₀₃: There will be not be a significant difference among the third and fifth grade female students' grade retention rates as a result of transitioning from a traditional based grading system to a standards-based grading system.

H₀₄: There will be not be a significant difference among the third and fifth grade male students' grade retention rates as a result of transitioning from a traditional based grading system to a standards-based grading system.

Identification of Variables

Independent Variables

For this study, the independent variables were the traditional grading system and the standards-based grading system. The independent variables were categorical variables because the researcher divided the students into two preexisting groups: students who received traditional-based instruction during the 2007-08 school year and students who received standards-based instruction during the 2008-09 school year. The students' scores on the CRCT were the promotional criteria for third and fifth grade students. Fourth grade in the state of Georgia does not have CRCT promotion criteria scores. Under an identical promotion criterion, the researcher compared third grade and fifth grade students' subject tests on the CRCT.

Dependent Variables

This study's dependent variables were each grade level's retention rate, which the

researcher compared to the students' reading and math CRCT scores under a traditional grading system and under a standards-based grading system.

Definitions

Definition of Key Terms

The Georgia Department of Education (GaDOE, 2011a) stated that the Criterion Referenced Competency Test (CRCT) “is designed to measure how well students acquire the skills and knowledge described in the Georgia Performance Standards (GPS)” (p. 8). In reference to the Education Reform Act of 2000, the CRCT tests students in grades one through eight in reading, English/Language Arts, and mathematics, along with social studies and science for grade three through eight. The CRCT is an assessment that measures whether or not students have learned the requisite state standards.

Adequate Yearly Progress (AYP): AYP is an accountability measurement that determines whether schools and school systems are making sufficient progress towards meeting their achievement benchmarks (USDOE, 2010).

Georgia Performance Standards (GPS): The Georgia Performance Standards provide academic performance expectations for student achievement in the academic subject areas (GaDOE, 2009).

Standards-based report card (SBRC): The SBRC is a grade-reporting system that measures a student's achievement based on his or her ability to make progress on specific learning goals or standards. Grades are not reported as an average or letter grade; instead, a student's progress is reported in the following manner: (a) does not meet, (b) meets, or (c) exceeds the standard (GaDOE, 2009).

Standards-based grading: A standards-based grading system does not measure student progress based on a traditional A, B, C, D, and F grading system. Instead, subject-area standards determine students' progress, and teachers report student progress in the following manner: (a) does not meet, (b) in progress, (c) meets, or (d) exceeds expectations. Grades are based on the student's achievement toward individual subject standards (Guskey, 2001).

Traditional Grading System: A traditional grading system provides a letter grade or a numerical score for each subject area (Cherniss, 2008). A traditional grading system also uses a letter grade and a corresponding scale. The scale could be as follows: A = 93-100, B = 85-92, etc. Unfortunately, students view their grades as numerical averages as opposed to a reflection of their performance. Cherniss (2008) found that with a traditional grading system, students' grades are a measure of their performance on assignments and tests and the amount of effort they put forth in class.

CHAPTER TWO: REVIEW OF THE LITERATURE

Introduction

Grade retention is a contentious and complex issue within education (Cannon & Lipscomb, 2011). Recent studies found grade retention had mixed results (Lorence & Dworkin, 2006; Poland, 2009). Educators and school administrators have two options: promote the student to the next grade or retain the student at his or her current grade (Early et al. 2003). According to Denton (2001), grade retention occurs because a student is unable to master grade-level skills or he or she has behavioral issues.

Jimerson (2001) compared the grade retention rates of America's public schools and those of other countries. Jimerson noted that Norway and Iceland had an automatic progression plan in place. Jimerson estimated that each year schools across the United States retained close to 10% of their students. With the exception of Rwanda and Togo, which are underdeveloped countries, Holmes (2006) found that the United States had the highest student retention rate.

Holmes (1989), Holmes and Matthews (1984), Jackson (1975), and Jimerson (2001) all concluded that grade retention is an ineffective strategy that negatively impacted students. Holmes (1989) found students developed a poor self-concept and a negative attitude towards school because of grade retention. Jimerson, Carlson, Rotert, Egeland, and Sroufe (1997) believed that retained students struggled emotionally and socially in the school setting. Jimerson (1999), Roderick (1994), and Shepard and Smith (1990) concluded that grade retention also negatively impacted a school's dropout rate.

Jimerson et al. (2005) determined that grade retention initially improved students' academic achievement, but students eventually fell behind their grade-level peers within three years. Gleason, Kwok, and Hughes (2007) found that grade retention had a positive impact on

students. In fact, Gleason et al. concluded that retained first grade students had a higher peer acceptance rate than their promoted peers.

In general, Roderick (1994) and Nagaoka and Roderick (2004) found that grade retention occurred more often with boys than girls as well as with African-American and Latino students than Caucasian students. When Wilson and Hughes (2009) investigated grade retention indicators, they concluded that a student's ethnicity and socioeconomic status were predictive factors. The higher grade retention rates among male students and certain ethnic groups prompted schools to reexamine their assessment strategies. Guskey (2001) stated that schools have focused on aligning state standards, teaching methods, and standardized assessments. However, schools should also consider a student's socioeconomic status, gender, developmental readiness, physical abilities, learning disabilities, school attendance, and chronological age (Grant, 1997). This study examined schools' grade retention rates, the impact of standardized testing on grade retention, and the effect of a traditional grading system and a standards-based grading system on students' grade retention rates.

Conceptual or Theoretical Framework

This study's theoretical framework was Vygotsky's social constructivist theory. The foundation for his theory, "the zone of proximal development," (Vygotsky, 1978, p. 352) underscores the importance of instructing students at their current instructional level. Vygotsky's theory is a sequential process in which teachers determine what academic tasks a student is able to complete independently and what academic tasks require teacher assistance. Once the teacher determines that a child needs assistance, the teacher adapts the student's instruction. Vygotsky (1978) believed in the interconnectedness of curriculum, instruction, and assessment, with a focus on the learning process and the ability of scaffolding to impact students'

learning.

Vygotsky (1978) also stressed the importance of teacher-student interaction in which the teacher models and instructs while the learner takes responsibility for learning, developing, and practicing a skill. During this process, the student demonstrates the core principles of standards-based education and assessments as the student has a direct role in his or her learning process (Hatch, 2010). The concept of utilizing assessment to guide instruction, in addition to measuring students' learning, provides the basis for standards-based grading (Guskey, 2007).

With this in mind, Hatch (2010) found the literature supported Vygotsky's theory: learning leads to development; in addition, the literature supports Piaget's theory: development leads to learning as elementary schools incorporate both theories. Hatch discussed and stressed the importance of using curriculum, instruction, and assessment as a framework for ensuring that elementary education is focused on "teaching for learning." Hatch stated, "Learning should be the stuff of early education, curriculum content should be the focus of what children learn, and teachers should use as many teaching strategies as necessary to maximize every child's opportunity to learn" (p. 264).

In reference to teachers incorporating various teaching strategies (Hatch, 2010), Thomas Guskey (2007), an authority on standards-based grading, discussed Benjamin Bloom's theory of mastery learning. Through the use of varied assessments (portfolios and projects), Bloom's goal was to close the achievement gap between high-achieving and low-achieving students while not compromising achievement standards (Guskey, 2007). Guskey (2007) stated:

A far better approach, according to Bloom, is for teachers to use their classroom assessments as learning tools, both to provide students with feedback on their learning progress and to guide the correction of learning errors. In other words, instead of using

assessments only as evaluation devices that mark the end of a unit, Bloom recommended they be used as part of the instructional process to diagnose individual learning difficulties and to prescribe remediation procedures. (p. 11)

Students' diverse socioeconomic/cultural backgrounds, learning styles, and experiences mandate that schools adopt multiple forms of assessment and help students master learning standards (Jimerson, 2001). According to the constructivist theory, it is unrealistic for students to master concepts in a similar manner or at an identical time. Instead, a student's learning should be individualized and determined by his or her environment (Palinesar, 1998).

Grade Retention: A Historical Perspective

Jackson (1975) defined grade retention as "the practice of requiring a student who has been in a given grade level for a full school year to remain at that level for a subsequent school year" (p. 613). Picklo and Christensen (2005) found grade retention's initial intent was to give students an additional year in their current grade. The central premise of grade retention was that exposing a student to the same material would improve his or her academic achievement (Picklo & Christensen, 2005).

From a historical perspective, Grant (1997) discovered that during the 1950s some students came to kindergarten unprepared to learn and were unable to acquire the requisite knowledge to move onto first grade. Grant found that if a second year at the same grade level did not increase the students' academic skills, then schools retained the students again. After schools retained students several times, they might dropout due to frustration or embarrassment (Grant, 1997).

Currently, the NCLB Act and its accountability mandates have forced schools' to reexamine their respective grade retention policies (Leckrone & Griffith, 2006). Parts of this

legislation have improved students' academic achievement and provided students with equal educational opportunities (Leckrone & Griffith, 2006). Interestingly, students who are retained tend to have positive peer relationships as their peers believe these students are smart; yet, their peers are unaware that these students are repeating a particular grade (Bonvin et al., 2008).

Legislative Impact and Standardized Testing

To illustrate its commitment to improving public education, the legislature passed the NCLB Act, which mandates that 100% of a school's students be proficient in reading and math by 2014 (Brown, 2007). Leckrone and Griffith (2006) stressed that school systems must monitor students' progress on the 100% proficiency in reading and math by 2014. According to Leckrone and Griffith, the NCLB's accountability standards impacted both grade retention and graduation rates. While NCLB's mandates have improved the quality of students' education, they have resulted in many students being "left behind" as grade retention rates have increased (Leckrone & Griffith, 2006).

Powell (2010) found a connection between the inception of NCLB and the increase in grade retention rates in schools. Regardless of NCLB's original intent, which was to ensure that no child was "left behind," Powell found grade retention rates are proliferating. Powell believed NCLB, and its AYP measure, has resulted in a restrictive, test-oriented curriculum that is devoid of ingenuity. Instead of creating a learning environment in which the teacher and student work together, the student has become solely responsible for learning outcomes.

Through increased standards and accountability, legislators' sought to improve education (Carnoy & Loeb, 2002). In order to ensure high levels of student achievement, states have implemented high-stakes tests (Craig, 2012; Holmes, 2006; Paepflow, 2011). Picklo and Christensen (2005) discussed the impact of recent educational reforms on grade retention and

social promotion decisions. They reasoned that student retention and social promotion decisions have become more complicated because of recent educational reforms. Needless to say, government entities have adopted a myopic viewpoint concerning the grade retention issue as schools have utilized high-stakes tests as the primary criterion for determining retention and promotion decisions (Sengupta, 1997). The state of Georgia utilized the CRCT and had cut-off scores for subjects such as math and reading. At certain grade levels, these cut-off scores determined whether a school promoted or retained a student.

Critics denounced the practice of retaining students based on their standardized test scores (Hauser, 1999; Holmes, 2006). Holmes (2006) recommended that schools use standardized test results to improve curriculum and not to determine student retention. “Though well-meaning individuals are looking to high-stakes testing results for promotion as a means of ensuring that all students learn, retaining students ‘has a much greater impact’ on minority and poor youths than on majority, middle-class children” (Hauser, 1999, p. 64). In contrast, policymakers do not believe high-stakes tests and standards-based reforms disproportionately target students from poor socioeconomic backgrounds (Brown, 2007). Brown (2007) concluded standards-based reforms and high-stakes tests ensure that students develop the requisite academic skills to succeed in life.

In summary, if schools use students’ standardized test scores as the criterion for a grade retention decision, then they fail to consider students’ wellbeing (Horn, 2003). School systems need to view students’ academic success in a holistic manner (Leckrone & Griffith, 2006). According to the Center for Policy Studies, Educational Research, and Community Development (CPSER, 2008), standardized testing should not be the only factor in a grade-to-grade promotion decision. The CPSER concluded that a student’s test scores might not reflect his or her overall

knowledge. The increase in grade retention rates underscores the importance of reassessing student assessment (Smith & Shepard, 1989).

Prior Studies

Prior research investigated the effect that instruction and curriculum had on students' CRCT scores (Adams, 2011; Phemister, 2009). Adams (2011) "used a causal-comparative design to determine if at-risk students' standardized tests were impacted by remedial math instruction and if gender differences in math abilities existed for the remedial students" (p. 44). The participants in Adam's study were a convenience sample. The inclusion criteria for Adams' study were current enrollment at one of three northeast Georgia middle schools and an at-risk student designation. Adams described an at-risk student as one who scored below 810 on the 2010 sixth-grade CRCT math test. The treatment group included students who scored below 810 on the 2010 sixth-grade math section of the CRCT. These students received both remedial and regular math group instruction. The control group included students who scored below 810 on the math section of the CRCT. These students received regular math group instruction, but they did not receive remedial math instruction.

Adams (2011) used a one-way analysis of covariance (ANCOVA) to determine if there was a difference between the mean achievement scores of the control and treatment groups. The students' 2010 CRCT scores were the dependent variables while the status of remediation (students enrolled in a remedial math course, students not enrolled in a remedial math course) were the independent variables. The students' 2010 CRCT scores were the covariates.

Adams (2011) conducted a second analysis of covariance and examined the difference in mean math achievement scores based on gender. Adams found students who received the remedial math course (independent variable) had significantly higher 2010 CRCT math scores

(dependent variable) when compared to the at-risk students who did not receive remedial math instruction. However, the study did not find a significant difference between the male and female students' 2010 Math CRCT scores for at-risk students.

Adams' 2011 study had several limitations in that he only included two rural northeast Georgia school districts; therefore, Adams' results may not be generalizable to other school districts that have different student demographics, math standards, or standardized tests. In addition, because of the study's non-equivalent groups, there was a selection threat as Adams was unable to randomly assign the control and treatment groups.

Phemister's (2009) causal-comparative study examined "the effectiveness of reading instruction through Georgia's Choice curriculum on third grade science CRCT scores" (p. 81). Phemister compared the third grade CRCT science scores for 105 Georgia elementary schools that implemented the Georgia's Choice Curriculum reform model and 105 Georgia elementary schools that did not implement the Georgia's Choice Curriculum reform model. In Phemister's study, the independent variable was the group that implemented Georgia's Choice Curriculum reform model. This group consisted of third grade students who attended 105 of Georgia's Choice elementary school as well as students previously enrolled in Georgia's elementary schools, irrespective of their socioeconomic, racial, academic, gender, geographic, or cultural factors. The second comparison group was the other independent variable. This group included third grade students who attended the 105 elementary schools that did not implement Georgia's Choice Curriculum. The dependent variable was the students' CRCT scores.

As to why the researcher selected a particular grade level, Phemister cited NCLB's mandate that required all third grade students to read on grade level. Phemister also found the Georgia Department of Education used the third grade students' CRCT Reading score as a factor

in a student retention or promotion decision. The researcher noted that there was no manipulation with respect to the study's independent variable. In order to decrease any threat to validity, Phemister utilized "a pretest-posttest with an untreated comparison group" (p. 81) and compared the groups' CRCT scores in the following manner: (a) 2002 pretest and 2004 posttest, (b) 2004 pretest scores and 2005 scores, (c) 2005 pretest scores and 2006 scores, and (d) 2007 pretest scores and 2008 scores.

To determine if there was a significant difference in the CRCT scores of Georgia Choice schools, which was the treatment group, and the CRCT scores of the group that did not receive the treatment, Phemister (2009) utilized an analysis of covariance. Subsequent analysis revealed that both groups' CRCT scores increased by four points when compared to their 2002 scores. Phemister also determined that Georgia schools that did not participate in Georgia's Choice reading program had higher reading scores than Georgia schools that participated in Georgia's Choice reading program.

A meta-analysis conducted by Jimerson (2001) summarized the results of 20 studies that investigated the effectiveness of grade retention. Jimerson's meta-analysis included studies that occurred between 1990 and 1999. Jimerson matched and compared a group of retained students and a group of promoted students on at least one of the following variables: IQ, academic achievement, socio emotional adjustment, SES, or gender. Jimerson was able to match the two groups on at least one variable in 19 of the 20 studies and at least 2 or more variables in 18 of the 20 studies in the meta-analysis. Jimerson found that 17 of the 20 studies investigated student outcomes in grades one through seven and 14 of the 20 studies examined student achievement over consecutive years.

When comparing the academic achievement outcomes of retained students and promoted

students (comparison group), Jimerson found that the 20 studies utilized 175 analyses. Jimerson investigated 91 of these analyses and found a statistically significant difference; however, only nine of the analyses were positive for the retained students while 72 analyses were positive for the promoted students. Jimerson also found that 84 did not find a statistical difference between the promoted and retained students. Jimerson determined 169 effect sizes from 18 distinct studies that included 1,249 retained students and 1,557 comparison students. Jimerson stated that the “retained group scored .39 of a standard deviation lower than the comparison promoted group” (p. 429). When Jimerson examined “the effect sizes for language arts, reading, math, composite scores, and grade point average, the comparison group was higher in all areas (.36, .54, .49, .20, and .18, respectively)” (p. 429).

Regarding the socio emotional and behavioral outcomes, Jimerson (2001) found that 16 of the 20 studies addressed these outcomes, resulting in 148 analyses. Jimerson concluded there were statistically significant differences between the groups in 127 of the 148 analyses. For the remaining 21 studies, 13 analyses were positive for the promoted students and eight were positive for the retained students. Based on 16 studies, Jimerson computed 246 effect sizes and categorized 77 effect sizes as social emotional and behavioral adjustment. Jimerson found a mean of -0.22 for the 77 effect sizes. Jimerson also determined that the effect size was lower for the retained groups of students in the following areas: (a) social = .08, (b) emotional = .28, (c) behavioral = .11, (d) self-concept = .04 and (e) ratings of adjustment = .15.

Finally, Jimerson examined the efficacy of grade retention. Jimerson discovered that only four studies (20%) concluded that grade retention resulted in positive outcomes for students. Further analyses found an effect size of .09 for retained students in the year following their retention. However, Jimerson’s longitudinal outcomes revealed a mean effect size of -.31 for the

retained groups of students. Jimerson noted that while grade retention benefits students in the short term (.09 the year following retention), these improvements were not sustainable (-.31 longitudinal results); in fact, Jimerson concluded that students regress years later.

Because previous studies found that grade retention had an immediate and positive effect on student's achievement, Jimerson stressed the importance of examining the long-term effects of grade retention on students. Jimerson suggested that educational professionals and lawmakers considered the purported connection between grade retention and high school dropout rates. Jimerson stated that schools "implement and examine remedial strategies that facilitate academic success" (p. 433). Forness, Kavale, Blum, and Lloyd (1997) found that mnemonics, augmented reading comprehension, direct instruction, behavior modification, formative evaluation, and early intervention are the best practices.

Grade Retention

For over a century, studies have examined grade retention's impact in the educational setting (Greene & Winters, 2006; Picklo & Christensen, 2005; Wilson & Hughes, 2009). Picklo and Christensen (2005) concluded that grade retention has no long-term benefits: academically, socially, or behaviorally. In its 2003 study, the NASP concluded that grade retention results in ephemeral gains for students; in fact, any achievement gains are generally lost within three years (Silberglitt, Jimerson, & Burns, 2006). David (2008) found that hundreds of studies have examined grade retention, with most focusing on retention at the elementary level.

While prior studies examined this topic from various perspectives and methodological approaches, David (2008) concluded studies that compared same grade retained and non-retained students were the most informative. David determined the majority of researchers found that grade retention did not benefit students: academically, socially, and personally. David suggested

that schools focused on the reasons why students fail rather than retention and promotion decisions.

Safer (1986) identified the factors that contributed to grade retention at both the elementary and junior high levels. For elementary school non-promotion, Safer found that an IQ below 90, grade-level achievement one year below a student's grade level, and behavioral problems were significant risk factors. When Safer compared the IQs and achievement discrepancies between non-promoted first grade students and students in grades two through five, the author found that first grade students had "lower IQs and greater achievement deficits than do those retained in grades two through five" (p. 500). For junior high students, Safer found "the school factors most closely associated with JHS grade retention were: persistent elementary school (ES) classroom misconduct, junior high school (JHS) suspension and excessive ES and JHS absenteeism" (p. 501).

Prior research investigated the following aspects of grade retention: (a) criteria for grade retention, (b) the decision process for grade retention, (c) comparisons between retained pupils and their peers, (d) efficiency of grade retention, (e) the effect of grade retention on academic achievement, and (f) the effects of grade retention on students, socially and emotionally (Holmes, 1989; Jimerson et al., 1997; Nagaoka & Roderick, 2004; Wang & Wang, 2007). Even though grade retention resulted in short-term benefits for students (Jimerson et al., 1997), research did not support its long-term benefits (Holmes, 1989; Nagaoka & Roderick, 2004; Wang & Wang, 2007). Researchers found that grade retention's current framework was initially positive for students, but it had negative academic and social outcomes years later (Alexander, Entwisle, & Dauber, 2003).

In 54 studies that compared retained and non-retained students, Holmes (1989) found that

retained students had lower standardized test scores. In addition, Holmes discovered that only nine of the fifty-four studies determined that retention had short-term positive outcomes for students. McCoy and Reynolds (1999) found an association between grade retention and lower reading achievement scores. For early elementary students who have learning gaps, Bonvin et al. (2008) believed that grade retention is a precautionary measure. Bonvin et al. asserted that a student's reading score at the start of the school year may predict a student's academic success and may influence a retention decision at the end of the school year.

Roderick and Nagaoka (2005) found that government entities as well as school systems vacillate on the issue of grade retention. According to NASP's 2003 study, public schools' grade retention rates have increased in the last 25 years. The NASP found that 15% of students in the United States are retained each year. Denton (2001), Hauser, Pager, and Simmons (2000), and Jimerson (2001) found grade retention rates were increasing, but research does not support its positive effect on students' academic achievement.

Evans (2012) also found that schools' policies regarding grade retention were inconsistent. Evans asserted that discrepancies in policies resulted in different practices, especially in larger school districts. In larger school districts, Schwager, Mitchell, Mitchell, and Hecht (1992) concluded that "increased objectivity increased retention rates, while in smaller school districts it significantly decreased student retention rates" (p. 7). Conversely, Schwager et al. found the following: "increasing administrative work in the small districts tend [*sic*] to increase retentions...whereas in large school districts, increasing administrative decreases retentions" (p. 7).

Horn (2003) provided a historical perspective of how the standardized testing movement has evolved since the early 1970s and influenced grade retention policies. Horn found "the

number of such state-level testing rose from 1 in 1972 to 34 in 1985” (p. 30). Because students’ scores on high-stakes tests could determine whether they were promoted or retained, students experienced test anxiety (Horn, 2003). Consequently, students do not perform to their potential. Horn stated that government mandates have endorsed using students’ scores on high-stakes testing for promotion and retention decisions.

Pertinent in a grade retention decision are the beliefs of school administrators. Evans (2012) sought to address a gap in the literature and used a comparative design “to find relationships between elementary and middle school principals’ knowledge, influence, involvement, and beliefs towards grade-level retention rates and effectiveness in their schools” (p. 55). Evans had 198 elementary and middle school principals complete the PRIMES survey. The principals’ administrative experience ranged from less than two years to more than 12 years, with principals who had more than 12 years of experience providing 84 of the 198 (42.6%) participants the largest sample.

Evans (2012) used a four-point Likert scale and the following intervening variables to assess the principals’ dispositions regarding retention: “beliefs, knowledge, influence, and involvement” (p. 68). Evans determined that the principals’ involvement in grade retention practices yielded the highest mean rating ($M = 3.60$, $SD = 0.585$). Evans believed the lower standard deviation reflected the importance of the principals’ involvement in retention decisions. Conversely, their beliefs regarding retention yielded the lowest mean score ($M = 1.95$, $SD = 0.756$), which demonstrated that the principals had different beliefs regarding retention. Even though the principals did not agree with three of the four of Tomchin and Impara’s (1992) retention philosophies, Evans (2012) found that only 9 of the 188 “strongly agreed with the statement that students should never be retained under any circumstance” (p. 69).

In regard to the principals' overall knowledge of retention research, Evans (2012) found the principals were knowledgeable about retention research ($M = 2.83, SD = 0.800$). Evans also determined that the principals' involvement in retention decisions was high "as child-team study members ($M = 3.81, SD = 0.468$), resource providers ($M = 3.36, SD = 0.772$) and mediators in retention practices ($M = 3.62, SD = 0.509$)" (p. 71). As for their influence on grade retention decisions, the PRIMES survey results revealed that principals influenced grade retention decisions ($M = 2.96, SD = 0.765$). Evans also used a t test to determine if schools with principals who had positive views of grade retention had higher retention rates than schools with principals who had less favorable views of grade retention. Evans found statistically significant differences between the grade retention beliefs of principals at non-retention and retention schools.

Grade retention: A positive perspective. Bowman (2005) discussed why student retention was a viable option for struggling students. Bowman cited a student's emotional and social immaturity, low academic achievement, and inability to master grade-level standards as potential reasons for grade retention. Yet, Bowman believed that other factors were pertinent when making a grade retention decision. Bowman stressed the importance of schools providing students with quality classroom instruction and teachers with professional development opportunities.

In the state of Florida, Greene and Winters (2006) found that grade retention decisions based on standardized test scores improved students' reading achievement. Greene and Winters' study conducted an "across-years" comparison of third grade students' reading scores that were below Florida's test-based promotion policy standard. Starting in the 2002-03 school year, they found that the Florida Legislature required third grade students to achieve a Level 2 score on the Florida's Comprehensive Assessment Test (FCAT) reading assessment, with a Level 2 score

representing the second to last level of achievement. Greene and Winters had a control group and a treatment group. The control group included the 2001-02 third grade students' reading scores, which was prior to Florida's test promotion policy, while the treatment group included the 2002-03 third grade students' reading scores, which was under the test promotional policy. Greene and Winters compared the two groups' reading scores one and two years after their retention and promotion. They found that the retained third grade students (2002-03 school year, the treatment group) had higher reading achievement scores, 4.1 higher developmental scale scores (DSS) score in year one and a 40.9 higher score DSS score in year two, when compared to their promoted peers (2001-02 school year, the control group).

In Lowery's (2010) study, the author compared third through eighth grade as well as the eleventh grade (seven grade levels) students' reading, math, and science scores from the school year preceding and the school year following their retention. Initially, Lowery's sample included close to 70,000 participants in both the spring of 2007 and 2008, with seven grades having close to 10,000 students each. Prior to conducting the three subject area comparisons, Lowery provided the study's inclusion criteria: first, the students must have been retained at their current grade at the conclusion of the 2007 school year; second, the retained students must have repeated the same grade during the 2008 school year; third, the students must have taken the state's same subject test in the 2007 and 2008 school years. Due to this study's inclusion criteria, Lowery's study had 516 participants.

After converting scale scores to z scores, "difference z scores were compared between the 2007 z score and the 2008 z score for the same student, and analyzed separately for each subject" (Lowery, 2010, p. 67). The four test score comparisons between 2007 and 2008 revealed the following:

- (a) The students' z scores in reading increased 0.362, which represents a significant difference;
- (b) The students' z scores in math increased 0.397, which represents a significant difference; and,
- (c) The students' z scores in science increased 0.304, which represents a significant difference.

In all three comparisons, Lowery (2010) identified that the retained students demonstrated significant improvement on each subject test from 2007 to 2008.

Grade retention and its negative implications. Bonvin et al. (2008) asserted that grade retention gave poor academic achievers an additional year to meet the requisite standards. Yet, Jimerson (2001) found that student retention has more negative than positive effects. Jimerson's meta-analysis found that close to 2.5 million students are retained each year, which represents between 5 and 10% of the overall student population in the United States. Of the 169 retention cases evaluated, Jimerson concluded only 5% of these cases yielded long-term positive outcomes for students.

To illustrate grade retention's negative impact, Leckrone and Griffith (2006) concluded that retained students had more academic difficulties than students who received supplemental academic assistance in their next grade. In some cases, the incidence of grade retention and high-stakes testing had increased proportionately. Retention based on high-stakes testing may not benefit students.

Like Jimerson (2001), Holmes (1989) conducted a meta-analysis of 63 "controlled studies" on grade retention. Fifty-four studies yielded negative implications regarding students' long-term academic success while only nine showed positive implications. Picklo and

Christensen (2005) asserted that intervention, as opposed to grade retention, is the most effective way to prevent academic failure. Expanding upon the concept of student intervention, Smirk (2001) stressed that identifying at-risk students and implementing intervention plans were effective ways to reduce student retention.

Jimerson (1999) concluded that grade retention increases high school drop-out rates. Apparently, Jimerson also found a positive correlation between the amount of times a student is retained and his or her chances of dropping out of school. Grade retention also led to students developing poorer self-images and attitudes towards school (Holmes, 1989). With the exception of dropping out of school, students' failure rates in courses were a prevalent and troubling issue for schools (Alexander et al., 2003).

Additionally, grade retention has a negative effect on students' emotional health. Byrnes and Yamamoto (1985) found that students viewed grade retention as a punitive measure; therefore, students experienced negative emotions. Anderson, Jimerson, and Whipple's 2005 study, which encompassed a 20-year period, examined sixth grade students' perceptions of grade retention. When the authors asked the students to make a stressful events list, the sixth graders rated grade retention as the third most stressful event, with only the loss of a parent and loss of vision as being worse than grade retention.

In addition to the above stated negative outcomes, Jimerson et al. (2005) stated that grade retention has been linked to higher dropout rates for students and lower self-esteem for both children and adults. They suggested that schools develop teacher-led student support teams which address students' academic needs, create tutoring and mentoring programs for students, enlist parental support, and ensure regular education and special education teachers develop students, academically and socially.

Holmes (1989) compared the long-term performance of retained and non-retained students. Subsequent analysis of the two groups' performance on the standardized tests revealed the retained students scored lower than their non-retained peers. Bowman (2005) encouraged schools to examine their grade retention and social promotion policies. When comparing the basic skill levels of students who were socially promoted and students who were retained, Greene and Winters (2006) found the students who were retained had lower basic skill levels.

Grade retention: financial impact. Irrespective of its long- or short-term impact on students, grade retention financially impacts society. The additional cost for educating retained students costs taxpayers more than \$14 billion each year (Dawson, 1998). According to the Center for Policy Research in Education (1990), the annual cost of retaining one student is about \$4,050.

Bowman (2005) also stated that grade retention has negative financial ramifications. Bowman asserted that grade retention is not only a costly measure for school systems, but it also costs most students academically and socially. Picklo and Christensen (2005) questioned the practice of spending money on retaining students when intervention strategies were more appropriate, both financially and educationally.

Likewise, Holmes (2006) asserted that school systems' financial expenditures should be geared towards pre-retention intervention. Holmes stressed grade retention's impact on a school system's budget. McCollum, Cortez, Maroney, and Montes (1999) highlighted grade retention's annual financial impact: \$10 billion per year. McCallum et al. opined that this money could be utilized to improve students' academic achievement. In essence, money should be used for individualized student instruction, assistance, and assessment as a preventative measure as opposed to using it as a post-intervention strategy.

Jimerson (2001) concluded that retention benefits students who receive individualized instruction. In light of the above stated positive impact individualized instruction has on retained students' academic achievement, Jimerson encouraged lawmakers, teachers, and schools to create learning strategies that help students who fail to meet a state's academic standards. Leckrone and Griffith (2006) stated that schools should focus on improving their instructional strategies and interventions instead of students' scores on high-stakes tests.

Social Promotion

In contrast to grade retention, Picklo and Christensen (2005) indicated social promotion's goal is to keep low-achieving students with their peers. They concluded that social promotion has more negative than positive outcomes, but it has fewer negative outcomes than grade retention. Brown (2007) stated the implementation of high-stakes testing policies has eliminated social promotion and ensured that students and teachers will be responsible for meeting a state's curriculum standards.

In a 1999 document, the United States Department of Education described social promotion as "the practice of allowing students, who have failed to meet performance standards, and academic standards to pass on to the next grade with their peers instead of completing or satisfying the requirements," (p. 259). Thompson and Cunningham (2000) outlined the main arguments against social promotion:

It frustrates promoted students by placing them in the grades where they cannot do the work, sends the message to all students that they can get by without working hard, forces teachers to deal with under-prepared students while trying to teach the prepared, gives parents a false sense of their children's progress, leads employers to conclude that diplomas are meaningless, and dumps poorly educated students into a society where they

can't perform. (p. 3)

Apparently, social promotion was a common practice during the 1970s as students could avoid the associated negative stigma of grade retention (Bowman, 2005). With states adopting standards-based practices, states have abandoned this practice and replaced it with high-stakes testing as a promotion criterion (Horn, 2003). The NCLB states that student assessments must align with state standards, and they must measure students' grade-level academic achievement (Davison et al., 2002).

According to Johnson and Rudolph (2001) and Nagaoka and Roderick (2004), legislators have focused on ending social promotion in schools and are using students' scores on high-stakes test to make grade retention decisions. Brown (2007) stated that policymakers were correcting a major flaw in the educational system: social promotion. Furthermore, as part of the standards-based accountability reform movement, Brown found policymakers were establishing specific performance standards that must be met before a student can be promoted.

Summarily, Bowman (2005) stated that high-stakes tests will be the criterion for promotion and retention decisions. The author concluded that school systems have different grade retention and social promotion policies. However, Bowman found the practice of promoting academically unprepared students had garnered criticism; therefore, schools have adopted new promotion and retention policies.

Grade Retention: Predictors

Studies indicated the strongest predictors of retention were low early school academic performance, being male, low parent participation, and high mobility between schools (Karweit, 1999; McCoy & Reynolds, 1999). Stipek (2001) concluded that school failure during the early years (K-3) is an indicator of long-term negative academic, behavioral, and occupational

outcomes for students.

Grant (1997) asserted that students' developmental readiness can vary up to one year and still be considered normal. Yet, Grant stressed students who were performing in the average range were not always successful. Grant found there were many factors that contribute to the above stated one-year difference in students' developmental readiness. First, Grant discussed the impact of physical development. Apparently, children whose physical development is behind that of their peers, struggle in the areas of reading and writing. Second, students in grades K-3 who are struggling academically are more likely to have a learning disability when compared to students in higher grades (grades 4 and higher). For students who are struggling academically, Grant stated that additional time at a particular grade level would not help these students; they need appropriate learning support.

In regard to a student's chronological age, Grant found students who were born close to the enrollment deadline were not only the youngest student in their respective grade but many times the child ended up being retained. If a parent decides to retain a child because he or she is born late or right at the cut-off date, then the child could experience additional problems (Grant, 1997).

In addition to a students' physical and academic development, a student's school attendance may result in grade retention. Grant (1997) countered that a student's attendance should not influence a grade retention decision. Grant further stated that if a student has an attendance problem, then it is the parent's fault. Furthermore, students are more likely to be retained if their parents have a negative perception of education (Wang & Wang, 2007). Lastly, a lack of parental involvement in school activities and the number of times that a student changes schools also increases his or her chances of being retained (McCoy & Reynolds, 1999).

Socioeconomic status. Lorence and Dworkin (2006) discussed the impact students' demographics have on grade retention. In particular, the authors stated that non-Caucasian students, or students from certain racial and ethnic groups, are more likely to be retained than Caucasian students. Jacobs and Lefgren (2004) determined that one urban school district's third and sixth grade Hispanic and African American students' grade retention rates were high. Similarly, Meisels and Liaw (1993) cited a 1988 National Educational Longitudinal study which found Hispanic and African American students had disproportionate grade retention rates.

Furthermore, Lorence and Dworkin (2006) referenced Meisels and Liaw's 1993 study that found students from low socioeconomic backgrounds have a greater probability of being retained. As a caveat to the previous studies that found a link between a student's demographics and retention, Lorence and Dworkin (2006) stated that these studies "did not control for cognitive skill levels and levels of academic performance" (p. 1001). With this in mind, Lorence and Dworkin cited two urban school districts' results that contradicted Meisels and Liaw's 1993 study as the school districts' findings "indicated that neither race nor family economic background were related to the likelihood of being required to repeat an early elementary grade" (p. 1001).

Grant (1997) made similar conclusions regarding students' socioeconomic status and race. Grant stated that students who live in poverty, especially minority students, are often beset with higher grade retention rates. Unfortunately, these students' living conditions—lack of a home computer, uninvolved and uneducated parents—are not conducive to academic success. Instead of retaining these students for something that is out of their control, schools should provide these students with supplemental academic support and resources (Grant, 1997).

When considering grade retention's effect on minority and poor students when compared

to middle-class students, McCoy and Reynolds (1999) discussed the impact a student's race, socioeconomic status, and gender had on grade retention. For example, McCoy and Reynolds found that if a student is male, Hispanic or African American, and has low academic performance, then he or she has a greater chance of being retained.

Similarly, Hauser et al. (2000) investigated grade retentions impact on minority students. They found that males and minority students had higher grade retention rates than females and Caucasian students. Overall, Hauser et al. concluded that female/Caucasian student group had the greatest success in school while the African-American males had the highest grade retention rates. According to Karweit (1999), economically disadvantaged students are more likely to be retained. However, Jimerson (2001) did not find a correlation between grade retention rates and students who received free and reduced lunch.

Expanding upon the notion that students' backgrounds and demographics increase their chances of being retained, Karweit (1999) examined the predictive factors of grade retention and classified them into three categories: risk factors, protective factors, and not significant factors in retention. Some of the risk factors for grade retention included if the student was male, had a disability, was classified as other race ethnicity, attended a high poverty school, or was from a larger family. The protective factors for grade retention included if the student was Hispanic, had a mother with high education and income, was from an urban area, was motivated, and attended nursery school. The not significant factors for grade retention included if the student was black, was from a single-parent household, and had a mother with particular occupation (Karweit, 1999).

In regard to parental influences on grade retention, Wilson and Hughes (2009) discussed the home environment's impact on grade retention, especially the parental aspect. Wilson and

Hughes cited a longitudinal study of 24 African-American parents of preschool through first grade students in which the parents deferred to the school. In their study, Wilson and Hughes asserted that parental naivety as it relates to the grade retention process could impact a grade retention decision. The authors concluded that parents who have a positive and uninformed perception of a school are willing to agree with a grade retention decision. This philosophy eschews shared responsibility for a student's education (Wilson & Hughes, 2009).

In addition to parents, teachers can influence a grade retention decision (Smith & Sheppard, 1988; Stipek & Byler, 1997; Tomchin & Impara, 1992). Smith and Sheppard (1988) found that classroom teachers depended on their professional experiences rather than empirical evidence when making grade retention decisions. Stipek and Byler (1997) determined that teachers who utilized "teacher-directed approaches" (p. 314) endorsed grade retention when students were unable to master basic concepts. Tomchin and Impara (1992) concluded that primary grade teachers (K-3) supported grade retention more than intermediate, fourth through sixth grade teachers. Tomchin and Impara found that primary grade teachers believed that mastery of grade-level concepts and skills were a prerequisite for advancing to the next grade level.

Grade retention related to gender. When determining if grade retention impacts males more than females, Grant (1997) found that boys had higher retention rates than girls. Similar to Grant (1997), Nagaoka and Roderick (2004) and Roderick (1994) concluded that boys were more likely to be retained than girls. The Texas Education Association (TEA) (2006) compared grade retention rates among Texas public school students based on gender and grade level. TEA's grade-level comparisons of retention rates based on gender revealed that male grade retention rates were higher than those of females at all grade levels (K-12). Further analysis of

the grade-level comparisons based on gender revealed that the largest discrepancy was at the ninth-grade level, with males having a 5.6% higher retention rate. TEA also concluded that first-grade males had the highest grade retention rate (7.5%) while sixth-grade females had the lowest (1.0%). Finally, Texas school districts retained more than twice as many males than females at the sixth-grade level (3354 = males, females = 1547).

Karweit (1999) utilized student data from 120 school districts from four regions to compare male and female students' grade retention rates, and she concluded that male students' grade retention rates were higher than their female peers. Within the analysis for predictors of grade retention, Karweit's sample was comprised of 8,695 students, with 4,457 males and 4,238 females. Karweit found that "the odds for being retained for boys are 979/3478 or .28. The odds for females are 625/3613 or .17" (p. 19). Initially, Karweit concluded that boys were 1.64 times more likely to be retained than girls, but inclusive of other factors, a subsequent reduction to 1.33 was determined.

According to Thomas and Stockton (n.d.), the Louisiana Department of Education (2001) analyzed its Student Information System (SIS) data from 1997-2001 in grades K-12 and found that male students were more likely to be retained than female students, and students on free lunch were twice as likely to be retained as students not receiving any food services (p. 8).

The Transition from Traditional-Based to Standards-Based Assessment

Because high-stakes tests determine if students and schools are meeting achievement benchmarks, schools have transitioned from traditional based to standards-based assessments (Nichols, 2007). For example, Colby (1999) found that teachers were utilizing district-developed tests (common assessments and benchmarks) to assess students' progress on state standards. Craig (2012) stated, "Standards-based report cards provide information on student

mastery of content and skills within the strands, standards, and indicators of the established curriculum and inform students of their progress along a continuum of proficiency” (p. 7). Craig asserted there must be a correlation between student assessments and instructional activities. Standards-based grading is a reflection of what a student learns rather than what he or she earns (Brookhart, 2011). The student’s ability to master a criterion-referenced standard is a significant component of standards-based grading (Paepflow, 2011).

Clearly, the NCLB Act’s accountability standards have prompted school systems and teachers to reexamine their assessment procedures and instructional strategies (Craig, 2012; Paepflow, 2011). With this in mind, Guskey (2001) concluded that a standards-based grading system allowed teachers to identify students’ learning standards and to adapt students’ instructional needs. Guskey stated that federal and state education agencies determine students’ learning standards.

Guskey (2007) further stated that with standards-based assessments, students’ grades are based on their ability to meet each individual standard, which is different from grading students’ knowledge on the whole topic. Guskey asserted the inherent challenge of a standards-based grading system involved clarifying the purpose of each instrument. Scriffiny (2008) stressed the importance of students participating in standards-based reporting, with each student understanding his or her grade level’s standards and expectations.

The passage of the NCLB ensured that low-achieving students received standards-based reform (Picklo & Christensen, 2005). Standards-based reform refers to “plans in place” as it improves educational quality by setting curriculum-based standards and holding schools, educators, and students accountable for meeting these standards (McDonnell, McLaughlin, & Morrison, 1997). Thurlow and Johnson (2000) stressed that high-stakes tests were critical

components of standards-based reform as all K-12 public school districts are held to the same accountability standard.

Prior to implementing a standards-based grading approach, Colby (1999) stated a school system needs to address any curriculum issues. “When high-quality standards drive instructional and assessment decisions and when teachers report on how well each student progresses according to each standard, then standards-based grading system will become essential” (Colby 1999, p. 52). Scriffiny (2008) concluded that standards-based grading allowed teachers to individualize student instruction, including students who required alternative assessments and students who needed more challenging tasks. Scriffiny also highlighted the standards-based grading system’s ability to provide students and teachers with comprehensive feedback.

Haptonstall (2010) compared the Colorado Student Assessment Program (CSAP) scores of students in a traditional grading system and the CSAP scores of students in a standards-based grading system. Haptonstall found that the correlation was stronger between the students’ CSAP scores and the standards-based grading system. Haptonstall also determined that students in a standards-based grading system had higher mean scores on the CSAP in the following areas: (a) reading, (b) writing, (c) math, and (d) science. He recommended that schools implement standards-based grading across all grade levels and subjects and that school systems encourage teachers to eliminate any grading criteria that considers non-academic items. Haptonstall asserted that the stronger correlations between the standards-based grading system and students’ CSAP scores, as well as the higher CSAP scores of the students in a standards-based grading system, compelled schools to align standards-based grading practices and standards-based teaching approaches.

Although researchers acknowledged the connection between standards-based grading

assessments and standards-based instructional practices (Harlen & Crick, 2003; Leckrone & Griffith, 2006), they found that their impact and original intent did not align. In order to prepare students for high-stakes tests, the authors found teachers reinforced concepts and skills, which precludes higher-order thinking and restricts the curriculum. Students, parents, and teachers are discouraged when a student's classroom performance is irrelevant and a single test determines a grade retention decision (Leckrone & Griffith, 2006).

Because high-stakes tests are critical components in determining the success of students, teachers need to utilize performance-based assessment and instructional strategies (Leckrone & Griffith, 2006). The U.S. Department of Education (1999) concluded that social promotion and retention were not viable options for improving students' academic achievement. When students fall behind academically, achievement and learning gaps occur (Craig, 2012; Guskey, 2001). A potential solution is to retain students who are unable to demonstrate proficiency on standardized tests (Wang & Wang, 2007). Yet, grade retention does not address low-achieving students' academic and social needs because low-achieving students require a learning environment that addresses their learning needs (Picklo & Christensen, 2005).

Summary

Cannon and Lipscomb (2011) stressed that grade retention is a divisive issue within education. Silberglitt et al. (2006) concluded that grade retention had more negative than positive outcomes. Researchers found that when non-academic factors were included in grading, a clear discrepancy existed between students' classroom achievement and their performance on high-stakes tests (Brennan, Kim, Wenz-Gross, & Siperstein, 2001). The above stated discrepancy between students' classroom performance and their standardized test scores has resulted in teachers and administrators developing multiple assessments that correspond to the

grade-level learning standards students must master.

Because the target school district replaced its traditional grading system with a standards-based grading system during the 2007-08 school year, the standards-based grading system's ability to reduce the target school system's grade retention rate had not been determined.

Scriffiny (2008) discussed and stressed the impact of standards-based assessments. Scriffiny found that standards-based assessments enabled educators to individualize student instruction and assessment.

There are factors that increase a student's probability of being retained. For example, McCoy and Reynolds (1999) concluded that students had a greater chance of being retained if they were male, Hispanic or African American, and had low academic performance. For gender, Karweit (1999) found boys were 1.33 times more likely to be retained than girls. Lorence and Dworkin (2006) found that non-Caucasian students were more likely to be retained. Grant (1997) investigated the impact of poverty on grade retention. Grant found that students from poor socioeconomic backgrounds had higher grade retention rates than students from middle class and upper middle class backgrounds.

When examining grade retentions' impact on students' academic achievement, Jimerson (2001) concluded that grade retention had an overall success rate between 5 and 10%. In addition, Jimerson found that grade retention had a positive effect on students' academic achievement in the year that followed the retention; however, longitudinal analyses (several years later) revealed that students either remained the same or they regressed academically. Similarly, other studies did not support the long-term benefits of grade retention (Holmes, 1989; Nagaoka & Roderick, 2004; Wang & Wang, 2007). Holmes' (1989) meta-analysis revealed that grade retention had negative implications in 54 studies and positive implications in only nine

studies.

Evans (2012) revealed that principals' dispositions were critical components in a grade retention decision. In particular, principals' levels of involvement in grade retention ($M = 3.60$, $SD = 0.585$), the principals' involvement in retention decisions as members of their schools' child study teams ($M = 3.81$, $SD = 0.468$), in their roles as resource providers ($M = 3.36$, $SD = 0.772$) and as mediators in their schools ($M = 3.62$, $SD = 0.509$). Evans also found that principals' involvement in schools' grade retention policies resulted in the study's highest mean rating ($M = 3.60$, $SD = 0.585$). This study addressed that gap in the literature by evaluating the number of students retained under a traditional grading system compared to a standards-based grading system in one of Georgia's public school districts.

CHAPTER THREE: METHODOLOGY

Introduction

This study utilized a non-experimental causal-comparative retrospective research design. Williams (2007) stated that in a causal-comparative research design “the researcher examines how the independent variables are affected by the dependent variables and involves cause and effect relationships between the variables” (p. 66). Similarly, Ellis and Levy (2009) found that causal-comparative research determines if there is a cause-effect relationship between the independent(s) and the dependent(s) variables. Schenker and Rumrill (2004) noted that “causal-comparative designs generally involve the use of pre-existing or derived groups to explore differences between or among those groups on outcome or dependent variables” (p. 117).

Because the researcher utilized preexisting groups’ data to determine whether the implementation of a standards-based grading system impacted student retention rates when compared to a traditional grading system, a causal-comparative research design was appropriate. Schenker and Rumrill (2004) stated, “Causal-comparative designs generally involve the use of pre-existing or derived groups to explore differences between or among those on outcome or dependent variables” (p. 117). Generally, the researcher does not manipulate or influence variables in causal-comparative designs (Schenker & Rumrill, 2004). Similar to this study, causal comparative designs can include more than one independent variable; moreover, “the independent variables are (a) categorical and (b) not experimentally manipulated” (Schenker & Rumrill, 2004, p. 118).

Ary, Jacobs, Razaviah, and Sorensen (2006) also described the characteristics of a causal-comparative, non-experimental research design by citing “the designation *ex post facto*” (p. 356) which means the variable of interest occurred prior to the study. Next, Ary et al. found groups

are non-randomized in a causal-comparative, non-experimental design. The two groups differ on both an independent and dependent variable, and the researcher is attempting to determine the cause of this discrepancy (Ary et al., 2006).

While causal-comparative and non-experimental research designs examine the relationship between the independent variable and the dependent variable, experimental studies allow the researcher “to obtain much more convincing evidence for causal (functional) relationships among variables than can be obtained with ex post facto studies” (Ary et al., 2006, p. 357). Ary et al. (2006) found the researcher manipulated the independent variable (the why) when conducting experimental research. By manipulating the independent variable, the researcher is able to make “strong inferences about causal relationships” (p. 358). In contrast, the researcher does not control the independent variable when using a causal-comparative design; therefore, “there is less basis for inferring a causal relationship between X and Y” (Ary et al., 2006, p. 358).

For this study, the researcher determined that a causal-comparative research design was appropriate because the researcher included preexisting groups’ data, assigned participants to groups (non-randomized), and did not manipulate the independent variables used in this study: the traditional grading system (2007-08 school year) and the standards-based grading system (2008-09 school year). Further, the researcher compared student retention rates under a traditional grading system and standards-based grading system. Next, the researcher determined which grading system decreased the target school system’s grade retention rate.

To improve their assessment strategies and increase students’ mastery of grade-level standards, school systems have transitioned from a traditional grading system to a standards-based grading system. NCLB’s mandate that 100% of a public school’s students demonstrate

grade-level proficiency by 2014 has prompted this shift to a standards-based grading system. The researcher sought to determine if the standards-based grading system was a valid and reliable predictor of student success.

The purpose of Chapter Three was to provide an overview of this study's methodology. This chapter includes the following: (a) a description of the sample population, (b) the instruments utilized for data collection, and (c) the data collection procedures.

Design

Causal-comparative research is defined "as a type of research investigation that attempts to discover the direction and magnitude of the relationship between variables" (Gall, Gall, & Borg, 2007, p. 636). Similarly, Ary et al. (2006) concluded that a causal-comparative design investigates between two variables. Gall et al. (2007) described the characteristics of the causal-comparative research design:

To identify cause-and-effect relationships by forming groups of individuals in whom the independent variable is present or absent... and then determining whether the groups differ on the dependent variable. The critical feature of causal-comparative research is that the independent variable is measured in the form of categories. (p. 306)

In this study, the researcher used a causal-comparative research design to examine two groups with different independent variables to determine the effect of different grading systems on retention rate. Specifically, the researcher determined whether a standards-based grading system or a traditional grading system had an effect on third and fifth grade students' grade retention rates.

This study's sample size, approximately 1600 students with both groups being equal in size, was large enough to detect significant differences between the two groups. The researcher

used a convenience sampling procedure because the sample student populations were accessible to the researcher and a part of pre-existing groups (Gall et al., 2007). The researcher controlled for variability by including students who attended schools that had similar demographics and were part of the same county school system.

While the researcher limited the scope of this study to upper elementary, third through fifth grade students, the researcher believes this study's findings could extend beyond these grades to middle school because the state of Georgia mandates that eighth grade students pass the CRCT. The students in these grades take both the CRCT and a summative test for each grade level, which compare the academic performance of Georgia's public school students at each grade level. Although the state of Georgia tests students (in both reading and math) in all three grades, the Georgia State Board of Education uses the CRCT results for different purposes. Third grade students must pass the reading component of the CRCT (at the meets expectations level) to be promoted to fourth grade, but their math score is an informational piece to guide instruction. Conversely, fourth grade students' reading and math CRCT scores are advisory, meaning passing scores in these areas are not required to be promoted to fifth grade. In the fifth grade, students must pass both the reading and the math CRCT tests (GaDOE, 2008).

The increased promotion criteria could increase the probability of grade retention for fifth grade students when compared to the third grade students. Pertinent was whether the relative grade retention rates change significantly based on grading systems, so it is important to control for this. This study's goal was to determine if relative grade retention rates changed as a result of transitioning to a standards-based grading system.

This study incorporated adjacent (in time) school years. Including adjacent school years was important, as many factors related to education are time-related. For example, what

occurred in 1930 or 1970 is often quite different than what occurs in the present. Because it was not possible to compare students in the same year under two different grading methods, the researcher included two successive years of student data from identical schools. Including successive years allowed the researcher to compare student performance under a traditional grading system and student performance under a standards-based grading system. Therefore, each grading system was a unique data set. The transition from the traditional grading system to the standards-based grading system occurred at the conclusion of the 2007-08 school year. With this in mind, the researcher compared grade retention rates of the 2007-08 school year (traditional grading system) and during the 2008-09 school year (standards-based grading). The four schools included in this study implemented the standards-based grading system during the 2008-09 school year.

Questions and Hypotheses

The following research questions guided this study:

RQ1. Is there a difference between the third grade students' grade retention rates with a traditional grading system and their grade retention rates with a standards-based grading system?

RQ2. Is there a difference between the fifth grade students' grade retention rates with a traditional grading system and their grade retention rates with a standards-based grading system?

RQ3. Is there a difference among the third and fifth grade female students' grade retention rates as a result of transitioning from a traditional based grading system in 2007-08 to a standards-based grading system in 2008-09?

RQ4. Is there a difference among the third and fifth grade male students' grade retention rates as a result of transitioning from a traditional based grading system in 2007-08 to a standards-based grading system in 2008-09?

Null Hypotheses

H₀₁: There will not be a significant difference between the 2007-08 third grade group's grade retention rate with a traditional-based grading system and the 2008-09 third grade group's grade retention rate with a standards-based grading.

H₀₂: There will not be a significant difference between the 2007-08 fifth grade group's grade retention rate with a traditional-based grading system and the 2008-09 fifth grade group's grade retention rate with a standards-based grading.

H₀₃: There will not be a significant difference among the third and fifth grade female students' grade retention rates as a result of transitioning from a traditional based grading system to a standards-based grading system.

H₀₄: There will not be a significant difference among the third and fifth grade male students' grade retention rates as a result of transitioning from a traditional based grading system to a standards-based grading system.

Participants

The participants in this study were third and fifth graders who attended four rural elementary schools in northeast Georgia during the 2007-08 and 2008-09 school years. The researcher included every third and fifth grade student who attended the four schools during the above-stated school years. The researcher also examines whether gender affects grade retention.

As stated above, this study's participants were third and fifth grade students who were assessed with a traditional-based grading system during the 2007-08 school year and students from identical grades who were assessed with a standards-based grading system during the 2008-09 school year. It is important to note the researcher conducted comparisons at each grade level: the third grade retention rates for the 2007-08 and 2008-09 school years and the fifth grade

retention rates for the 2007-08 and 2008-09 school years. Because the researcher compared two different years of grade retention data for two distinct grade levels, each grade level.

The researcher included 70 students from each grade level. Therefore, both the 2007-08 sample (traditional-based grading system) and the 2008-09 sample (standards-based grading system) each had 840 students who attended four schools. While the 2007-08 sample and the 2008-09 had identical sample sizes, the researcher included students in both the 2007-08 sample and the 2008-09 sample. For example, the researcher included 2007-08 third grade students who will be included in the study as a 2008-09 fourth grade students. Hence, there was a discrepancy between this study's sample sizes and the overall number of students. The four elementary schools in this study were Title I schools. For example, the researcher compared the 2007-08 fourth grade students and 2008-09 first grade students' grade retention rates. The researcher followed an identical procedure for the third and fifth grade students.

Setting

The researcher conducted this study in a rural county located in northeast Georgia. The four Title 1 elementary schools included in this study were part of the same rural Northeast Georgia school district. The school district is comprised of eight elementary schools, three middle schools, and two high schools. The racial demographics for the four schools in this study are as follows: White = 80%, Hispanic = 12%, Black = 5%, and other/multiracial = 3%. This Northeast Georgia County had a total population of 61,620 and a median annual household income of \$52,029 (U. S. Census Bureau, 2008). The socioeconomic status of the students who attend the four schools was slightly below the national average as over 50% of students were eligible for free or reduced-price lunches.

During the 2007-08 school year, the four elementary schools utilized a traditional-based

grading system. The traditional grading system was a scale grading system. For example, a 93-100 = A, 85-92 = B, 78-84 = C, 70-77 = D, and below a 70 = F. At the beginning of the 2008-09 school year, the four elementary schools implemented a standards-based grading system, with common curriculum maps, pacing guides, rubrics, and performance indicators. The target school system established curriculum teams for each grade level. These teams periodically reviewed standards, instructional methods and pacing, and assessments to ensure that teachers were teaching the standards-based learning standards.

Georgia Department of Education (2011) provided the following guidelines regarding promotion and grade retention:

Beginning with the current 2004-05 school year, the implementation of the Georgia Promotion, Placement, and Retention law (O.C.G.A. §§ 20-2-282 through 20-2-285) and State Board of Education Rule (160-4-2-.11) will take effect for students in Grade 5. All fifth grade students must achieve grade level scores on the Georgia Criterion Referenced Competency Tests (CRCT) in Reading and Mathematics in order to be promoted to the sixth grade. (p. 21)

Instrumentation

Georgia Criterion-Referenced Competency Test

“The CRCT program is designed to measure how well students have acquired the skills and knowledge prescribed in the Georgia Performance Standards (GPS) and the Quality Core Curriculum (QCC)” (Georgia CRCT Guide, 2008, para. 1). The CRCT is a summative assessment that assesses students’ knowledge and skills proficiency levels on state standards (GaDOE, 2008). The Georgia State Department of Education (2008) described the CRCT in the following manner:

A series of tests, designed to measure how well students acquire the skills and knowledge described in the Georgia Performance Standards (GPS). The assessments yield information on academic achievement at the student, class, school, system, and state levels. This information is used to diagnose individual student strengths and weaknesses as related to the instruction of the GPS, and to gauge the quality of education throughout Georgia. (para. 1)

The GaDOE (2011a) stated that CRCT testing began in 2000, and it “is designed to measure how well students acquire the skills and knowledge described in the Georgia Performance Standards (GPS)” (p. 8). Students in grades 1-8 are testing in English, language arts, and math; in addition, students in grades 3-8 also take science and social studies tests (GaDOE, 2008). The GaDOE (2011a) utilized a GPS scoring system in which a student’s raw score (the number of correct answers) is converted to a CRCT scale score. The GaDOE stated, “Scale scores are comparable across all test forms administered for the same content/grade” (p. 1). Therefore, a student who scores 850 on one test administration is identical to another student who scores an 850 on grade level administration.

The GaDOE reports scaled CRCT scores ranging in value 650 to 900 and above, with different lower and upper limits (lowest possible score/highest possible score) for each grade level and content area: (a) does not meet standard = below 800; (b) meets the standard = 800-849; (c) exceeds the standard = 850 and above (GaDOE, 2008). While the GaDOE used the students’ CRCT score, the required raw score varied based on the subject and the student’s grade level. The following were the scale cut/raw cut scores for the 2011 CRCT test administration:

- (a) Reading = Grade 3—meets = 800/18, exceeds = 850/33; Grade 4—meets = 800/21, exceeds = 850/35; Grade 5—meets = 800/18, exceeds = 850/33

- (b) English/Language Arts = Grade 3—meets = 800/24, exceeds = 850/42; Grade 4—meets = 800/27, exceeds = 850/44; Grade 5—meets = 800/24, exceeds = 850/24
- (c) Math = Grade 3—meets = 800/34, exceeds = 850/48; Grade 4—meets = 800/35, exceeds = 850/50; Grade 5—meets = 800/31, exceeds = 850/47. (GaDOE, 2011, p. 2)

The CRCT is not a norm-referenced test. However, Georgia school systems may test students in grades three through five and six through eight in reading, mathematics, science, and social studies. If schools decide to test these students, students will take the Iowa Tests of Basic Skills (ITBS/A). There was not a testing window for the ITBS/A (GaDOE, 2011a).

In this study, the researcher compared the CRCT scores of students who received instruction under a traditional-based grading system and students who received instruction under a standards-based grading system. From a technical standpoint, it is important to determine an assessment's reliability and validity. Ary et al. (2006) concluded that reliability is the degree to which an instrument measures and produces reliable results. When determining reliability for a criterion-referenced test, Ary et al. underscored the importance of participants demonstrating proficiency during two identical tests. The authors identified and described "administering two equivalent forms of the test, or the same test, on two occasions and finding the consistency of the decisions reached" (Ary et al., 2006, p. 272). The GaDOE (2008) stressed that no test is infallible; therefore, each student's scaled score had a standard error of measure (SEM). SEMs for the spring 2008 CRCT were as follows: (a) third grade—reading = 2.46, English/language arts = 2.87, math = 3.10; (b) fourth grade—reading = 2.42, English/language arts = 2.77; math = 3.18; (c) fifth grade—reading = 2.60, English/language arts = 2.74, math = 3.19.

In addition to the SEM, the GaDOE (2008) provided Cronbach's alpha coefficient scores for the above-stated CRCT tests. Cronbach's alpha measures the strength of the relationship

between items on a test (Tavakol & Dennick, 2011). Cronbach's alpha also "describes the extent to which all the items in a test measure the same concept or construct" (Tavakol & Dennick, 2011, p. 53). Cronbach's alpha coefficient scores for the CRCT were as follows: (a) third grade—reading = 0.89, English/language arts = .90, math = .93; (b) fourth grade- reading = 0.89, English/language arts = 0.90, math = 0.91; (c) fourth grade reading = 0.86, English/language arts = 0.89, math = 0.92 (GaDOE, 2008).

Validity is the other key component of technical quality in testing, which begins with the purpose of the assessment and continues through item writing and review. The GaDOE (2008) stated that the CRCT's purpose is to assess students' ability to develop the skills and learn the requisite material "described in the Georgia Performance Standards (GPS) and the Quality Core Curriculum" (p. 1). Specifically, the CRCT assesses first through eighth grade students' academic achievement in the following areas: (a) reading, (b) English/language arts, and (c) mathematics; in addition, it assesses third through eighth grade students in science and social studies.

In addition to assessing students' performance in the above stated subject areas, the GaDOE (2008) identified that the CRCT determines whether Georgia's school systems, schools, and classrooms are meeting the state's and NCLB's achievement benchmarks. Regarding the test development process, and in order to determine content validity, the GaDOE (2008) acknowledged that committees of Georgia public school teachers "are formed to review the curriculum and establish which concepts, knowledge, and skills will be assessed and how they will be assessed" (p. 2). After developing the items, curriculum specialists and committees review all of the test items for overall quality and clarity, content coverage and appropriateness, curriculum alignment, and grade appropriate stimuli, with an emphasis on higher-order thinking

skills. In addition, there should be one clear correct answer with appropriate, relevant, and reasonable distractors.

During the item development process, GaDOE (2008) indicated that a significant emphasis is placed on developing a test that represents all of Georgia's students and is without bias. To ensure that the CRCT meets the highest standards of technical quality, the Testing Division meets with an independent panel of experts—Georgia's Technical Advisory Committee (TAC)—on a quarterly basis. TAC members are experts in the field of educational measurement who review all aspects of the test development and implementation process on a continual basis. The Georgia Department of Education concludes that the CRCT tests are both reliable and valid.

Standards-based report cards. The four schools in this study, along with the county school system, used an identical standards-based report card. The county school system implemented the standards-based assessment and grade reporting during the 2007-08 school year. Therefore, all of the students in this study were assessed with identical rubrics, assessments, standards, and performance indicators. Each one of the county's schools utilizes a common-standards report card.

The progress indicators on the SBRC were as follows: DNM (does not meet = less than 50%), IP (in progress = 50-90%), M (meets = 90-100%), and E (exceeds = above grade level expectations). The SBRC allows teachers to identify students who have not demonstrated mastery of a particular standard. Through this identification, teachers make students aware of the standards they have mastered and the standards that need more instruction. This is recorded on their report cards (i.e. second grader knows division and multiplication, even though they are only required to add and subtract). Teachers measured and used a time sheet to measure students' performance. The students' scale scores range between 650 and 910 (GaDOE, 2011a).

Procedures

Prior to collecting and analyzing data, the research conducted a comprehensive review of the related literature. Next, the researcher completed all of Liberty University's and the target school system's procedures for collecting student data. First, the researcher completed all of the Liberty University's IRB approval process, including the university's IRB forms and the target school system's permission to collect student data form. After completing Liberty University's IRB approval process (see Appendix A), the researcher presented the proposed study to the school district's coordinator of elementary education.

After the school district's coordinator of elementary education approved the study, the researcher contacted the target school district's director of curriculum and accountability and requested. Next, the researcher requested traditional report card data, standards-based report card data, and CRCT results as well as grade retention data for third and fifth grade students who attended the four elementary schools during the 2007-08 and 2008-09 school years. The school system's data portal, Infinite Campus, provided report card and grade retention data. The school district's director of curriculum and accountability had access to Infinite Campus and provided the researcher with Excel spreadsheets that included all of the above-stated data. The director of curriculum and accountability also de-identified all student data. There was a distinct spreadsheet for each grade level; therefore, the researcher received a spreadsheet for the third grade students and one for the fifth grade students. Within each spreadsheet, each de-identified student had a distinct row. In total, there were four Excel spreadsheets: two (third and fifth grade levels) for the 2007-08 school year and two (third and fifth grade levels) for the 2008-09 school year.

The researcher secured the collected data in a locked filing cabinet, to which only the

researcher has a key. Because central office personnel de-identified all student data, the researcher did not obtain parental consent to participate. To comply with Family Educational Rights and Privacy Act (FERPA) regulations, the researcher did not provide any identifying information concerning the school district, the four elementary schools, the teachers, and the students.

The researcher compared the 2007-08 grade retention rates with a traditional grading system and 2008-09 grade retention rates with a standards-based grading system. The third and fifth grade students included in this study attended the four elementary schools during the 2007-08 and 2008-09 school years and participated in CRCT testing in both of these school years.

Comparing the student data from the 2007-08 and 2008-09 school years in this study was appropriate because the four elementary schools utilized a traditional-based grading system for the 2007-08 school year and a standards-based grading system for the 2008-09 school year. In addition to comparing the third and fifth grade retention rates under a traditional grading system (2007-08 school year) and a standards-based grading system (2008-09 school year), the researcher compared the third and fifth grade students' 2007-08 (traditional-based grading system) and 2008-09 (standards-based grading system) CRCT scores.

This study had two independent variable groups. The first independent variable group included students who were assessed under a traditional grading system (2007-08 school year). The second independent variable group included students who were assessed under a standards-based grading system (2008-09 school year). The dependent variables were the third and fifth grade students' grade retentions at four rural Title 1 elementary schools in northeast Georgia that used different grading systems during the 2007-08 and 2008-09 school years.

By comparing the 2007-08 and 2008-09 grade retention rates, the researcher determined

if grade retention rates increased, decreased, or remained unchanged. The target school system's elementary testing coordinator provided the requisite data for this comparison and coded retained students as a "0" and promoted students as a "1".

Data Analysis

The researcher provided a graphic representation of the comparisons. The horizontal axis represented the 2007-08 and 2008-09 school years whereas the vertical axis represented the number of students retained during each school year. The researcher also investigated (a) traditional grading system retention rates, (b) standards-based grading retention rates, and (c) student CRCT scores. The researcher used a Fisher's Exact test to determine if there was a difference between students' retention rates under a traditional grading system and a standards-based grading system. The researcher conducted same grade-level comparisons for distinct years. For example, the researcher compared the third grade retention rate for 2007-08 school year (traditional grading system) and the third grade retention rate for the 2008-09 school year. The researcher conducted similar comparisons for the fifth grade level. In addition, the researcher compared retention rates for all grade levels under a traditional grading system and a standards-based grading system.

The researcher utilized a Fisher's Exact test to determine if there was a statistically significant difference between the students' retention rates under a traditional grading system and a standards-based grading system. This study also determined if there was a difference in grade retention rates based on gender. The inclusion of categorical variables such as grading systems (traditional, standards-based) and gender (male, female) and the utilization of a 2 x 2 contingency table prompted the researcher to consider the Pearson's chi-square statistic (Howell, 2010). The chi-square statistic determines differences between categorical variables (Bolboacă,

Jäntschi, Sestraş, Sestras, & Pamfil, 2011; Howell, 2010).

The use of the Pearson's chi-square statistic presupposes that a sample population is large enough ($n > 40$), that cell values are more than five ($n > 5$), and that the two categorical variables are independent from each other (Bolboacă et al., 2011; Howell, 2010; Warner, 2013). If cell values do not exceed five ($n > 5$), then researchers should utilize a Fisher Exact test rather than a Pearson chi-square test (Bolboaca et al., 2011; Lin, McClintock, & Williamson, 2011). The Fisher's Exact test calculates a probability value, represented as p , for the differences between independent variables (Bolboacă et al., 2011; Lin et al., 2011). The probability value the Fisher's Exact test calculates is generally more conservative than that of the Pearson chi-square statistic (Lin & Yang, 2008).

This study compared the grade retention rates of students prior to and after the implementation of a standards-based grading system. The threat of maturation was a concern in this study because student retention may have occurred at lower grade levels for both maturation as well as academic achievement issues (Karweit, 1999).

CHAPTER FOUR: FINDINGS

School reform has prompted schools to transition from a traditional grading system to a standards-based grading system (Craig, 2012; Guskey, 2001; Nichols, 2007). Schools across America are using a standards-based grading system to ensure that students master grade-level content before they promote them to the next grade (Colby, 1999). The NCLB Act mandates that school districts and their schools meet yearly AYP benchmarks established by the governing state. As stated by the GaDOE (2011a), AYP is one of the foundations of the NCLB Act of 2001. AYP measures the yearly academic achievement of a school's students based on predetermined benchmarks. The state of Georgia uses the CRCT as one of its benchmarks for assessing a school's effectiveness. In an effort to improve their assessment strategies and increase students' master grade level content, school systems have transitioned to a standards-based grading system.

As school systems make this transition, they need to determine if a standards-based grading system is a valid and reliable way to measure student success. Previous research suggested that grade retention generally results in negative outcomes for students (Jimerson, 2001; Xia & Glennie, 2005). Even students who demonstrated minimal academic gains in the year following their grade retention tend to lose these gains over time (Hauser, 1999; Holmes, 1989; Karweit, 1991; NASP, 1998; Thompson & Cunningham, 2000).

The purpose of this study was to determine whether a traditional or a standards-based grading system improved third and fifth grade retention rates. The target schools' transition from a traditional system to a standards-based grading system underscored the importance of this inquiry. The significance of this study was its ability to provide quantitative data regarding the standards-based grading system's impact on grade retention rates. To determine the effect of the

standards-based grading system on grade retention rates, the researcher compared the effect of the traditional grading system and a standards-based grading system on the target school district's grade retention rates. This study's results could persuade or dissuade the future implementation of a standards-based grading system. Grade retention rates have a direct effect on a school's AYP (GaDOE, 2010). Schools that are unable to meet their achievement benchmarks, as measured by AYP, could lose federal funding.

Research Question and Hypotheses

This study used the following research questions to determine if there was a difference between either a traditional grading system or a standards-based grading system and retention rates for third and fifth grade students to identify any associations based on gender.

RQ1. Is there a difference between the third grade students' grade retention rates with a traditional grading system and their grade retention rates with a standards-based grading system?

H₀₁: There will not be a significant difference between the 2007-08 third grade students' grade retention rates with a traditional-based grading system and the 2008-09 third grade students' grade retention rate with a standards-based grading.

RQ2. Is there a difference between the fifth grade students' grade retention rates with a traditional grading system and their grade retention rates with a standards-based grading system?

H₀₂: There will not be a significant difference between the 2007-08 fifth grade students' grade retention rates with a traditional grading system and the 2008-09 fifth grade students' grade retention rates with a standards-based grading.

RQ3. Is there a difference among the third and fifth grade female students' grade retention rates as a result of transitioning from a traditional grading system in 2007-08 to a standards-based grading system in 2008-09?

H₀₃: There will not be a significant difference among the third and fifth grade female students' grade retention rates as a result of transitioning from a traditional grading system to a standards-based grading system.

RQ4. Is there a difference among the third and fifth grade male students' grade retention rates as a result of transitioning from a traditional grading system in 2007-08 to a standards-based grading system in 2008-09?

H₀₄: There will not be a significant difference among the third and fifth grade male students' grade retention rates as a result of transitioning from a traditional grading system to a standards-based grading system.

Data Summary

Tables 1 and 2 provide the student demographic information for the third and fifth grade students who attended the four Title 1 schools during the 2007-08 and 2008-09 school years.

Table 1

Student Demographic Information for the 2007-2008 School Year

	Grade			Total
	3rd	4th	5th	
Gender				
Female	109	117	102	328(46.07%)
Male	129	133	122	384(53.93%)
Race				
White	194	198	192	584(82.02%)
Hispanic	22	25	18	65(9.13%)
Black	11	9	12	32(4.49%)
Multiracial	0	3	3	6(0.84%)
Asian	3	5	4	12(1.69%)
Am. Indian	5	5	3	13(1.83%)

Note. Traditional Grading System

Table 2

Student Demographic Information for the 2008-09 School Year

	Grade			Total
	3rd	4th	5th	
Gender				
Female	130	129	135	394(51.17%)
Male	130	122	124	376(48.83%)
Race				
White	207	211	209	627(81.43%)
Hispanic	25	25	30	80(10.39%)
Black	11	9	12	32(4.15%)
Multiracial	0	3	3	6(0.78%)
Asian	3	5	4	12(1.56%)
Am. Indian	5	5	3	13(1.69%)

Note. Standards-Based Grading System

The researcher collected two sets of CRCT math and reading data for four Title 1 elementary schools that transitioned from a traditional grading system to a standards-based grading system. The researcher included these sets of data because they provided the criteria for a retention or promotion decision. Therefore, the researcher did not conduct any statistical analyses that included these data sets. The first set of data were students' CRCT math and reading scores under a traditional grading system while the second set of data were the students' CRCT reading and math scores under a standards-based grading system. Math and reading data sets for the 2007-08 school year included 666 individual student scores; math and reading

datasets for the 2008-09 school year included 731 observations. The researcher merged the 2008 CRCT math and reading data (traditional grading system) and the 2009 CRCT math and reading data (standards-based grading system) by using the student’s de-identifiable number, which meant the researcher combined two data sets for each student. The researcher followed an identical procedure for the 2009 CRCT math and reading dataset.

The researcher referred to the levels of the CRCT score variable as D (Does not meet), I (In progress), M (Meets), and X (Exceeds standards). The “End Action” variable, which provided each student’s retention status, had two categories: R (retained) and NR (not retained). Tables 1 and 2 provide summaries of retention rates under a traditional grading system and a standards-based grading system. In these tables, the *N* is the total sample of observations and percent is the percent of students who make up the respective categories. The researcher determined that the target school retained about 2% of the students under the traditional grading system (see Table 3) and less than 1% of the students under the standards-based grading system (see Table 4).

Table 3

Overall Retention Rate for the 2007-2008 School Year: Traditional Grading System

Retention	<i>N</i>	Percent%	Cumulative%
NR	653	98.05	98.05
R	13	1.95	100.00

Note. NR=Not retained; R=retained

Table 4

Overall Retention Rate for the 2008-09 School Year: Standards-Based Grading System

Retention	<i>N</i>	Percent%	Cumulative%
NR	726	99.32	99.32
R	5	.68	100.00

Note. NR=Not retained; R = retained

Tables 5 and 6 provide the frequency distribution of retention based on grade and gender. Table 5 provides students' retention data under a traditional grading system while Table 6 includes students' retention data under a standards-based grading system.

Table 5

Retention Rate in 2007-08 by Grade and Gender, and Traditional-Based Grading System

Retention	Grade			Total
	3rd	4th	5th	
Females				
NR	104(97.2%)	120(100%)	101(95.3%)	98.2% (avg)
R	3(2.8%)	0	3(2.83%)	1.8% (avg)
Total	107	120	104	331
Males				
NR	99(97.1%)	122(100%)	101(98%)	97.9% (avg)
R	5(4.9%)	0	2(1.9%)	2.1% (avg)
Total	104	122	103	329

Note. NR=Not retained; R = retained

Table 6

Retention Rate in 2008-09 by Grade and Gender, and Standards-Based Grading System

Retention	Grade			Total
	3rd	4th	5th	
Females				
NR	127(100%)	121(100%)	113(99.1%)	99.7% (avg)
R	0	0	1	0.3% (avg)
Total	127	121	114	362
Males				
NR	117(98.3%)	125(100%)	123(98.4%)	98.9% (avg)
R	2(1.7%)	0	2(1.6%)	1.1% (avg)
Total	119	125	125	370

Note. NR=Not retained; R = retained

Tables 7 and 8 provide frequency distributions for the students' CRCT math and reading scores under a traditional grading system and a standards-based grading system. The researcher categorized the variable score based on a traditional grading system as A, B, C, and F, and the variable based on a standards-based grading system as D (D=Does Not Meet), I (I=In Progress), M (M=Meets), and X (X=Exceeds). The majority of the students (approximately 75%) received an A (37.09%) or a B (37.69%) in math using the traditional grading system (see Table 7).

Table 7

Math Frequency Scores: Traditional and Standards-Based Grading System

Grading System	<i>N</i>	Percent%	Cumulative%
Traditional			
A	247	37.09	39.04
B	251	37.69	76.70
C	106	15.92	92.64
F	49	7.36	100.00
Standards-Based			
D	9	1.23	1.23
I	146	19.97	21.20
M	445	60.88	82.08
X	131	17.92	100.00

Note. Missing cases in traditional math scores = 9, D=Does Not Meet, I= In Progress, M=Meets, and X=Exceeds

Table 8

Reading Frequency Scores: Traditional and Standards-Based Grading System

Grading System	<i>N</i>	Percent%	Cumulative%
Traditional			
A	337	50.60	51.95
B	232	34.83	86.79
C	76	11.41	98.20
F	12	1.80	100.00
Standards-Based			
D	11	1.50	1.50
I	123	16.83	18.33
M	465	63.61	81.94
X	132	18.06	100.00

Note. Missing cases in traditional math scores = 9, D=Does Not Meet, I= In Progress, M=Meets, and X=Exceeds

The majority of the students (approximately 85%) received an A (50.60%) or a B (34.83%) in reading using the traditional grading system (see Table 6); conversely, approximately 61% of the students meet (M) and about 18% exceed (X) the standards for math when using the standards-based grading system (see Table 5). Table 6 shows that about 64% of the students meet (M) and about 18% exceed (X) the standards for reading when using the standards-based grading system.

CRCT Score Summary

Tables 9-10 provide the summary statistics for the students' CRCT scores.

Table 9

CRCT Math Scores for 3rd, 4th, and 5th Grade Students

	<i>N</i>	<i>M(SD)</i>	Minimum	Median	Maximum
Traditional	666	834.08(39.12)	723	834	955
Standards-Based	731	835.32(40.99)	733	833	990

Notes: D=<799, M=800-849, X=>850

Table 10

CRCT Reading Scores for 3rd, 4th, and 5th Grade Students

	<i>N</i>	<i>M(SD)</i>	Minimum	Median	Maximum
Traditional	666	831.74(26.40)	746	833	920
Standards-Based	731	829.47(25.09)	749	828	920

Notes: D=<799, M=800-849, X=>850

The boxplots in Figures 1 and 2 provide a visual summary of CRCT scores based on retention status for math and reading for the traditional and standards-based grading systems. Each of the two boxplots provides the CRCT scores for retained and non-retained students. The length of the box is the distance between the 25th and the 75th percentiles, which is the interquartile range. The horizontal line inside the box represents the median CRCT score. The vertical lines emerging from the box extend to the minimum and maximum values of the CRCT scores. The small circles beyond the maximum and minimum point are outliers. The top and bottom circles indicate very high scores and very low scores when compared to others.

In these boxplots, light gray and dark gray boxplots are for both non-retained and retained students under a traditional grading system and a standards-based grading system.

Figures 1 and 2 show that non-retained students have higher median CRCT scores than those of retained students in both reading (non-retained = 833 vs. retained = 793) and math (non-retained = 834 vs. retained = 792) and under a traditional grading system and in both reading (non-retained = 828 vs. retained = 796) and math (non-retained = 834 vs. retained = 773) under a standards-based grading system.

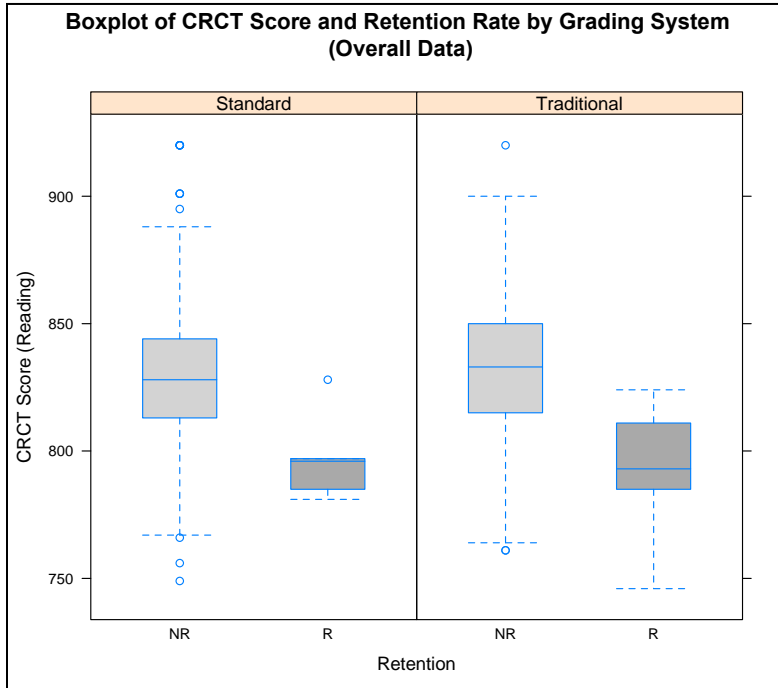


Figure 1. CRCT score on reading test by retention status by grading system

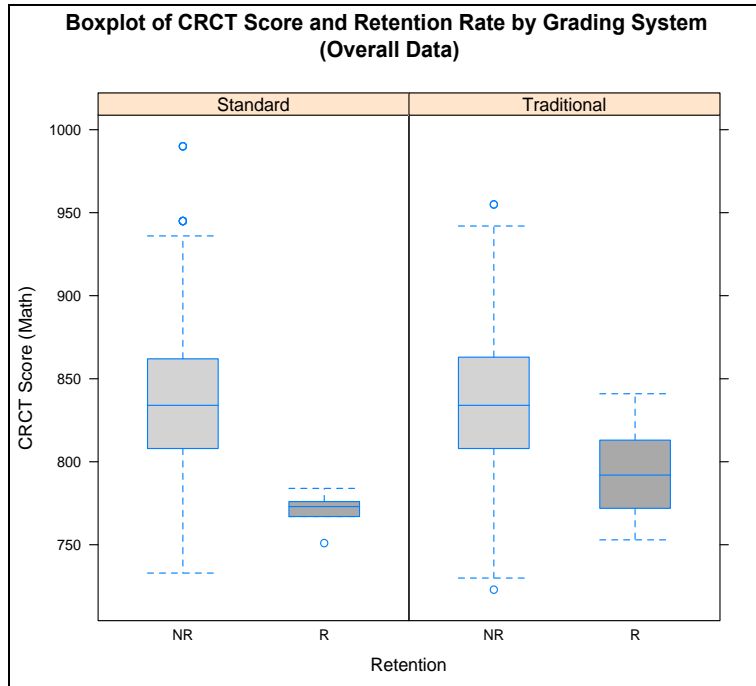


Figure 2. CRCT score on math test by retention status by grading system

Figures 3 and 4 provide students CRCT scores in reading and math based on retention status as well as for gender (male and female) across the entire data set. Overall, the researcher determined that median CRCT scores for non-retained male or female students (light gray box plots) were very close in both math and reading (non-retained females: reading = 833, math = 834; non-retained males: reading = 831, math = 835). The researcher found that the median CRCT scores of the females were higher than those of the males; however, the researcher did not determine if there was a statistically significant difference between the two groups (retained females: reading = 789, math = 786; retained males: reading = 796, math = 773).

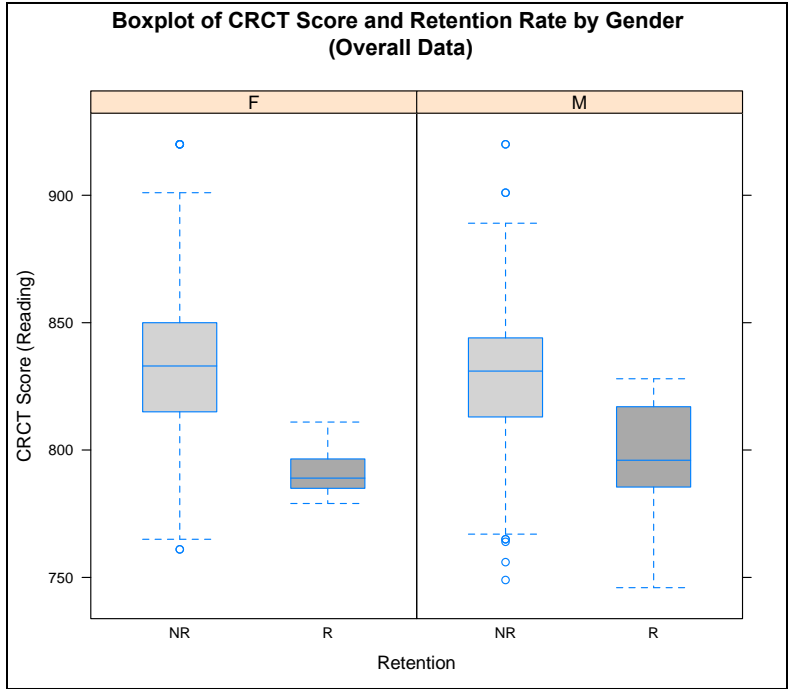


Figure 3. CRCT score on reading test by retention status and gender

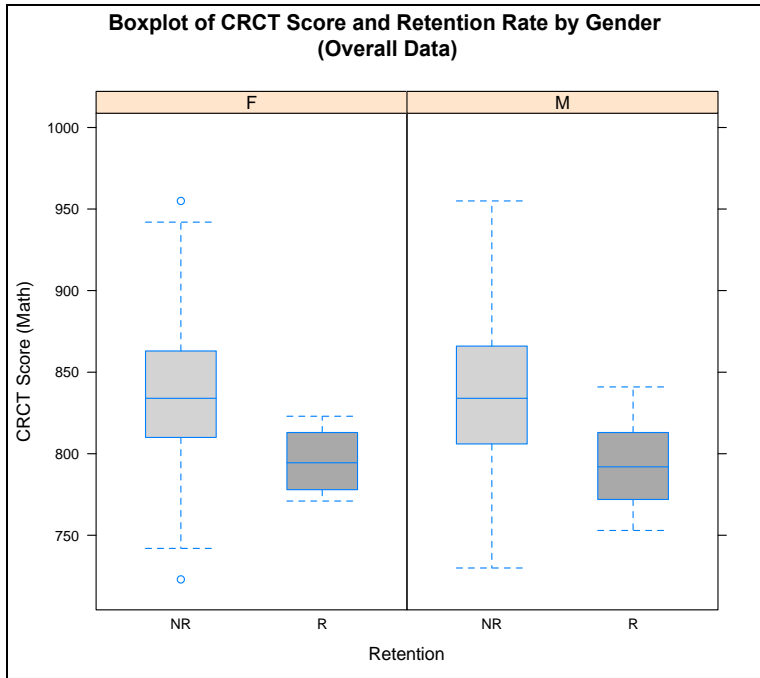


Figure 4. CRCT score on math test by retention and gender

Figures 5 and 6 provide CRCT scores in reading and math based on both retention status and grade level. The median CRCT scores in reading and math for the non-retained students were higher than the scores for the retained students in both grades (non-retained 3rd grade students: reading = 834, math = 834 vs. retained third grade students: reading = 791, math = 803) (non-retained fifth grade students : reading = 828, math = 844 vs. retained fifth grade students: reading = 796, math = 777). The singular boxplots in grade 4 are due to the fact that no students were retained in fourth grade.

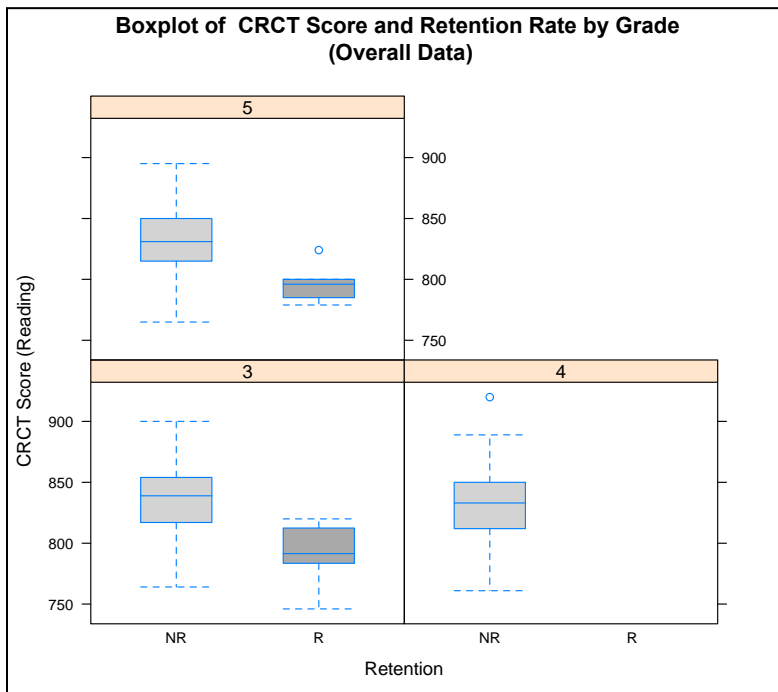


Figure 5. CRCT score on reading test by retention rate and grade

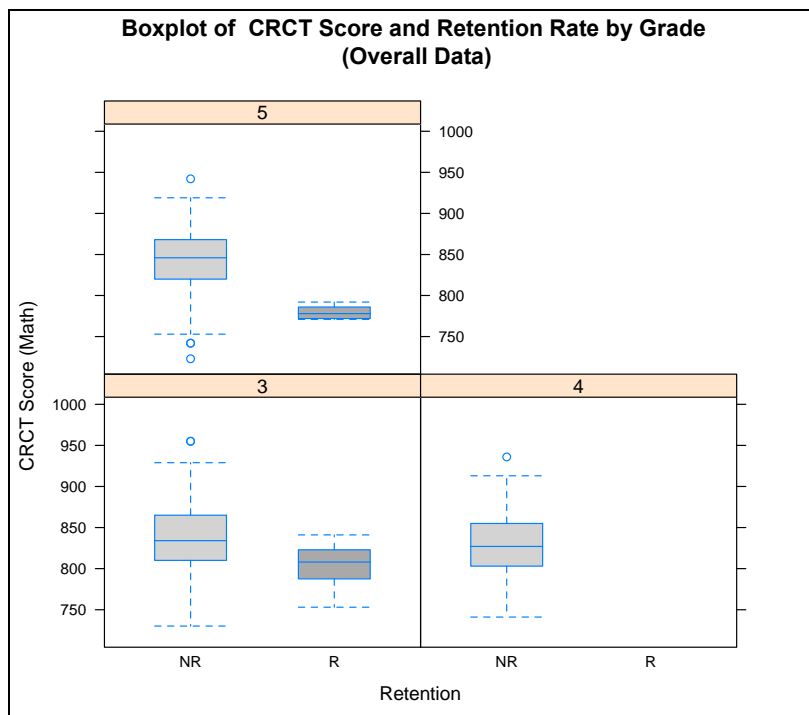


Figure 6. CRCT score on math test by retention rate and grade

Noteworthy was the school district's retention criterion for students' reading scores, which required grade retention for any third grade student whose score was below 800. Twenty students had reading scores that were below 800 under a traditional grading system, but the school district only retained five of the 20 students. The school district retained three additional students who had CRCT reading scores higher than 800 (811, 814, and 820). The school district retained these students because they were unable to read a grade-level text or had poor performance in the regular classroom. Utilizing an identical retention criterion (> 800) for reading scores under a standards-based grading system, the researcher found that 22 third grade students had reading scores below 800 under the standards-based grading system, but the school district only retained two of the 22 students. The researcher also concluded that the school district did not retain any student who had reading scores higher than 800.

The target school district's retention policy for fifth grade students was as follows: The

school district retained any student who had a score below 800 on either the CRCT reading or CRCT math tests. Thirty-eight students were unable to meet the above-stated criteria under the traditional grading system. Of these 38 students, 17 had math scores below 800, five had reading scores below 800, and 16 had both math and reading scores below 800. However, the school district only retained five of the 38 students (three with low math and reading scores, and two with low math scores only). Under the standards-based grading system, 40 students had scores below 800 on either the CRCT reading or CRCT math tests. Of these 40 students, 12 had math scores below 800, 11 had low reading scores below 800, and 17 had both reading and math scores below 800. The school district retained three of the 40 students (two had low math and reading scores, and one had a low math). Further analysis revealed that the school district did not retain some third and fifth grade students under either grading system, even though they failed to meet the district's promotion criteria. The researcher was unable to determine why the school district did not retain these students. The researcher also determined that the school district retained more students under a traditional grading system than under a standards-based grading system.

RQ1: Is there a difference between the third grade students' grade retention rates under a traditional grading system and their grade retention rates under a standards-based grading system?

H₀₁: There will be not be a significant difference between the 2007-08 third grade group's grade retention rate under a traditional grading system and the 2008-09 third grade group's grade retention rate under a standards-based grading.

RQ2: Is there a difference between the fifth grade students' grade retention rates under a traditional grading system and their grade retention rates a standards-based grading system?

H₀₂: There will be not be a significant difference between the 2007-08 fifth grade group's grade retention rate under a traditional grading system and the 2008-09 fifth grade group's grade retention rate under a standards-based grading.

RQ3: Is there a difference among the third and fifth grade female students' grade retention rates as a result of transitioning from a traditional grading system in 2007-08 to a standards-based grading system in 2008-09?

H₀₃: There will not be a significant difference among the third and fifth grade female students' grade retention rates as a result of transitioning from a traditional grading system to a standards-based grading system.

RQ4: Is there a difference among the third and fifth grade male students' grade retention rates because of transitioning from a traditional based grading system in 2007-08 to a standards-based grading system in 2008-09?

H₀₄: There will not be a significant difference among the third and fifth grade male students' grade retention rates as a result of transitioning from a traditional grading system to a standards-based grading system.

Analysis

The researcher determined if a difference existed in grade retention rates by comparing student retention rates when transitioning from a traditional grading system to a standards-based grading system. The researcher initially planned to utilize the Pearson chi-square test to determine the difference between categorical variables. The assumptions of the Pearson chi-square are that each participant contributes data to one cell and that the expected frequencies are greater than five (Field, 2009; Warner, 2013). If researchers' data do not meet these assumptions, then they should conduct a Fisher's Exact test (Field, 2009). The Fisher's Exact

test provides a precise statistical significance for small sample sizes (Field, 2009; Freeman & Julious, 2007; Warner, 2013).

The Fisher's Exact statistic also allows researchers to test the differences between the counts of categorical response variables of two independent groups (Freeman & Julious, 2007; Warner, 2013), which in this study included determining the difference and comparing the count of retained students among grades and between genders. The researcher utilized the Fisher's Exact test of independence to answer each of the research questions and the corresponding null hypotheses because more than 20% of cells had values less than five (Bolboacă et al., 2011; Howell, 2010).

To determine the statistical significance of a result, the researcher used a significance level of .05; therefore, if a two-tailed p -value was less than .05 ($p < .05$), then the researcher determined the result was statistically significant. When determining the p -value, the researcher assumed there was no relationship between the explanatory variable (grading system) and the response variable (retention rates) for the entire student population. If a p -value was less than .05, then the researcher inferred there was a less than 5% chance that results found in the sample were due to chance. By convention, this chance was small enough to make the inference that there was in fact some relationship of the explanatory variable to the response variable in the population.

RQ1. Is there a difference between the third grade students' grade retention rates under a traditional grading system and their grade retention rates under a standards-based grading system?

For Research Question 1, there was no statistically significant difference between the traditional and standards-based grading system for reducing the frequency of retention for third

graders [$\chi^2 (1, N = 467) = 4.377, p = .052$]. The two-tailed p -value from the Fisher's Exact test $p = .052$, was slightly larger than the predetermined significance level of $\alpha = .05$; therefore, the researcher failed to reject H_{01} , and the researcher concluded there was no difference between third grade students' retention rates and the traditional and standards-based grading systems. Based on post-hoc power analysis, the researcher also determined that had the target school system retained one more student under a traditional grading system, then the researcher would have had a p -value = .029. Conversely, if the target school system retained one more student under a standards-based grading system, then the researcher would have had a p -value = .015. Table 11 provides the data summary for the Pearson chi-square test and the Fisher's Exact test. Table 12 displays the percentages of third grade students retained under a traditional grading system (3.62%) and a standards-based grading system (0.81%).

Table 11

Pearson Chi-Square Test and Fisher's Exact Test

Test	<i>DF</i>	χ^2	<i>p</i>
Chi-Square	1	4.377	.036
Fisher's	1		.052

Note. Fisher's p value used to test the hypothesis.

Table 12

Overall Retention Rate for Third Grade Students Based on Grade Level

Grade 3		
Retention	Traditional	Standards-Based
NR	213(96.38%)	244(99.19%)
R	8(3.62%)	2(.81%)

Note. NR=Not retained; R = retained

RQ2. Is there a difference between the fifth grade students' grade retention rates under a traditional grading system and their grade retention rates under a standards-based grading system?

For Research Question 2, there was no statistically significant difference between the traditional and standards-based grading system for reducing the frequency of retention for 5th graders [$\chi^2 (1, N = 443) = 0.8874, p = .48$]. The two-tailed p -value from the Fisher's Exact test, $p = .48$, was greater than the predetermined significance level of .05 and resulted in the researcher failing to reject H_{02} as there was not enough evidence to conclude a difference existed between fifth grade students' retention rates and the grading systems. Table 13 provides the data summary for the Pearson chi-square test and the Fisher's Exact test. Table 14 displays the percentages of fifth grade students retained under a traditional grading system (2.45%) and a standards-based grading system (1.26%).

Table 13

Pearson Chi-Square Test and Fisher's Exact Test

Test	DF	χ^2	p
Chi-Square	1	0.8874	.346
Fisher's	1		.480

Note. Fisher's p value used to test the hypothesis.

Table 14

Overall Retention Rate for Fifth Grade Students Based on Grade Level

Grade 5		
Retention	Traditional	Standards-Based
NR	199(97.55%)	236(98.74%)
R	5(2.45%)	3(1.26%)

Note. NR=Not retained; R = retained

RQ3. Is there a difference among the third and fifth grade female students' grade retention as a result of utilizing a traditional grading system during the 2007-08 school year and transitioning to a standards-based grading system during the 2008-09 school year?

For Research Question 3, there was no statistically significant difference among third and fifth grade female students' grade retention when transitioning from a traditional grading system to a standards-based grading system [$\chi^2 (1, N = 452) = 4.3523, p = .054$]. The two-tailed p-value from the Fisher's Exact test, $p = .054$, was greater than the predetermined significance level of

.05 and resulted in the researcher failing to reject H_{03} as there was not enough evidence to conclude that a statistically significant difference existed between fifth grade students' retention rates and the grading systems. Table 15 provides the data summary for the Pearson chi-square test and the Fisher's Exact test. Displayed in Table 16 are the percentages of retained female students under a traditional grading system (2.84%) and the standards-based grading system (0.41%).

Table 15

Pearson Chi-Square Test and Fisher's Exact Test

Test	<i>DF</i>	χ^2	<i>p</i>
Chi-Square	1	4.3523	.037
Fisher's	1		.054

Note. Fisher's *p* value used to test the hypothesis.

Table 16

Overall Retention Rate by Gender and Grading System

Retention	Traditional	Standards-Based
Female		
NR	205(97.16%)	240(99.59%)
R	6(2.84%)	1(.41%)
Male		
NR	207(96.73%)	240(98.36%)
R	7(3.24%)	4(1.64%)

Note. NR = Not retained; R = retained

RQ4. Is there a difference among the third and fifth grade male students’ grade retention rates as a result of utilizing a traditional grading system during the 2007-08 school year and transitioning to a standards-based grading system during the 2008-09 school year?

Table 16 provides the percentages of male students retained under a traditional grading system (3.24%) and the standards-based grading system (1.64%). For Research Question 4, the researcher found no statistically significant difference among third and fifth grade male students’ grade retention when transitioning from a traditional grading system to a standards-based grading system [$\chi^2 (1, N = 458) = 1.2949, p < .361$]. The two-tailed p -value from the Fisher’s Exact test, $p = .361$, was greater than the predetermined significance level of .05. This finding resulted in the researcher failing to reject H_{04} as there was not enough evidence to conclude that a statistically significant difference existed among third and fifth grade male students’ grade retention rates as a result of transitioning from a traditional grading system to a standards-based

grading system. Table 17 provides the data summary for the Pearson chi-square test and the Fisher's Exact test.

Table 17

Pearson Chi-Square Test and Fisher's Exact Test

Test	DF	χ^2	p
Chi-Square	1	1.2949	.255
Fisher's	1		.361

Note. Fisher's p value used to test the hypothesis.

Summary

The researcher compared students' retention rates at the third grade level as well as at the fifth grade level under a traditional grading system and a standards-based grading system. The school district utilized a traditional grading system during the 2007-08 school year and standards-based grading system during the 2008-09 school year, which allowed the researcher to conduct same grade-level comparisons under two distinct grade systems. While the percentage of students retained under a traditional grading system was higher at both grade levels (third grade traditional grading system = 3.68% versus standards-based grading system = 0.81%; fifth grade traditional grading system = 2.45% versus standards-based grading system = 1.26%), the researcher concluded there was not a significant difference under the two grading systems at either the third grade level ($\chi^2 = 4.377, p = .052$) or the fifth grade level ($\chi^2 = 0.8874, p = .48$).

In subsequent comparisons, the researcher found that retention rate was higher for both females (2.84% = traditional grading system and 0.41% = standard based grading system) and

males (3.24% = traditional grading system and 1.64% = standards-based grading system) under a traditional grading system. However, the researcher found there was not a significant difference under the two grading systems for either the female students ($\chi^2 = 4.3523, p = .054$) or the male students ($\chi^2 = 1.2949, p = .36$).

Pertinent in both of these overall comparisons was the exclusion of the fourth grade because the target school district did not retain fourth grade students during the 2007-08 and 2008-09 school years. To determine whether the differences found in the data set was the result of something other than chance for the 2007-08 and 2008-09 school years, the researcher conducted statistical analyses across grade levels and based on gender.

The researcher found that the third and fifth grade students' retention rates had no significant differences between the grading systems and the retention rates. Retention rates for third grade students were 0.81% and 3.62% when using the traditional grading system and a standards-based grading system; conversely, retention rates for fifth grade students were 1.26% under a traditional grading system and 2.45% under a standards-based grading system.

The researcher also compared male and female grade retention rates under both a traditional grading system (3.27% = males and 2.84% = females) and a standards-based grading system (1.64% = males and 0.41% = females). While the male and female students' overall grade retention rates were lower under a standards-based grading system, the researcher concluded there was not a difference between each gender's grade retention rate and the two grading systems. With this in mind, the researcher concluded that the educational practitioners' decisions to retain or to promote students functioned independently from both the traditional grading system and the standards-based grading system.

CHAPTER FIVE: DISCUSSION

Introduction

Since the NCLB Act, states across America have been using a standards-based grading system to ensure that students master grade-level content. To ensure mastery of standards, many schools have eliminated traditional report cards and have adopted standards-based reporting. The increased utilization of both standards-based instruction and assessments within America's public schools will hopefully decrease student retention and improve student achievement and success (Forman & Sanders, 2001). Researchers found that grade retention had both a positive and a negative impact on students (Bowman, 2005; Jimerson et al., 1997). Holmes (1989) found that retained students developed poor self-esteem and a negative attitude towards school while Jimerson et al. (1997) asserted that retained students struggle emotionally and socially in the school setting. Jimerson (1999), Roderick (1994), and Shepard and Smith (1990) concluded that grade retention negatively impacted a student's chance of graduating high school. Conversely, Gleason et al. (2007) determined that grade retention had a positive impact as retained students had a higher peer acceptance rate than their promoted peers.

Restatement of the Problem

To ensure mastery of standards, many schools have eliminated traditional report cards and have utilized standards-based reporting. Grade retention has a negative impact on student's academic achievement (Brown, 2007). Researchers found that grade retention was an ineffective intervention strategy (Jimerson, 2001; Xia & Glennie, 2005). While some students demonstrated achievement gains in the year following a grade retention decision, they relinquished these gains within three years (Hauser, 1999; Holmes, 1989; Karweit, 1991; NASP, 1998; Thompson & Cunningham, 2000). With this in mind, this study examined whether a traditional-based grading

system or a standards-based grading system improved grade retention rates.

Research Questions

RQ1. Is there a difference between the third grade students' grade retention rates with a traditional grading system and their grade retention rates with a standards-based grading system?

The two-tailed p -value from the Fisher's Exact test was 0.0522, which was slightly larger than the pre-specified significance level $\alpha=0.05$; therefore, the researcher failed to reject H_{01} , and the researcher concluded there was no difference between third grade students' retention rates and the traditional and standards-based grading systems. However, the researcher determined that third grade students' retention rates were higher under a traditional grading system, but the differences between the numbers of students retained under the two grading systems was not statistically significant. While the differences between the groups was not statistically significant, the two-tailed p -value (0.052) was only slightly greater than the pre-determined level of significance ($p < 0.050$) that the researcher needed to reject H_{01} , suggesting that including additional students in the sample could reveal that a significant difference existed between the groups.

RQ2. Is there a difference between the fifth grade students' grade retention rates with a traditional grading system and their grade retention rates with a standards-based grading system?

The two-tailed p -value from the Fisher's Exact test was 0.48, which was greater than the pre-specified significance level of 0.05. The researcher failed to reject H_{03} as there was not enough evidence to conclude a difference existed between fifth grade students' retention rates and the grading systems. The researcher found that the percentage of fifth grade students who were retained was higher under a traditional grading system. The researcher concluded that the school district retained 1.26% of the students using a standards-based grading system and 2.45%

of students using a traditional grading system.

RQ3. Is there a difference among the third and fifth grade female students' grade retention rates as a result of transitioning from a traditional based grading system in 2007-08 to a standards-based grading system in 2008-09?

The two-tailed p -value from the Fisher's Exact test was 0.054, which was greater than the pre-specified significance level of 0.05; hence, the researcher failed to reject H_{04} as there was not a statistically significant difference in the female students' grade retention rates when transitioning from a traditional grading system to a standards-based grading system. However, the researcher found that the school district retained 0.41% of the female students using a standards-based grading system and 2.84% of the female students using a traditional grading system. Although the difference between the groups was not statistically significant, the two-tailed p -value (0.054) was only slightly greater than the pre-determined level of significance ($p < 0.050$) that the researcher needed to reject H_{03} , suggesting that including additional female students in the sample could reveal that a significant difference existed between the groups.

RQ4. Is there a difference among the third and fifth grade male students' grade retention rates as a result of transitioning from a traditional grading system in 2007-08 to a standards-based grading system in 2008-09?

The two-sided p -value from the Fisher's Exact test was 0.36, which was greater than the pre-specified significance level of 0.05; therefore, the researcher failed to reject null hypothesis H_{04} as there was not a statistically significant difference in the male students' grade retention rates when transitioning from a traditional grading system to a standards-based grading system. However, the researcher determined that third grade students' retention rates were higher under a traditional grading system (3.24%) than under a standards-based grading system (1.64%), but the

difference between the numbers of students retained under the two grading systems was not statistically significant.

Discussion

For the rural school system in Northeast Georgia, as well as school systems across the state of Georgia and throughout the nation, grade retention impacted the student and the school system on many different levels. Schools have transitioned from a traditional grading system to a standards-based grading system to promote students' mastery of grade-level standards system and to comply with the guidelines of the Common Core State Standards (Craig, 2012; Guskey, 2001). In order to achieve greater promotion rates, school districts must identify students' learning gaps and support these students. This study occurred in a school district that was proactively engaged in student learning, promotion, and achievement. The school district's overall student retention rates were well below the national average.

The researcher found that student retention rates were higher under a traditional grading system than under a standards-based grading system for third and fifth grade students and based on gender (males and females). However, the researcher did not find a statistically significant difference in grade retention rates when transitioning from a traditional grading system (2007-08) to a standards-based grading system (2008-09) for third and fifth grade students and for males and females. The findings of West (2012), who compared grade retention rates before and after implementing standards-based grade retention criteria, may provide insight into why the researcher did not find a statistically significant difference. West found that in the school year following the implementation of the standards-based testing criteria, grade retention rates increased; however, in each of the next six school years (2002-2008), student retention rates decreased from 13.5% in 2002 to 5.6% in 2008. This researcher asserts that comparing student

retention rates under a traditional grading system and student retention rates several years after implementing a standards-based grading system could reveal significant differences.

The researcher also concludes that the findings of West (2012) support the belief that a standards-based grading system creates the requisite learning environment for increasing students' achievement. The standards-based grading system requires teachers to determine where students are on the learning continuum, implement appropriate teaching strategies, and promote students toward a deeper level of understanding (Blake & Pope, 2008; Haptonstall, 2010; Scriffiny, 2008). By determining where students are on the learning continuum, and by implementing teaching strategies that consider students' distinct learning styles and needs, teachers create impactful and engaging learning experiences for students (Haptonstall, 2010).

Determining each student's level of knowledge and his or her distinct learning needs are central concepts of Vygotsky's zone of proximal development (Blake & Pope, 2008; Hatch, 2010). Craig (2010) stated, "The increased opportunities to learn within the zone of proximal development are evidenced by expanded growth and improved performance" (p. 127). The zone of proximal development is the foundation for Vygotsky's social constructivism and is the essence of a standards-based assessment system (Hardedge, 2012). Teaching to students' zone of proximal development facilitates student learning, and students' cognitive development increases when student learning occurs (Guskey, 2007; Hatch, 2010).

The researcher did not determine if a statistically significant difference existed between the retention rates of male and female students; however, the researcher found that males had higher grade retention rates than those of females under both the traditional grading system (2007-08: males = 7 [3.24%]; females = 6 [2.84%]) and the standards-based grading system (2008-09: males = 4 [1.64%]; females = 1 [.41%]). This finding is consistent with other

researchers who found that grade retention rates were typically higher for boys than for girls (Grant, 1997; Nagaoka & Roderick, 2004; Hauser et al., 2000; Roderick, 1994).

As states across the nation implement rigorous academic standards, school systems must evaluate their grading systems. School systems may look to other school systems that have already undergone the transition from traditional to a standards-based grading system. The goal of any grading system should be to promote student learning and achievement as well as to decrease grade retention rates. This study determined whether a traditional grading system or a standards-based grading system was most effective on grade retention rates and whether there was a significant difference in gender retention rates.

Limitations of the Study

This study compared the grade retention rates at each grade level with a traditional grading system (2007-08) and a standards-based grading system (2008-09). Limitations that inhibited the extent to which the researcher generalized findings were the threat of maturation, and this study's sampling procedure. The threat of maturation was a concern as student retention may have occurred because of chronological age or a developmental disability rather than a traditional or a standards-based grading system. With this in mind, a potential threat to validity was the researcher's inability to determine why the target school district retained a student.

Additional limitations include this study's sample size, the similar demographics of the schools (all Title 1 schools), the study's sampling procedures, and the researcher's inability to determine which students received special education services and to discern if teachers' beliefs influenced grade retention decisions. In regard to the sample size of each grade level, the researcher would need to recruit 918 students to have an 80% chance (power = 0.8) that a random sample from this population would result in a statistically significant *p*-value. For the

test to be significant at $p < .05$, the researcher would need to recruit 520 participants, which would be 53 more participants than the researcher included in this study ($N = 467$). This finding presupposes that any increase in population would not change the proportions for the groups. Assuming that the researcher determined the true proportions in the population, the probability of drawing a random sample from this study's population resulting in a statistically significant p -value (power) would be 0.445 (44.5%).

Next, the researcher only included Title 1 schools in this study, precluding the researcher's ability to conduct between school differences. When comparing grade retention rates between Title 1 and non-Title 1 schools, the Collier County Public Schools (2012) found that grade retention rates were higher at Title 1 schools for grades K-2, but were also similar to retention rates at non-Title 1 schools for grades 4-10. McGill-Franzen, Zmach, Solic, and Zeig (2006) provided a possible reason for the discrepancy in grade retention rates between Title 1 and non-Title 1 schools for grades K-2, concluding that students develop their reading skills during these grades. Douglas and Montiel (2008) determined that the baseline reading scores of kindergarteners from low-income socioeconomic backgrounds were at least three months behind their peers who were from high-socioeconomic backgrounds. This achievement discrepancy continued in first grade as students from wealthy socioeconomic backgrounds had higher basic reading skills than those of students from poor socioeconomic backgrounds.

The utilization of a convenience sampling procedure was another limitation because this type of sampling procedure introduced sampling bias. The researcher could have reduced this bias by utilizing random sampling procedures (Creswell, 2008). While the researcher included special education students in this study, the researcher was unable to determine how many students received special education services or had a 504 plan during the 2007-08 and 2008-09

school years. The researcher was able to determine that if students had an IEP for a disability, then the target school district promoted them, even if they failed the CRCT. Students with 504 plans may have received testing accommodations, including extended time, small group setting, prompting upon request, but the school district retained these students if they did not pass the CRCT. Finally, the researcher did not determine if or how teachers influenced grade retention decisions. Researchers found that teachers' pedagogical strategies and beliefs influenced grade retention decisions (Smith & Sheppard, 1988; Stipek & Byler, 1997; Tomchin & Impara, 1992).

These pedagogical strategies included teachers individualizing instructional practices and assessments, adapting the frequency of assessments based on students' ability levels, grouping students based on their learning preferences and interests, and providing students with peer tutors and mentors (Smith & Sheppard, 1988; Stipek & Byler, 1997). Tomchin and Impara (1992) concluded that teachers' beliefs concerning grade retention was greatly influenced by the grade they taught; for example, primary grade level teachers (K-3) believed that student mastery of basic skills was a prerequisite for academic success in later years whereas upper grade teachers (grades 4-7) considered factors "such as motivation, behavior, absenteeism, and work habits" (p. 218) when making grade retention decisions.

Implications for Future Research

The current study is significant because it provides information about grade retention rates under two different grading systems. While the researcher determined there was not a statistically significant difference in third and fifth grade students' retention rates under a traditional grading system and standards-based grading system, the researcher found the school district retained more students under a traditional grading system at both grade levels. The

higher retention rate under the traditional grading system, while seemingly negligible, could impact the target school district's ability to meet individual AYP benchmarks.

Title 1 schools, which is how each of the schools included in this study are classified, that fail to meet AYP benchmarks must inform parents that they failed to meet AYP benchmarks, offer school choice to parents and children, provide supplementary instruction to students, and reassign teachers and administrators. Carifio and Carey (2010) identified several shortcomings of traditional grading systems, including their proclivity to limit academic success to the brightest students, inability to cultivate students' higher-order thinking skills, and failure to consider students' learning styles; moreover, they concluded, "traditional grading does not so much ignore motivation theory as much as it simply predates it" (p. 228).

In contrast, standards-based grading systems provide a comprehensive understanding of students' progress, facilitate students' success by giving them additional opportunities to demonstrate mastery, encourage teachers to move fundamental to abstract concepts, and enable "the school to provide more differentiated instruction, in large part because they track precisely what students know and don't know" (Spencer, 2012, p. 9). When considering student achievement with traditional and standards-based grading systems, Craig (2012) cited Bandura's theory of self-efficacy when discussing how the type of grading system influenced both students' effort and belief in their ability to complete a task. With traditional grading practices, students' "sense of self-efficacy is impacted by the recognition that others readily receive good grades for similar effort, or less, than they expend to receive failing grades" (Craig, 2012, p. 24). This type of learning experience may result in struggling students expending less effort, exhibiting less motivation, and perpetuating a pattern of repeated failure (Craig, 2012). Under standards-based grading systems, students have the potential to increase self-efficacy by providing consistent and

specific feedback regarding their potential thereby increasing their effort and motivation to complete a task (Craig, 2012).

Urich (2012) also supported the use of standards-based grading, asserting that it provides an authentic representation of student knowledge. Teachers who utilize differentiated instructional strategies recognize that: (a) learning preferences and life experiences influence students' readiness to learn, (b) different learning modalities engage students, (c) students need varying levels of support during instructional activities, (d) identifying where students are on the learning continuum and moving them forward is the essence of effective teaching, and (e) connecting academic content and real-life experiences makes learning meaningful (Landrum & McDuffie, 2010). Sieling (2013) concluded that a standards-based grading system required a paradigm shift concerning student assessment.

While researchers investigated the relationship between students' standardized tests scores and standards-based report cards, compared teachers' preferences between traditional grading and standards-based grading systems, and compared students' standard test scores under traditional and standards-based grading systems (Craig, 2012; Haptonstall, 2010; Paepflow), researchers have not compared students' retention rates under traditional and standards-based grading systems. However, researchers found that a standards-based grading system promoted an in-depth knowledge of academic content among students, provided explicit and detailed feedback regarding students' performance, improved teachers' instructional practices, and allowed teachers to provide students with individualized support and instruction (Craig, 2012; Sieling, 2013; Urich, 2012). Teachers who integrate the instructional, pedagogical, and assessment strategies associated with standards-based learning provide a foundation for improving student achievement and decreasing student retention (West, 2012).

The researcher included schools that were racially and socioeconomically homogeneous as all four elementary schools were Title 1 schools and the majority of students who attended each of these schools were Caucasian (over 80%). The researcher recommends that future studies investigate grade retention under traditional and standards-based grading systems in school districts that have a more racially diverse student population as doing so could determine if grade retention is more prevalent in students who represent a particular racial group (Jimerson et al., 2006). Future studies should include a larger sample population by including students from other grades (K-2, 6-8). The researcher also discovered that the target schools did not retain a single fourth grade student under either grading system. This might not have been a random event because the state of Georgia did not establish promotion criteria for fourth grade students. However, the researcher was unable to determine if the target schools retained any fourth grade students in the school years prior to 2007-08 and after 2008-09. Researchers found that teachers' beliefs influenced grade retention decisions (Smith & Sheppard, 1988; Stipek & Byler, 1997; Tomchin & Impara, 1992). The researcher recommends that researchers interview fourth grade teachers from the target school district to better understand the degree to which teachers influence grade retention decisions.

These studies would also investigate a system's grade retention practices and policies as researchers found that school systems' retention policies are inconsistent (Evans, 2012; Roderick & Nagaoka 2005). A very interesting implication of this study is that the target school system did not adhere to the state of Georgia's promotion criteria for both third grade students on the reading portion of the CRCT and fifth grade students on the reading and math portions of the CRCT. By not adhering to the state of Georgia's promotion criteria, school systems convey to stakeholders (parents, teachers, and community members) that state policies lack merit, which

cultivates distrust among stakeholder groups in both state and school system policies.

Investigating grade retention policies allows school systems to ensure that their respective grade retention policies align with the state-mandated policies.

Common Core State Standards (CCSS) have forced schools and teachers to align curriculum and assessment, to ensure that their grade retention policies align with those of the state, and to provide academic support to struggling students (Craig, 2012; GaDOE, 2011b; Paepflow, 2011). In the state of Georgia, public schools must follow the retention and promotion criteria set forth by the Georgia Department of Education; schools must also provide research-based interventions and support for at-risk students as well as monitor their academic progress while receiving interventions and support (GaDOE, 2011b). According to the GaDOE (2011b), “an at-risk student is a student with specific needs that may hinder academic achievement” (p. 8). To monitor students’ academic progress, public schools across the state of Georgia have “implemented a four-tier Response to Intervention (RTI) model in a) identifying and b) addressing students’ academic and/or behavioral needs” (p. 3).

At the four Title 1 elementary schools, Tier 1 of RTI included all regular education students who received grade-level instruction. At Tier 1, teachers differentiated their instructional strategies for and monitored the academic progress of all students. If a student was unable to meet predetermined academic benchmarks, the RTI team placed the student at Tier 2 and referred him or her for hearing and vision screening (GaDOE, 2011b). At Tier 2, the student received research-based interventions such as small group (3-5 students) and one-on-one instruction for a 12-week period. During this 12-week period, the team met every four weeks in order to determine the student’s progress and if the student needed different interventions.

If the student was unable to meet grade-level benchmarks within a 12-week period, then the RTI team moved the student to Tier 3 and the student received more intensive research based interventions. These small group and individual interventions were more intensive and frequent than those provided at Tier 2. The RTI team also referred the student for either special education or psychological testing. If a student did not qualify for special education services, then the RTI team considered grade retention. Grade retention was a viable option if the team determined that the child needed an extra year of maturation and exposure to the grade-level curriculum. A student who was on Tier 4 had an IEP and received special education services. In addition to at-risk students, RTI addressed the educational needs of students dismissed from special education, providing them with support as they transitioned from special education to regular education. The RTI team closely monitored these students' academic progress and ensured that they received individualized interventions.

This study found that neither a traditional grading system nor a standards-based grading system decreased grade retention rates. The Title 1 elementary schools included in this study provided remedial support for third and fifth grade students who failed either the math or the reading portion of the CRCT. The Early Intervention Program (EIP) was one of the remedial support services provided by the target school district. EIP provided students with subject-specific remedial support aimed at preparing them to retake the CRCT. The researcher learned that the target school district began the EIP process in kindergarten as doing so allowed the target school district to identify students who needed additional assistance.

Based on standardized assessments, student portfolios, or EIP checklists, the target school district found that some students in grades K-2 had academic deficits but did not qualify for special education services because the discrepancy between their achievement and intellectual

ability was not significant. However, the discrepancy between their achievement and intellectual ability may have increased as the academic content became more difficult in grades 3-5, resulting in them qualifying for special education services (Lynch, 2009). All of the Title 1 elementary schools also offered an after-school tutoring program. The after-school tutoring program gave students specific and individualized academic assistance. For example, a student identified with a math disability received supplementary math instruction.

Conclusions

The purpose of this quantitative non-experimental, causal comparative, ex post facto research study was to determine whether a traditional or a standards-based grading system improved third and fifth grade retention rates of Title 1 elementary schools that were part of one North Georgia school system. Specifically, this study examined whether there was a difference in retention rates of third and fifth grade students and a traditional and standards-based grading system. This study compared CRCT math and reading scores for third grade and fifth grade students because of state mandates in regard to retention policies in these two particular grades. The state of Georgia uses the students' CRCT scores as one of its benchmarks for assessing a school's effectiveness and for grade retention practices.

The participants in this study were third and fifth graders who attended four rural elementary schools in northeast Georgia during the 2007-08 and 2008-09 school years. This dataset contained two sections, one for the traditional grading system and one for the standards-based grading system. The analysis did not identify a significant difference in grade retention rates between the two grading systems. The study examined whether there was a significant difference in grade retention rates between a traditional grading system and a standards-based grading system by using CRCT scores in areas of math and reading and determining if there was

a difference between gender and CRCT scores. Third and fifth grade students were important to this study because the state of Georgia uses students' CRCT scores as the criteria for promoting or retaining students.

The study revealed that grade retention rates with a traditional grading system were not significantly different from the grade retention rates with a standards-based grading system. Under a traditional grading system in third grade, 20 of the students did not make the state of Georgia's promotion criteria by scoring an 800 on the CRCT, but the target school district only retained eight of those students. Of the eight students retained by the target school district, only two scored below the required 800 while three of the students scored above the required score (811, 814, and 820). Under a standards-based grading system, 22 of the students did not make the state of Georgia's promotion criteria by scoring an 800 on the CRCT; however, the target school system only retained two of those students. None of the students retained scored below an 800 on the reading CRCT.

Under a traditional grading system in fifth grade, 38 students did not make the state of Georgia's promotion criteria by scoring an 800 on the CRCT in both reading and math, but the target school system only retained five of those students. Of the five students retained by the target school district, three of the students did not make the required 800 score in reading and math and two students scored below the required 800 score in reading only. Under a standards-based grading system, 40 of the students did not make the state of Georgia's promotion criteria by scoring an 800 on the CRCT in reading and math, but the target school district only retained three of these students. Of the three students retained, two did not meet the required 800 reading and math and one did not achieve the 800 math score. The study revealed that there was not a significant difference between grade retention rates based on gender. The researcher found that

the school district retained six females and seven males under a traditional grading system and one female and four males under a standards-based grading system.

This study indicated that the grade retention criteria did not align with what occurred with third and fifth grade students who attended this northeast Georgia school system. This indicates that the grade level expectations for third and fifth grade students were not as strict as the guidelines set forth by the state of Georgia. Noteworthy was the fact that the target school district did retain any fourth grade students either year. All of this was noticeable under each of the included years and during a traditional grading system and a standards-based grading system.

In summary, the expected outcome of a decreased grade retention rate using a standards-based grading system did not occur. While there was a slight decrease in the number of students retained under a standards-based grading system, the significance was not enough to reject the null hypothesis. While it may be impossible to have a grade retention policy that applies and affects all students the same in any grading system, it is clear that not all students lacking the required CRCT are actually retained. This can be misleading to students and their families. By focusing on student learning and mastery of required standards, educators continue to improve the learning environment and the success of each student. This in turns helps each child to have a successful learning career.

REFERENCES

- Adams, C. L. (2011). *The effects of a remedial on standardized test scores in Georgia middle schools* (Unpublished doctoral dissertation). Liberty University, Lynchburg, VA.
- Alexander, K., Entwisle, D., & Dauber, S. (2003). *On the success of failure: A reassessment of the effects of retention in the primary grades* (2nd ed.). Cambridge, MA: Cambridge University Press.
- Allington, R., McGill-Franzen, A. M., & Schick, R. (1997). How administrators understand learning difficulties. *Remedial and Special Education, 18*(4), 223-232.
doi:10.1177/074193259701800404
- Ames, L. B. (1966). *Is your child in the wrong grade?* New York: Harper Row.
- Ames, L. B. (1980). Retention: A step forward. *Early Years, 11*, 10-11.
- Anderson, G., Jimerson, S., & Whipple, A. (2005). Student ratings of successful experiences at home and school: Loss of parent and grade retention as superlative stressors. *Journal of Applied School Psychology, 21*(1), 1-20. doi:10.1300/J008v21n01_01
- Ary, D., Jacobs, L. C., Razavieh, A., & Sorensen, C. (2006). *Introduction to research in education* (7th ed.). Belmont, CA: Thomson Wadsworth.
- Blake, B., & Pope, T. (2008). Developmental psychology: Incorporating Piaget's and Vygotsky's theories in classrooms. *Journal of Cross-Disciplinary Perspectives in Education, 1*(1), 59-67.
- Bloom, B. S., Madaus, G. F., & Hastings, J. T. (1981). *Evaluation to improve learning*. New York, NY: McGraw-Hill.
- Bolboacă, S. D., Jäntschi, L., Sestraş, A. F., Sestras, R. E., & Pamfil, D. C. (2011). 'Pearson-Fisher chi-square statistic revisited information. *Information, 2*(4), 528-545.

doi:10.3390/info2030528

- Bonvin, P., Bless, G., & Schuepbach, M. (2008). Grade retention decision-making and effects on learning as well as social and emotional development. *School Effectiveness and School Improvement, 19*(1), 1-19. doi:10.1080/09243450701856499
- Bowman, L. (2005). Grade retention: Is it a help or hindrance to student academic success. *Preventing School Failure, 49*(3), 42-46. doi:10.3200/PSFL.49.3.42-46
- Brennan, R. T., Kim, J., Wenz-Gross, M., & Siperstein, G. M. (2001). The relative equitability of high-stakes testing versus teacher-assigned grades: An analysis of the Massachusetts Comprehensive Assessment System (MCAS). *Harvard Educational Review, 71*(2), 173-216.
- Brookhart, S. M. (2011). Starting the conversation about grading. *Educational Leadership, 69*(3), 10-14.
- Brown, C. (2007). Examining the streams of retention policy to understand the politics of high-stakes reform. *Education Policy Analysis Archives, 15*(9), 1-28. Retrieved from <http://epaa.asu.edu/epaa/v15n9/>
- Byrnes, D., & Yamamoto, K. (1985). Academic retention of elementary pupils: An inside look. *Education, 106*(2), 208-214.
- Cannon, J. S., & Lipscomb, S. (2011). *Early grade retention and student success: Evidence from Los Angeles*. Public Policy Institute of California: San Francisco, CA. Retrieved from http://www.ppic.org/content/pubs/report/R_311JCR.pdf
- Carifio, J., & Carey, T. (2010). Do minimum grading practices lower academic standards and produce social promotion. *Educational Horizons, 88*(4), 219-230.
- Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-

- state analysis. *Educational Evaluation and Policy Analysis* 24(4), 305–331.
- Center for Policy Research in Education (1990). *Repeating grades in school: Current practice and research evidence*. CPRE Policy Briefs.
- Cherniss, A. (2008). *Standard based report card teachers perception on the development, transition, and implementation* (Doctoral Dissertation). Retrieved from ProQuest Dissertations and Theses. (UMI No. 3311106)
- Colby, S. (1999). Grading in a standards-based system. *Educational Leadership*, 56(6), 52-55. Retrieved from http://www.ascd.org/publications/educational_leadership/mar99/vol56/num06/Grading_in_a_Standards-Based_System.aspx
- Collier County Public Schools. (2012). *LEA Title I program evaluation and needs assessment*. Retrieved from <https://www.collierschools.com/fsgrants/docs/CCPS%20Title%20I%20Evaluation%20and%20Needs%20Assessment%20FY12.PDF>
- Craig, T. (2012). *Effects of standards-based report cards on student learning* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (UMI No. 3498282)
- Crawford, S. L. (2006). Correlation and regression. *Journal of the American Heart Association*, 114, 2083-2088. doi:10.1161/CIRCULATIONAHA.105.586495
- Creswell, J. W. (2008). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.
- Darling-Hammond, L. (1998). Alternatives to grade retention. *The School Administrator*, 55(7), 18-21. Retrieved from http://www.aasa.org/publications/sa/1998_08/Darling-Hammond.htm

- David, J. L. (2008). Grade retention. *Educational Leadership*, 65(6), 83-84.
- Davenport, S., Delgado, A., Meisels, M., & Moore, D.R. (1998, November). *Rethinking retention to help all students succeed: A resource guide*. Paper presented at the Rethinking Retention to Help All Students Succeed Conference, Chicago, IL. Retrieved from <http://www.designsforchange.org/pdfs/rethink.pdf>.
- Davison, M. L., Davenport, E. C., Kwak, N., Peterson, K., Irish, M, Seo, Y., et al (2002). *The "No Child Left Behind" Act and Minnesota's standards, assessments, and accountability: 2002 policy brief*. Minneapolis: University of Minnesota, College of Education and Human Development, Office of Educational Accountability.
- Dawson, P. (1998). A primer on student grade retention: What the research says. *Communique*, 26(8), 28-30. Retrieved from <http://search.ebscohost.com/proxy.liberty.edu>
- Denton, D. (2001). *Finding alternatives to failure: Can states end social promotion and reduce retention rates?* Southern Region Education Board: Atlanta, GA. Retrieved from <http://www.sreb.org/programs/srr/pubs/Alternatives/AlternativesToFailure.asp>
- Douglas, K., & Montiel, E. (2008). Learning to read in American elementary school classrooms: Poverty and the acquisition of reading skills. *International Reading Association*. Retrieved from www.reading.org/Libraries/SRII/ECLS-K_SES_Report.sflb.ashx
- Early, D., Bushnell, M., Clifford, D. Konanc, E., Maxwell, K., Palsha, S., & Roberts, L. (2003). Kindergarten Readiness Issues Group, Partners in Research Forum. *North Carolina early grade retention in the age of accountability*, Chapel Hill, N.C.: The University of North Carolina, FPG Child Development Institute.
- Ellis, T. J., & Levy, Y. (2009). Towards a novice guide for novice researchers on research methodology: Review and proposed methods. *Issues in Informing Science and*

- Information Technology*, 6, 323-337. Retrieved from <http://iisit.org/Vol6/IISITv6p323-337Ellis663.pdf>
- Evans, M. T. (2012). *How knowledge, agency, and beliefs influence grade level k-8 student retention* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (UMI No. 3487821)
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Thousand Oaks, CA: Sage.
- Forman, J., & Sanders, M. E. (2001). Critical issue: Beyond social promotion and retention—five strategies to help students succeed. Retrieved from <http://www.ncrel.org/sdrs/areas/issues/students/atrisk/at800.htm>
- Forness, S. R., Kavale, K.A., Blum, I. M., & Lloyd, J. W. (1997). Mega-analysis of meta-analyses: What works in special education and related services. *TEACHING Exceptional Children*, 29(6), 4-9.
- Fraenkel, J. R., & Wallen, N. E. (2006). *How to design and evaluate research in education* (6th ed.). Boston, MA: McGraw-Hill.
- Freeman J. V., & Julious, S. A. (2007). The analysis of categorical data. *Scope Online*, 16(1), 1821.
- Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction*. Boston, MA: Pearson Education.
- Gay, L. R., Mills, G. E., & Airasian, P. (2006) *Educational research: competencies for analysis and application* (8th ed.). Upper Saddle River, NJ: Prentice Hall.
- Georgia Department of Education. (2008). *Validity and reliability for the 2008 criterion referenced CRCT*. Assessment Research and Development Division of the Georgia Department of Education.

- Georgia Department of Education. (2009). *Technical brief: Accommodation usage on the 2009 competency-referenced competency tests and Georgia high school graduation tests*. Assessment Research and Development Division of the Georgia Department of Education. Retrieved from <http://archives.gadoe.org/DMGetDocument.aspx/2009%20Accommodations%20Technical%20Brief%20Final.pdf?p=6CC6799F8C1371F69F24BC99BEC56A98CE99F28C0DF8B764CB5F8A462EE6F759&Type=D>
- Georgia Department of Education. (2011a). *An accountability and assessment brief: 2011 CRCT scale scores and cut scores*. Assessment Research and Development Division of the Georgia Department of Education. Retrieved from http://www.gadoe.org/ci_testing.aspx?PageReq=CI_TESTING_CRCT
- Georgia Department of Education. (2011b). Response to intervention: Georgia's student achievement pyramid of interventions. Retrieved from http://public.doe.k12.ga.us/ci_services.aspx?PageReq=CIServRTI
- Gesell Institute of Human Development. (1982). *A gift of time*. New Haven, CT: Gesell Institute of Human Development.
- Gleason, K. A., Kwok, O., & Hughes, J. N. (2007). The short-term effect grade retention on peer relations and academic performance of at-risk first graders. *Elementary School Journal*, 107(4), 327-340. doi:10.1086/516667
- Grant, J. (1997). *Retention and its prevention: Making informed decisions about individual children*. Chicago, IL: Modern Learning Press.
- Greene, J., & Winters, M. (2006). Getting farther ahead by staying behind: A second-year evaluation of Florida's policy to end social promotion. *Civic Report*, 49, 1-14. Retrieved

- from http://www.manhattaninstitute.org/html/cr_49.htm
- Guskey, T. R. (2001). Helping standards make the grade. *Educational Leadership*, 59(1) 20-27.
Retrieved from <http://www.ascd.org/publications/educational-leadership/sept01/vol59/num01/Helping-Standards-Make-the-Grade.aspx>
- Guskey, T. R. (2005). A historical perspective on closing achievement gaps. *NASSP Bulletin*, 89(644), 76-89. doi:10.1177/019263650508964405
- Guskey, T. R. (2007). Multiple sources of evidence: An analysis of stakeholders' perceptions of various indicators of student learning. *Educational Measurement: Issues and Practices*, 26(1), 19-27. doi:10.1111/j.1745-3992.2007.00085.x
- Haptonstall, K. (2010). *An analysis of the correlation between standards-based, non-standards-based grading systems and achievement as measured by the Colorado Student Assessment Program (CSAP)* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (UMI No. 3397087)
- Hardegree, A. C. (2012). *Standards-based assessment and high stakes testing: Accuracy of standards-based grading* (Unpublished doctoral dissertation). Liberty University, Lynchburg, VA.
- Harlen, W., & Crick, R. D. (2003). Testing and motivation for learning. *Assessment in Education*, 10(2). 169-207.
- Hatch, J. A. (2010). Rethinking the relationship between learning and development: Teaching for learning in early childhood classrooms. *The Educational Forum*, 74, 258-268.
doi:10.1080/00131725.2010.483911
- Hauser, R. M. (1999). *Should we end social promotion? Truth and consequences* (CDE Working Paper No. 99-06). Center for Demography and Ecology: Madison, WI.

- Hauser, R. M., Pager, D. I., & Simmons, S. J. (2000, August). *Race-ethnicity, social background, and grade retention*. Paper presented at the American Sociological Association, Washington, D.C.
- Holmes, C. T. (1989). Grade-level retention effects: A meta-analysis of research studies. In L. A., Shepard, & M. L. Smith (Eds.), *Flunking grades: Research and policies on retention* (pp. 16–33). London, England: The Falmer Press.
- Holmes, C. T. (2006). Low test scores + high retention rates = more dropouts. *Kappa Delta Pi Record*, 42(2), 56-58. Retrieved from <http://search.ebscohost.com/proxy.liberty.edu>
- Holmes, C. T., & Matthews, K. M. (1984). The effects of nonpromotion on elementary and junior high school pupils: A meta-analysis. *Review of Educational Research*, 54(2), 225-236. doi:10.3102/00346543054002225
- Holmes, C. T., & Saturday, J. (2000). Promoting the end of retention. *Journal of Curriculum and Supervision*, 15, 300-314.
- Horn, K. (2003). High stakes testing in students: Stopping or perpetuating a cycle of failure? *Theory into Practice*, 42(1), 30-41. doi:10.1207/s15430421tip4201_5
- Howell, D. C. (2010). Chi-square test: Analysis of contingency tables. In Lovric, M. (Ed.), *International Encyclopedia of Statistical Science* (pp. 250-252). Berlin, Germany: Springer-Verlag.
- Jackson, G. (1975). The research evidence of the effects of grade retention. *Review of Educational Research*, 45(4), 613-635. doi:10.3102/00346543045004613
- Jacobs, B., & Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *The Review of Economics and Statistics*. 86(1), 226-244. doi:10.1162/003465304323023778

- Jimerson, S. (1999). On the failure of failure: Examining the association of early grade retention and late adolescent education and employment outcomes. *Journal of School Psychology, 37*(3), 243-272. Retrieved from <http://search.ebscohost.com/proxy.liberty.edu>
- Jimerson, S. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review, 30*(3), 420-437.
- Jimerson, S., Carlson, E., Rotert, M., Egeland, B., & Sroufe, L. (1997). A prospective, longitudinal study of the correlates and consequences of early grade retention. *Journal of School Psychology, 35*(1), 3-25. Retrieved from <http://search.ebscohost.com/proxy.liberty.edu>
- Jimerson, S. R., Pletcher, S. M., Graydon, K., Schnurr, B. L., Nickerson, A. B., & Kundert, D. K. (2006). Beyond grade retention and social promotion: Promoting the social and academic competence of students. *Psychology in the Schools, 43*(1), 85-97. doi:10.1002/pits.20132
- Jimerson, S., Pletcher, S., & Kerr, M. (2005). Alternatives to grade retention. *Principal Leadership, 5*(6), 11-15.
- Johnson, B. (2000, April). *It's (beyond) time to drop the terms causal-comparative and correlational research in educational research methods textbooks*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. Abstract retrieved from <http://www.eric.ed.gov/PDFS/ED445010.pdf>
- Johnson, D., & Rudolph, A. (2001). *Critical issue: Beyond social promotion and retention—five strategies to help students succeed*. North Central Regional Educational Laboratory. Retrieved from <http://www.ncrel.org/sdrs/areas/issues/students/atrisk/at800.htm>
- Karweit, N. L. (1999). *Grade retention: prevalence, timing, and effects* (Report No. 3). Washington D.C.: Office of Educational Research and Improvement.

- Landrum, T. J., & McDuffie, K. A. (2010). Learning styles in the age of differentiated instruction. *Exceptionality, 18*(1), 6-17. doi:10.1080/09362830903462441
- Leckrone, M., & Griffith, B. (2006). Retention realities and educational standards. *National Association of Social Workers, 28*(1), 53-58.
- Lin, C., & Yang, M. (2008). Improved p-value tests for comparing two independent binomial proportions. *Communications in Statistics - Simulation and Computation, 38*(1), 78-91. doi:10.1080/03610910802417812
- Lin, H., McClintock, S. K., & Williamson, J. M. (2011). Correction for two-group sample size calculation with uncertain group membership. *Journal of Data Science, 9*, 155-170.
- Lorence, J., & Dworkin, A. G. (2006). Elementary grade retention in Texas and reading achievement among racial groups. *Review of Policy Research, 23*(5), 999-1033. doi:10.1111/j.1541-1338.2006.00247
- Lowery, D. R. (2010). *The effect of grade retention on academic achievement* (Unpublished doctoral dissertation). University of South Dakota, Vermillion, SD.
- Lynch, M. A. (2009). Determining eligibility. *Faculty Publications*. Retrieved from 258.<http://digitalcommons.ric.edu/facultypublications/258>.
- Mann-Whitney U test. (n.d.). In *WINKS SDA: Statistical data analysis manual* (6th ed.). Retrieved July 1, 2010, from <http://www.texasoft.com/winkmann.html>
- Marzano, R. J., & Kendall, J. S. (1997). *The fall and rise of standards-based education; a National Association of School Boards of Education (NASBE) issues in brief*. Aurora, CO: Mid-Continent Research for Education and Learning.
- 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.

McCoy, R., & Reynolds, A. (1999). Grade retention and school performance: An extended investigation. *Journal of School Psychology, 37*(3), 272-298.

McDonnell, L., McLaughlin, M., & Morrison, P. (1997). *Educating one and all: Students with disabilities and standards-based reform.* Washington, DC: National Academy Press.

McGill-Franzen, A., Zmach, C., Solic, K., & Zeig, J. (2006). The confluence of two policy mandates: Core reading programs and third-grade retention in Florida. *The Elementary School Journal, 107*(1), 67-91. doi:10.1086/509527

Meisels, S. J., & Liaw, F. (1993). Failure in grade: Do retained students catch? *Journal of Educational Research, 87*(2), 69-77. doi:10.1080/00220671.1993.9941169

Nagaoka, J., & Roderick, M. (2004). Ending social promotion: The effects of retention.

Consortium on Chicago School Research. Retrieved from <http://www.consortium-chicago.org/publications/p70.html>

National Academy of Education. (2009). *Education policy white paper on standards, assessments, and accountability.* L. Shepard, J. Hannaway, E. Baker (Eds.). Washington, DC.

National Association of School Psychologists (NASP). (2003). *Position statement: Student grade retention and social promotion.* Retrieved

http://www.nasponline.org/about_nasp/positionpapers/whitepaper_graderetentionandsocialpromotion.pdf

Nichols, A. (2007). Causal inference with observational data. *Stata Journal, 7*(4), 507–541.

Olson, B., Mead, R., & Payne, D. (2002). *A report of standard setting method for alternate assessments with students with significant disabilities.* Minneapolis, MN: University of

- Minnesota, National Center on Educational Outcomes.
- Paeplow, C. J. (2011). *Exploring the implementation of standards-based grading in Wake County Elementary Schools* (Doctoral Dissertation). Retrieved from ProQuest Dissertations and Theses. (UMI No. 3497209)
- Palinesar, A., (1998) Social constructivist perspectives on teaching and learning. *Annual Review Psychology*, 49, 345-375. doi:0066-4308/98/0201-0345
- Phemister, A.W. (2009). *The effects of Georgia's Choice Curricular Reform Model on third grade science scores on the Georgia Criterion Referenced Competency Test* (Unpublished doctoral dissertation). Liberty University, Lynchburg, VA.
- Picklo, D., & Christensen, S. (2005). Alternatives to retention and social promotion: The availability of instructional options. *Remedial and Special Education*, 26(5), 258- 268. doi:10.1177/07419325050260050101
- Poland, S. (2009). Grade retention: School districts are leaving too many children behind. *District Administration*, 45(10), 104. Retrieved from <http://www.districtadministration.com/viewarticle.aspx?articleid=2212>.
- Powell, P. (2010). A perilous policy path: Grade retention in the age of NCLB. *eJournal of Education (JEP)*. Retrieved from <https://www4.nau.edu/cee/jep/journalterms.aspx?term=1&year=2010>
- Roderick, M. (1995). Grade retention and school dropout: Policy debate and research questions. *Phi Delta Kappa Research Bulletin*, 15, 1-6.
- Roderick, M., & Nagoake, J. (2005). Retention under Chicago's high-stakes testing program: Helpful, harmful, or harmless? *Educational Evaluation and Policy Analysis*, 27(4), 309-340. doi:10.3102/01623737027004309

- Safer, D. J. (1986). Nonpromotion correlates and outcomes at different grade levels. *Journal of Learning Disabilities, 19*(8), 500-503.
- Schenker, J. D., & Rumrill, P. D. (2004). Causal-comparative research designs. *Journal of Vocational Rehabilitation, 21*(3), 117-121.
- Schwager, M., Mitchell, D., Mitchell, T., & Hecht, J. (1992). How school district policy influences grade level retention in elementary schools. *Educational Evaluation & Policy Analysis, 14*(4), 421-438.
- Scriffiny, P. (2008). Seven reasons for standards-based grading. *Educational Leadership, 66*(2), 70-74. Retrieved from http://www.ascd.org/publications/educational_leadership/oct08/vol66/num02/Seven_Reasons_for_Standards-Based_Grading.aspx
- Sengupta, S. (1997, August 6). 14 Children in North Carolina sue a district over testing policy. *The New York Times*.
- Shepard, L. A., & Smith, M. L. (1990). Synthesis of research on grade retention. *Educational Leadership, 47*(8), 84-88.
- Sieling, C. J. (2013). *Standards-Based grading in mathematics: Effects on student achievement and attitude* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (UMI No. 1523840)
- Silberglitt, B., Jimerson, S., & Burns, M. (2006). Does the timing of grade retention make a difference? Examining the effects of early versus later retention. *School Psychology Review, 35*(1), 134-141. Retrieved from <http://www.nasponline.org/publications/spr/pdf/spr351silberglitt.pdf>
- Smirk, J. (2001). Alternatives to retention. *NASSP Bulletin, 85*(629), 3-15.

- Smith, M. L., & Shepard, L. A. (1988). Kindergarten readiness and retention: A qualitative study of teachers' beliefs and practices. *American Educational Researcher*, 25(3), 307-333.
doi:10.3102/00028312025003307
- Smith, M. L., & Shepard, L. A. (1989). *Flunking grades: Research and policies on retention*. New York, NY: Falmer Press.
- Spencer, K. (2010). Standards-based grading. *Education Digest*, 78(3), 4-10.
- Stipek, D. (2001). *Pathways to constructive lives: The importance of early school success*. Washington, D.C.: American Psychology Association.
- Stipek, D., & Byler, P. (1997). Early childhood education teachers: Do they practice what they preach? *Early Childhood Research Quarterly*, 12(3), 305-325.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53-55. doi:10.5116/ijme.4dfb.8dfd
- Texas Educational Agency. (2006). *Grade-level retention in Texas public schools, 2004-05* (No. GE07 601 03). Austin, TX: Texas Education Agency.
- Thomas, J., & Stockton, C. (2003). Socioeconomic status, race, gender, and retention: impact on student achievement. *University of South Carolina Aiken*, 7. Retrieved from <http://www.usca.edu/essays/vol72003/stockton.pdf>.
- Thompson, C. L., & Cunningham, E. K. (2000). *Retention and social promotion: Research and implications for policy*. ERIC Digest #161/ED449241: Retrieved from <http://www.ericdigests.org/2001-3/policy.htm>
- Thurlow, M., & Johnson, D. (2000). High-stakes testing of students with disabilities. *Journal of Teacher Education*, 51(4), 305-314. doi:10.1177/0022487100051004006
- Tomchin, E. M., & Impara, J. C. (1992). Unraveling teachers' beliefs about grade retention.

- American Educational Research Journal*, 29(1), 199-223.
- doi:10.3102/00028312029001199
- Urich, J. L. (2012). *Implementation of standards-based grading at the middle school level* (Unpublished doctoral dissertation). Iowa State University, Ames, IA.
- U. S. Census Bureau, population division (2008). Retrieved from <http://www.census.gov/popest/states/NST-ann-est2008.html>
- U. S. Department of Education. (1999). Taking responsibility for ending social promotion: A guide for educators and state and local leaders. Washington, D.C.: U.S. Government Printing Office.
- U. S. Department of Education. (2010). U. S. Secretary of Education Duncan announces winners of competition to improve student assessments. Washington, D.C.: U.S. Government Printing Office.
- Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
- Wang, J., & Wang H. (2007). *Should grade retention decisions be indicators-based or model-driven* (Report no. 716)? Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Warner, P. (2013). Testing association with Fisher's Exact test. *Journal of Family Planning and Reproductive Health Care*, 39(4), 281-284. doi:10.1136/jfprhc-2013-100747
- West, M. R. (2012). *Is retaining students in the early grades self-defeating* (CCF Brief No. 49)? Washington, D.C.: The Brookings Institution. Retrieved from <http://www.brookings.edu/research/papers/2012/08/16-student-retention-west>.
- Williams, C. (2007). Research methods. *Journal of Business & Economic Research*, 5(3), 65-72.
- Wilson, V., & Hughes, J. (2009). Who is retained in first grade? A psychosocial perspective. *The*

Elementary School Journal, 109(3), 252-266. doi:10.1086/592306

Xia, C., & Glennie, E. (2005). *Grade retention: The gap between research and practice*.

Durham, NC: Duke University, Center for Child and Family Policy, Terry Sanford

Institute of Public Policy.

APPENDIX A: IRB Letter



The Graduate School at Liberty University

January 8, 2013

[Mishea Griffeth Porter]

IRB Exemption 1482.010813: Grade Retention Rates: Impact of Standards-Based Grading as Compared to a Traditional Grading System in 3rd, 4th, and 5th Grades

Dear Mishea,

The Liberty University Institutional Review Board has reviewed your application in accordance with the Office for Human Research Protections (OHRP) and Food and Drug Administration (FDA) regulations and finds your study to be exempt from further IRB review. This means you may begin your research with the data safeguarding methods mentioned in your approved application, and that no further IRB oversight is required.

Your study falls under exemption category 46.101 (b)(4), which identifies specific situations in which human participants research is exempt from the policy set forth in 45 CFR 46:

(4) Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

Please note that this exemption only applies to your current research application, and that any changes to your protocol must be reported to the Liberty IRB for verification of continued exemption status. You may report these changes by submitting a change in protocol form or a new application to the IRB and referencing the above IRB Exemption number.

If you have any questions about this exemption, or need assistance in determining whether possible changes to your protocol would change your exemption status, please email us at irb@liberty.edu

Sincerely,



Liberty University | Training Champions for Christ since 1971