ABSTRACT

Gail T. Pylant, THE EFFECTS OF EARLY VERSUS LATER GRADE RETENTION ON STUDENT ACHIEVEMENT (Under the direction of Dr. James McDowelle). Department of Educational Leadership, April 2011.

In an era of increased accountability for student learning, administrators and teachers are forced to address the needs of struggling students. Grade retention is used as a practice to help low-performing students overcome their academic deficits. The objective of this study was to determine the effects of grade retention on student achievement. The goal of this study was to determine whether there was a significant difference in the growth performance scores and EVAAS predicted scores of sixth grade students retained in early grades first through third as compared to students retained in later grades fourth through sixth utilizing the NC End-of-Grade Tests in reading and mathematics.

Separate Fisher's exact tests for reading and mathematics were used to determine the frequency of distribution of student growth performance scores in these subjects as measured by the within and/or above expectations versus below expectations categories from the NC ABCs model to compare early retainees and later retainees to see if the relationship of performance to category of retention differed more than expected by chance. Also, separate Fisher's exact tests for reading and mathematics were conducted to examine the frequency of distribution of student growth scores in these subjects as measured by the below expectation versus within or above expectations categories from the EVAAS predicted scores to compared early retainees and later retainees to see if the relationship of performance to category of retention differed more than expected by chance.

According to these findings in this study, when examining academic change scores for reading and mathematics a significant relationship does not exist between the growth

performance of early and later retainees. When examining EVAAS predicated scores, it was also indicated that a significant relationship does not exist between the growth performance of early and later retainees. However, the data did indicate that a higher percentage of later retainees met and/or scored above expectations in reading when compared to the early retainees.

THE EFFECTS OF EARLY VERSUS LATER GRADE RETENTION ON STUDENT ACHIEVEMENT

A Dissertation

Presented to

The Faculty of the Department of Educational Leadership

East Carolina University

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

by

Gail T. Pylant

April, 2011

THE EFFECTS OF EARLY VERSUS LATER GRADE RETENTION ON STUDENT ACHIEVEMENT

by

Gail T. Pylant

APPROVED BY:

DIRECTOR OF DISSERTATION:_____
James McDowelle

COMMITTEE MEMBER:_____
Lane Mills

COMMITTEE MEMBER:_____
Kermit Buckner

COMMITTEE MEMBER:_____
Mark Burlis

INTERIM CHAIR OF THE DEPARTMENT OF EDUCATIONAL LEADERSHIP:

_____
William Rouse, Jr.

DEAN OF THE GRADUATE SCHOOL:

_____
Paul Gemperline

ACKNOWLEDGEMENTS

I would like to acknowledge and thank the following people for their support and encouragement throughout my dissertation process. First, my sincere gratitude is extended to my entire dissertation committee of Dr. Mills, Dr. Buckner, and Dr. Bulris. I thank you for your time, your support, and your wisdom. To my dissertation chair, Dr. James McDowelle, I extend my infinite appreciation for your advice, your patience, your positive outlook, your guidance that pushed me further than I thought was possible, and the many hours of assistance you provided. This dissertation would not have been possible without your help.

I would also like to thank Onslow County Schools, and more specifically, the district superintendent, Dr. Kathy Spencer, the district test coordinator, Lisa Thompson, and the district's middle school principals for their support and assistance in completing this research. I hope the conclusions reached in this research study will assist you as you make important decisions about student placement. I would also like to thank my professional colleague and friend, Donna Lynch. Donna kept me focused and motivated.

Finally, I would like to thank my family, for their encouragement during this entire process. Foremost, I would like to thank my husband, Herschel Pylant, for his love, support, and encouragement. There were many weekends that we had to forgo spending time together so I could write. No matter what I choose to pursue in life my husband is always there to support me. I could not have made it without his patience, understanding, and constant encouragement. Thank you so much for believing in me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

This dissertation is a study regarding whether the effects of early grade retention grades first through third are different from those of later grade retention grades fourth through sixth on student achievement in reading and mathematics. The study is undertaken to help educators and parents have a better understanding of the effects of grade retention at the different grade levels. This study is based upon an analysis of the growth performance scores earned by sixth grade students who have been retained in grades first through sixth using the North Carolina End-of-Grade (NCEOG) reading and mathematics tests at the end of the 2008-2009 school year. The analysis will also include the Educational Value-Added Assessment System (EVAAS) predicted scores in reading and mathematics for the retained students. Data will be collected from eight middle school sites located in the Onslow County School System.

The first chapter of the dissertation presents the background of the study, specifies the problem and purpose, and presents the research questions and significance of the study. The chapter concludes by noting delimitations, defining terms, presenting the conceptual framework, describing how the dissertation will be organized, and summarizing the introductory material.

Background of the Study

The purpose of public schools is to educate all students to become productive citizens. School leaders want students to be successful and graduate from high school and either seek employment or continue their education. However, at the end of each school year school principals, given the authority to grade and place, are faced with the dilemma of whether to retain a student who does not possess the necessary skills to progress to the next grade level or socially promote the student instead (Powers and Duties of Principal, 115C-288(a)).

High stake accountability models are a factor in the resurgence of the practice of retention (David, 2008; North Carolina School Psychology Association [NCSPA], 2005). Despite concerns regarding potential negative impacts of retention, the practice has persisted and even increased in recent years. President Clinton challenged the nation in his 1997 State of the Union address to adopt high national standards to end social promotion (Heubert & Hauser, 1999). He stated that by 1999 every state should test every third grader in reading and every eighth grader in math to make sure these standards are met. It was believed that good tests would show educators who needed help, what changes in teaching needed to be made, and which schools needed to improve. Clinton further stated that no child should move from grade school to junior high or junior high to high school until they were ready. As a result of the mandate by President Clinton, many school systems across the nation including North Carolina have enacted policies that require students in particular grades to demonstrate a benchmark level of mastery in basic skills by passing a standardized test before they can be promoted (Heubert & Hauser, 1999).

The ABCs of Public Education mandated in May 1995 in North Carolina reflect the notion of grade retention being implemented to influence higher educational standards (North Carolina Department of Public Instruction [NCDPI], 2008). The ABCs focus on strong accountability with an emphasis on high educational standards, teaching the basics, and maximum local control. Student achievement is measured through state standardized tests. Elementary and middle grades (grades three through eight) use the end-of-grade tests in reading and math, and writing (grades four and seven) to measure student achievement. In 2006-2007, the high school model consisted of the end-of-course test in core subject areas. For students in grades kindergarten through second, special assessments are used in reading, writing, and

mathematics to chart students' academic progress. Students who fail to perform satisfactorily on classroom work, tests, and assessments are targeted for grade retention (NCDPI, 2008).

Proponents of high stake accountability models think that schools do students a disservice by promoting them to the next grade if they do not possess the skills to be successful (David, 2008; Greene & Winters, 2004; Lorence & Dworkin, 2006). They argue that students will be unable to learn the more advanced material in the next grade if they do not understand the subject matter of the current grade. Therefore, the effect of grade retention should be to improve student academic outcomes. On the other hand, those supporting social promotion point to the wide body of research suggesting that students who are retained for an extra year are academically and emotionally harmed by the experience (Burkam, LoGerfo, Ready, & Lee, 2007; Holmes, 1989; Holmes & Matthews, 1984; Hong & Raudenbuah, 2005; Jackson, 1975; Jimerson, 2001; Meisels & Liaw, 1993; Roderick & Nagaoka, 2005).

The research on grade retention indicates mixed results. The strong beliefs about the ineffectiveness of grade retention are primarily derived from four seminal studies conducted by Holmes (1989), Holmes and Matthews (1984), Jackson (1975), and Jimerson (2001). These studies are frequently cited as definitive studies that demonstrate that requiring low-performing students to repeat a grade is a futile educational practice (Lorence & Dworkin, 2006). However, other studies conclude that retention is beneficial; helping students learn the material they missed the first time through the grade (Alexander, Entwisle, & Dauber, 2003; Gleason, Kwok, & Hughes, 2007; Greene & Winters, 2004, 2007; Jacob & Lefgren, 2004; Karweit, 1999; Lorence & Dworkin, 2006; Peterson, DeGracie, & Aybe, 1987; Pierson & Connell, 1992). There is evidence to suggest that retention of students in grades first through third is probably more favorable than for students in grades fourth through sixth (Alexander et al., 2003; Gleason et al.,

2007; Greene & Winters, 2004, 2007; Jacob & Lefgren, 2004; Karweit, 1999; Lorence & Dworkin, 2006; Peterson et al., 1987; Pierson & Connell, 1992).

Problem Statement

In an era of increased accountability for student learning, administrators and teachers are forced to address the needs of struggling students. Grade retention is used as a practice to help low-performing students overcome their academic deficits. The objective of this study is to determine the effects of grade retention on student achievement. The goal of this study is to determine whether there is a significant difference in the growth performance scores and EVAAS predicted scores of sixth grade retained in early grades first through third as compared to students retained in later grades fourth through sixth utilizing the NC End-of-Grade Tests in reading and mathematics. Most of the studies cited in Chapter 2 that support retention indicate that students retained in the early grades first through third demonstrate positive achievement growth.

Purpose of the Study

There are two purposes of this study. The first is to determine whether the timing of retention in early grades first through third as compared to later grades fourth through sixth demonstrate differences in sixth grade reading and mathematics growth performance. The second purpose is to determine whether there are significant differences in the EVAAS predicted scores of sixth grade students who have been retained in grades first through third and grades fourth through sixth in reading and mathematics. To analyze growth performance, the North Carolina End-of- Grade (NCEOG) reading and mathematics tests in the 2008-2009 school year will be used. To analyze EVAAS predicted scores, the SAS Educational Value-Added Assessment System (EVAAS) reporting system from the 2008-2009 school year will be utilized.

Existing research focusing on grade retention and its impact on student achievement indicates mixed results. Some studies indicate that retention is harmful to students and reflects no academic benefits (Holmes, 1989; Holmes & Matthew, 1984; Jackson, 1975; Jimerson, 2001; Meisels & Liaw, 1993; Roderick & Nagaoka, 2005). Other research studies suggest that retention actually increases academic achievement and helps students catch up (Alexander et al., 2003; Gleason et al., 2007; Green & Winters, 2004, 2007; Jacob & Lefgren, 2004; Karweit, 1999; Lorence & Dworkin, 2006; Peterson et al., 1987; Pierson & Connell, 1992). According to Rice (2003), teacher quality is the most important school-related factor influencing student achievement. However, no research studies are available examining teacher effectiveness and grade retention.

Research indicates that students retained in grades first through third demonstrate academic benefits from retention as compared to students in grades fourth through sixth (Alexander et al., 2003; Gleason et al., 2007; Green & Winters, 2004, 2007; Jacob & Lefgren, 2004; Karweit, 1999; Lorence & Dworkin, 2006; Peterson et al., 1987; Pierson & Connell, 1992). Based on these studies this dissertation will analyze the impact of retention on first through sixth grade student achievement at the end of sixth grade to determine whether retention shows greater academic benefits for students retained in early grades as compared to students retained in later grades.

Research Questions

This study is designed to compare the academic achievement in the areas of reading and mathematics of sixth grade students who have been retained in early grades first through third to those who have been retained in later grades fourth through sixth. The following research questions will provide direction for this study:

1. Is there a relationship in the growth performance as measured by the ABC's academic change score categories of within and/or above expectations versus below expectations between sixth grade students retained in grades first through third as compared to sixth grade students retained in grades fourth through sixth in reading in the Onslow County School System?
2. Is there a relationship in the growth performance as measured by the ABC's academic change score categories of within and/or above expectations versus below expectations between sixth grade students retained in grades first through third as compared to sixth grade students retained in grades fourth through sixth in mathematics in the Onslow County School System?
3. Is there a relationship in the growth performance as measured by the EVAAS predicted score categories of met and/or above versus below expectations between sixth grade students retained in grades first through third as compared to sixth grade students retained in grades fourth through sixth in reading in the Onslow County School System?
4. Is there a relationship in the growth performance as measured by the EVAAS predicted score categories of met and/or above versus below expectations between sixth grade students retained in grades first through third as compared to sixth grade students retained in grades fourth through sixth in mathematics in the Onslow County School System?

Significance of the Study

National, state, and local legislation and policies mandate an increase in academic standards and prohibit social promotion. Grade retention is used as a remediation practice to

increase student achievement. Research examining the effects of grade retention on academic achievement indicates mixed results. However, there is some indication that students retained in first, second, or third grade show positive gains in achievement as compared to students in grades four through six (Alexander et al., 2003; Gleason et al., 2007; Greene & Winters, 2004, 2007; Jacob & Lefgren, 2004; Karweit, 1999; Lorence & Dworkin, 2006; Peterson et al., 1987; Pierson & Connell, 1992). Based on these studies retention in grades first through third might be an effective early intervention strategy to use to help underachieving students.

In the Onslow County School System located in North Carolina, the total number of students retained in grades first through sixth from 2001-02 to 2005-06 consisted of 1,992 students (1,355 in grades first through third) and (637 in grades fourth through sixth) (NCDPI, 2001, 2005). Examining the academic achievement of sixth grade students retained in early grades first, second, or third as compared to later grades fourth, fifth, or sixth could aid in decisions regarding student educational placement for that school district which will benefit future students.

Delimitations of the Study

The limitations of this study follow:

1. All data collected is from one school system. Since this study is conducted using participants from one North Carolina School District results may not be generalized to other locales. The advantages of using results from one school district outweigh the drawbacks. Policies mandated by local school boards for retaining students vary widely across locales. Using data from one school system eliminates the need to attempt to compensate for the inevitable differences. Thus, the criteria for retaining students are the same.

2. In this study, there is no way to account for teacher effectiveness. However, in North Carolina teachers are required to hold a license for the subject or grade level they teach. To obtain a license and be considered highly qualified, teachers must meet minimum requirements related to content knowledge and teaching skills (NCDPI, 2010). By having this uniformed method of standards in place, each teacher that holds a license is deemed to possess a minimal level of effectiveness.
3. In this study, curriculum differentiation for reading and mathematics will not be considered.

Definition of Terms

The following terms are defined for the purpose of this study.

Achievement: In this study, achievement will be measured by performance on the North Carolina End-of-Grade tests.

Developmental scale score (DSS): The developmental scale score is a converted raw score (number of questions answered correctly) that allows for comparison of the student's end-of-grade scores by subject from one grade to the next. The DSS is like a ruler that measures growth in reading and mathematics from year to year. The DSS is the scale score measure assigned to each student score on standardized achievement tests.

EVAAS Predicted Score: A student's predicted score is an expected score, based on his or her performance on previous tests, assuming the student is in an average school in the state (North Carolina Resource Guide for Value-Added Reporting, 2007).

North Carolina End-of-Grade tests: Tests designed to measure student performance on the goals, objectives, and grade level competencies specified in the North Carolina Standard Course of Study.

Raw Score: The number of questions a child answers correctly on the NCEOG tests. The raw score is converted to a developmental scale score.

Retention: The practice of having a student repeat the entire grade just completed.

SAS EVAAS: A customized software system used by NC school districts to analyze student progress each year.

Social Promotion: The practice of advancing a student to the next grade level based on age rather than achievement.

Standardized tests: Tests constructed by measurement experts over a period of years. They are designed to measure broad, national objectives, and have a uniform set of instructions that are adhered to during each administration. Most also have tables of norms, to which a student's performance may be compared to determine where a student stands in relation to a national sample of students at his or her grade or age level (Kubiszyn & Borich, 1987, p. 21).

Conceptual Framework

Research conducted by Bransford, Brown, and Cocking (1999) on the transfer of learning states that all new learning involves transfer based on previous learning. The first factor that influences successful transfer is the degree of mastery of the original subject. Without an adequate level of initial learning, transfer cannot be expected. This research implies that learning cannot be rushed; the complex cognitive activity of information integration requires time. This theory is supported by Piaget's view on how children's minds work and develop that suggests that children cannot undertake certain tasks until they are psychologically mature enough to do so (Atherton, 2005). Piaget proposes that children's thinking does not develop entirely smoothly; instead, there are certain points at which it takes off and moves into completely new areas and capabilities. Based on how learning is transferred and Piaget's theory on how

children's minds work, retention appears to be an appealing method to use to deal with underachieving students.

Organization of the Study

This study is organized in five chapters including (a) introduction, (b) review of the literature, (c) methodology, (d) results, and (e) conclusions. The literature review highlights key research examining issues surrounding retention. A synthesis of the available literature regarding the effects of retention on student achievement is presented including (a) seminal research, (b) kindergarten research, (c) opposing research, and (d) supporting research. The study's methodology focuses on the achievement of sixth grade students who have been retained in early grades first through third as compared to students retained in later grades fourth through sixth to determine whether there is a significant difference in the students' growth performance scores and EVAAS predicted scores in reading and mathematics. The results section includes all relevant findings that emerge from this research and the final portion of the study includes a discussion of those findings.

Summary

Schools are held accountable by federal, state, and district policy makers for student performance on the end of grade tests. Growth performance for students by subject from one grade to the next is a source of information used to help determine a student's placement in grade. For this reason, the growth performance a student makes is important to educators and parents. Grade retention is a strategy used to remediate underachieving students that fail to meet expected growth on the end of grade tests. Examining the growth performance scores and EVAAS predicted scores of sixth grade students who have been retained in early grades first

through third as compared to later grades fourth through sixth could aid in decisions regarding future student educational placement for the students in one school district.

CHAPTER 2: REVIEW OF THE LITERATURE

Introduction

This review of literature is presented in five sections. The first section defines retention and social promotion. The second section reviews the history of retention in the United States. The third section is reviewed under two headings: (a) the North Carolina Accountability Program; and (b) the frequency and timing of retention in North Carolina. The fourth section deals with the reasons for and prevalence of retention. The final section deals with the effects of retention on student achievement and is reviewed under four headings: (a) seminal research; (b) kindergarten research (c) opposing research in grades first through sixth; and (d) supporting research in grades first through sixth. The chapter concludes with a summary of the literature.

Retention and Social Promotion Defined

Grade retention occurs when students are not allowed to progress to the next grade with their age-mates in an attempt to provide them with a year to mature or a year to improve their academic performance (Jimerson, 2001; Lorence & Dworkin, 2006). The practice of retention has been studied since the early part of the twentieth century with mixed results in regard to academic achievement outcomes. Some research studies indicate that retention is harmful to students and shows no academic benefits (Burkam et al., 2007; Holmes, 1989; Holmes & Matthews, 1984; Hong & Raudenbush, 2005; Jackson, 1975; Jimerson, 2001; Meisels & Liaw, 1993; Roderick & Nagaoka, 2005). Other studies suggest that grade retention especially in grades first through third increases academic achievement and is an effective strategy to use to close the achievement gaps (Alexander et al., 2003; Gleason et al., 2007; Greene & Winters, 2004, 2007; Jacob & Lefgren, 2004; Karweit, 1999; Lorence & Dworkin, 2006; Peterson et al., 1987; Pierson & Connell, 1992).

Schools and communities use grade retention as a strategy to remedy academic failure and increase academic standards. In the No Child Left Behind (NCLB) legislation passed in 2001, retention was used as a measure to respond to teacher accountability and inadequate student progress (Lorence & Dworkin, 2006; NCSPA, 2005; U.S. Department of Education, 2008). The NCLB legislation outlines minimum qualifications needed by teachers to demonstrate competency. The requirements include passing a rigorous state test on content knowledge and teaching skills (U.S. Department of Education, 2004).

Statewide accountability testing is one of the foremost reasons for the resurgence of the practice of retention (David, 2008; NCSPA, 2005). Supporters of high stake accountability models believe that grade retention will help maintain or increase educational standards by compelling children to meet certain grade level competencies before moving on to the next grade (Darling-Hammond, 1998; David, 2008; Gleason et al., 2007; Lorence & Dworkin, 2006). Retention has been used as a way to foster educational reform (David, 2008).

Social promotion is the practice of promoting students to the next grade regardless of their academic achievement. Social promotion has been viewed as one of the reasons for a perceived decline in educational standards (Lorence & Dworkin, 2006). Therefore, by implementing stricter standards by developing the promotion gateways at some grade levels will increase the worth of public education (Greene & Winters, 2004; Lorence & Dworkin, 2006).

History of Retention

The history of retention as an educational practice to remediate underachieving students began in schools in the 19th century. Schools were first legislated in Massachusetts in 1647 to ensure that children could learn to read the bible. Because families lived in small settlements far away from each other, children of all ages were taught in a single class in a one-room ungraded

school. The major focus of education during this time was on morality and religion and little attention was paid to reading excellence; therefore, children who could not read were merely absorbed into the classroom community (Frey, 2005).

As the population in the United States increased in the 1800s, so did the focus in education. The schools were beginning to move away from just teaching about religion and saw education as a means of equalizing citizens and reorienting immigrants to the philosophy of democracy. Although what was being taught in schools was changing, little changed in how the schools delivered instruction. Compulsory education was still a policy in only a few communities. The policy excluded African Americans, and rarely included girls or children older than 10. Students who failed to achieve would drop out of school and go to work (Frey, 2005).

The advent of the Industrial Revolution and the influx of immigrants and freed slaves changed education and gave rise to the practice of retention. Compulsory education was implemented which helped to supply educated workers to staff the factories and mills. At the same time, the population in the urban areas was growing rapidly which changed the structure of schools. Increased enrollment allowed schools to specialize both in what was taught and in how it was taught. More subjects were added into the curriculum and children were placed in graded classrooms according to their chronological age (Frey, 2005).

With the replacement of the one-room schoolhouse by the graded school in the mid-19th century, students began to be promoted based on merit. Merit promotion is the practice of requiring students to demonstrate mastery of an inflexible academic standard for each grade level (Steiner, 1986). During this time, grade retention was so common that it had been estimated that

approximately half of all students were retained at least once during their first eight years of school (Steiner, 1986).

In the 1930s, attitudes began to change about the role of schooling. Interest in child psychology became a focus for educators and retention was used more sparingly, making social promotion a common practice (Fager & Richen, 1999; Steiner, 1986). Social promotion is the practice in which children pass to the next grade with their age peers who have failed to master the material of the previous grade, receiving remedial academic help when necessary (Steiner, 1986). One reason for this policy change based on years of research was that retention did little to help the low performing students and was actually more harmful in emotional and psychological terms (Fager & Richen, 1999; Steiner, 1986).

By the 1980s, a backlash against social promotion appeared. Opponents of social promotion argued that a lack of fixed academic standards symbolized a disregard for achievement and influenced children's motivation to learn (Steiner, 1986). The backlash was "intensified in *A Nation at Risk: The Imperative for Educational Reform*, a 1983 report by President Reagan's National Commission on Excellence in Education that brought an increased awareness to the school reform movement" (Fager & Richen, 1999, p. 5). The report called for an end to social promotion and advocated for stricter promotion policies to reverse the trend of declining national test scores. The push to end social promotion was supported by Americans which was reflected in the 1986 Gallup Poll, where 72% of the U.S. citizens favored stricter grade-to-grade promotion standards (Fager & Richen, 1999).

State testing and increased standards were the focus for educational reform in the mid-1990s. President Bill Clinton in his 1997 State of the Union Address added political pressure to the demand for merit promotion by directing states to implement policies to end social

promotion. Clinton stated that students needed to meet rigorous academic standards at key transition points in their schooling (Heubert & Hauser, 1999). On a national level, in-grade retentions grew dramatically during the early 1990s affecting approximately 2.6 million children each year (Jimerson, 2001). By 2000, 49 states had implemented strict promotion policies (Owings & Kaplan, 2001).

Research indicates that certain classifications of children are retained more frequently. Students most at risk for retention are male; African American or Hispanic; have a late birthday; have delayed development and/or attention problems; have behavior problems; show signs of immaturity; live in poverty or in a single-parent household; parents are not well educated; have parents who are less involved in their education; or who have changed schools frequently (Anderson, Whipple, & Jimerson, 2002; Position statement on student grade retention and social promotion, 2003).

In research conducted by the Institute of Education Sciences (IES) National Center for Education Statistics (NCES) in 1992 and 1995 on the dropout rates in the United States, the Current Population Survey (CPS) data indicated that the number of students retained increased from 11.1% in 1992 to 13.3% in 1995 (U.S. Department of Education, 1995). Nearly all of this increase occurred in the early elementary grades kindergarten through third. African American males were retained more frequently than students from other race-ethnicity groups. While retention rates increased for nearly all regions and income groups between 1992 and 1995, the distribution remained relatively unchanged. Children living in the South and from families with the lowest incomes were at the greatest risk of retention.

Between 1995 and 2004, the percentage of students retained in kindergarten through grade five decreased from 11% to 5%, while the percentage of students retained in grades six

through twelve was not measurably different between the two years 7% and 5%. In 1995, more students were retained in the early grades kindergarten through third than in the later grades four through twelve, but in 1999 and 2004, there were no measurable differences by grade level (U.S. Department of Education, 2006).

In 2004, a greater percentage of males than females 13 vs. 6% and a greater percentage of African Americans than Whites 16 vs. 8% were retained (U.S. Department of Education, 2006). More students whose families were in the lowest income quarter were retained than students whose families were in the middle or highest income quarter. This data are consistent with the 1995 IES NCES data (U.S. Department of Education, 1995; U.S. Department of Education, National Center for Education Statistics, 2006).

North Carolina Accountability Program

North Carolina was recognized in Education Week's 1999 state-by-state report card for being one of the only two states close to having all the components of a complete accountability system (Jerald & Boser, 1999). This system the ABCs of Public Education mandated in May 1995 relies heavily upon high stakes testing to measure student achievement. At the time of this study, elementary and middle grades used the end-of-grade tests in reading, writing (grades four and seven), and mathematics to measure school performance. In high school, the end of course (EOC) test results in core subject areas were used as part of the student's final grade. For students in grades kindergarten through second, the K-2 Assessments were used in reading, writing, and mathematics to monitor student performance and make promotion decisions (NCDPI, 2008).

In the 1996-1997 school year, the ABCs implementation began in grades kindergarten through eight. At the end of the year, the first ABCs Accountability Report was submitted to the

State Board of Education focusing on school performance. Low performing schools were identified and assistance teams were assigned to try to increase student performance. Schools that achieved exemplary growth standards received incentive awards. The high school accountability model was implemented in the 1997-1998 school year, and the results of all schools under the ABCs program were reported in A Report Card for the ABCs of Public Education, Volume I (NCDPI, 2008).

On January 27, 1999, another component was added to the ABCs Accountability program. Proficiency benchmarks referred to as the gateways were put in place to make sure students in grades three, five, and eight were ready to progress to the next grade level. The gateway from grade five to grade six was implemented in 2000-2001. The gateway from grade three to grade four and from grade eight to grade nine was implemented in 2001-2002. The high school gateway involved meeting local and state graduation requirements, getting a passing score on an exit examination, and garnering a passing score on a computer test. This policy took effect with the graduating class of 2003.

The gateways require that students meet local promotion standards as well as receive a level III or IV on the end-of-grade tests in both reading and mathematics. Students who fail to perform satisfactorily on the student assessment after three attempts are retained. There are some special considerations where students may be promoted to the next grade if they meet one of the following criteria: (1) achieve a score of l standard error of measurement below Level III in reading and/or mathematics; (2) achieve a score of 2 standard error of measurement below Level III in reading and/or mathematics, with supporting documentation; (3) retained previously; or (4) an exceptional child who has state tests administered by the NC alternate Assessment Academic Inventory or NC Alternate Assessment Portfolio or is making adequate progress on

the Individual Education Program (IEP). If a student scores a Level I or II on the EOG tests, they are to receive focused intervention the following school year. Individual schools determine what focused intervention strategies are to be used (NCDPI, 2008).

In 2010, the NC State Board of Education did away with the gateways. However, the Onslow County School System's promotion requirements mandate that students in grades three through eight demonstrate proficiency with an achievement score at Level III or above in both reading and mathematics on the NC End-of Grade test to be promoted to the next grade. In this study, the Onslow County School System's local promotion standards will be used to serve as the gateways.

Frequency and Timing of Retention in NC

In North Carolina in the ten years from 1991-92 to 2001-02, annual retentions have been climbing steadily from 3.2% to 5.3% (NCDPI, 2004b). In 2001-02, 68,000 students from grades kindergarten through twelve were retained. The data revealed that over 22,000 students were retained in grades kindergarten through third. From 1991-92 to 2001-02, the rates of retention doubled in grades kindergarten through third (NCDPI, 2003). However, from 2001-02 to 2005-06 there was a slight decline in the number of students retained in each grade kindergarten through third (NCDPI, 2006a). In grades four through eight from 2001-02 to 2005-06, the number of students retained decreased from 16,055 students to 11,000 students. The North Carolina Department of Public Instruction (2006a) data indicated that in grades nine through twelve from 2001-02 to 2005-06 all grades had an increase in the number of students retained. During the five years from 2001-02 to 2005-06 (the most recent data available) a total of 80,706 ninth grade students were not promoted, over twice as many as the 33,297 kindergartners, the second largest group of retainees (NCDPI, 2006a).

In the Onslow County School District located in North Carolina, of the total number of students reported in kindergarten through grade twelve in the 2001-02 school year, 1,273 or 5.6% were retained (NCDPI, 2006a). The total number of students retained decreased from 2001-02 to 2005-06 to 4.2%. The percentage of students retained varied by grade (see Table 1). The highest percentage of students retained was in grade nine. This trend showed little variation over the five year period (see Figure 1). In the elementary grades for all five years, the highest number of students retained was in first grade followed by kindergarten. In the elementary grades from 2001-02 to 2005-06, 2,088 students were retained. In the secondary grades, 3,157 students were retained. The total number of students retained in all grades over the five year period was 5,245. At an average per pupil cost of $6,264, the school district spent over $33 million for the extra year of schooling for the retained students (NCDPI, 2006a).

Reasons for and Prevalence of Retention

Not all children struggle at school for the same reason (Alexander et al., 2003; Karweit, 1999), as highlighted in a U.S. Department of Education (1999) report on social promotion:

> Some students have learning disabilities, others have behavioral problems, are not ready for school, or face other challenges in their families and in their lives outside school. Some students barely miss meeting the standards, while others perform at levels considerably behind their peers. The point is that in order to help all students meet standards, educators must understand the nature of children's difficulties, and they must do so early. (p. 32)

According to Alexander et al. (2003), "across the board solutions will not work and the available research offers little guidance for determining individual placements" (p. 249). To find the best course for each individual child requires understanding of what underlies a particular

Table 1

Retentions in Onslow County School District Five Year Review by Grade Level

	K	Gr1	Gr2	Gr3	Gr4	Gr5	Gr6	Gr7	Gr8	Gr9	Gr10	Gr11	Gr12
2001	98	130	77	77	51	31	101	95	43	315	130	86	39
2002	121	127	68	59	34	15	67	31	17	234	125	76	33
2003	103	118	70	55	15	11	65	34	22	232	99	60	25
2004	102	156	87	45	24	12	89	61	23	248	112	63	20
2005	79	166	73	47	28	9	85	48	47	237	105	64	26

Note. (North Carolina Department of Public Instruction, 2001, 2005). Retrieved from http://www.ncpublicschools.org/fbs/resource/data/

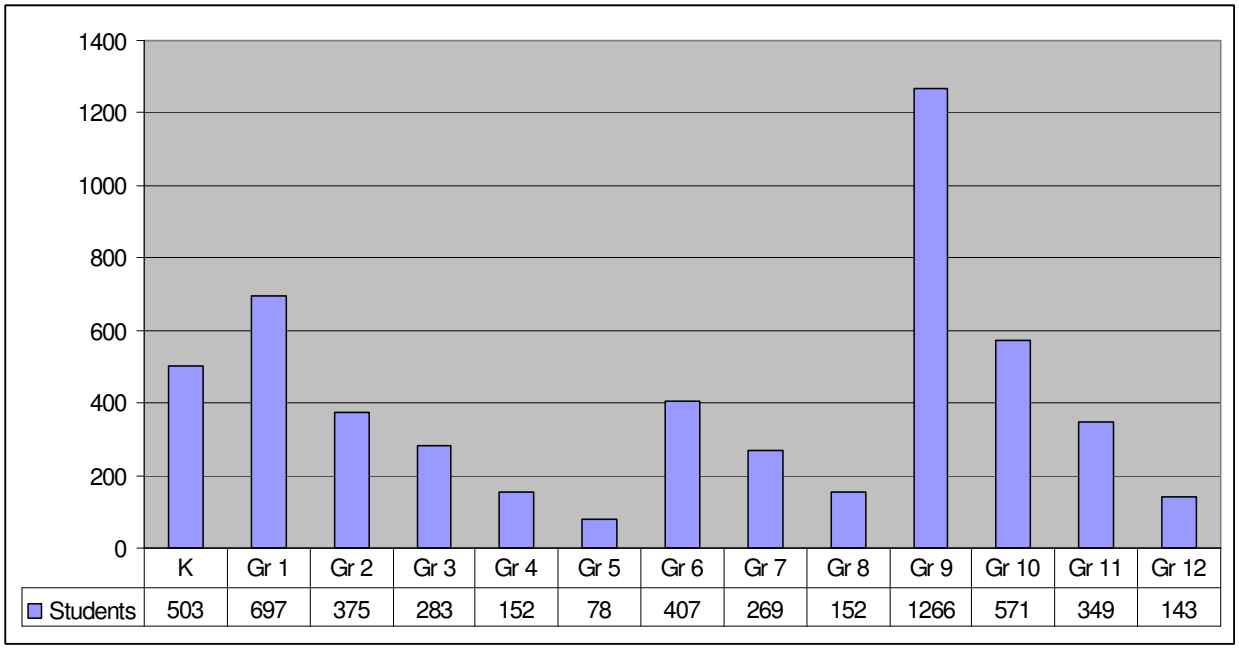

Figure 1. Total retentions (5,245) by grade in Onslow County School District 2001-2005.

Note. (North Carolina Department of Public Instruction, 2001, 2005). Retrieved from http://www.ncpublicschools.org/fbs/resource/data/

child's difficulties. This must be factored in for each child. For example, what may put a child at risk in a high poverty school of failure may be different for low poverty schools where few children fail.

The Retention/Promotion Checklist is a research-informed diagnostic inventory (Grant & Richardson, 1998). The checklist encourages educators to think about retention especially when it is advised that additional learning time is needed for misplaced students. The children who are most likely to be candidates for additional learning time are usually "younger students in a class, emotionally immature children of average or high ability, and children who are small for their age" (Kelly, 1999, p. 2).

Grade retention is as much as an organizational problem as it is personal (Alexander et al., 2003). Grade retention allows time, but the extra time is within the framework of the school with traditional scheduling and calendar constraints. All children are expected to be ready for first grade by age six. They are expected to move in a lockstep system annually thereafter from one grade to the next; and master the set curriculum in the same time frame. This calendar-driven model of schooling sets a rigid pace and if students are not ready to learn the information they are left behind (Alexander et al., 2003).

The primary grade curriculum is believed to be the foundation for all later learning so children need to master it before they move on (David, 2008; Lorence & Dworkin, 2006). But children do not always mature at the same time; learn at the same pace; or develop in the same ways. Mastering the curriculum takes longer for some children than others.

Effects of Retention

The evidence is quite clear and nearly unequivocal that the achievement and adjustment of retained children are no better and in most instances are worse than those of comparable

children who are promoted. Retention is one part of the current reform packages that does not work (Smith & Shepard, 1987, p. 134).

Those who continue to retain pupils at grade level do so despite cumulative research evidence that the potential for negative effects consistently outweighs positive outcomes---- the burden of proof legitimately faces on proponents of retention plans to show there is compelling logic indicating success of their plans where so many other plans have failed (Holmes & Matthews, 1984, p. 232).

These are the conclusions most often indicated when the results of research on grade retention are shared (Alexander et al., 2003). The judgments quoted reflect the point of view that prevailed in academic circles in the wake of the Excellent Commissions back to basics call for rigorous enforcement of high academic standards.

Beginning with the publication of A Nation at Risk and culminating with the passage of the No Child Left Behind Act of 2001; educators have been forced to address the needs of academically challenged students who struggle to learn the necessary material to be promoted to the next grade (Gleason et al., 2007; Lorence & Dworkin, 2006; National Commission on Excellence in Education, 1983; U.S. Department of Education, 2008). Although there is overwhelming agreement that students failing to meet grade level requirements should be helped, there are little consensuses as to which remediation strategies are best for enabling low-performing students to meet the accountability standards. Grade retention is a commonly used practice to help low-performing students catch up academically. Educators believe that giving students an additional year to learn the material will provide them with the foundation needed to proceed successfully through the remainder of their education (David, 2008; Greene & Winters, 2004; Lorence & Dworkin, 2006). It is often assumed that students will be unable to learn the

more advanced material in the next grade if they do not understand the subject matter of the current grade (Lorence & Dworkin, 2006). Therefore, the earlier a child is retained the better.

Opponents of retention contend that making low-performing students repeat a grade is detrimental to their academic achievement. "Retention is viewed as ineffective because gains in academic achievement during the repeated year are presumed to be either negligible or quickly fade if they do occur" (Lorence & Dworkin, 2006, p. 999). Instead of requiring students to repeat a grade, it is better to place students in the next grade even if they have not mastered all the material required for promotion. It is believed that if appropriate supplemental instruction is provided during the promoted year, low-performing students will be able to catch up with their classmates who were not experiencing academic difficulty. This practice of social promotion is presumed to be more beneficial than grade retention because students can remain with the same classmates and the students are not perceived as failures (Lorence & Dworkin, 2006). Proponents of social promotion argue that grade retention damages the child's self esteem causing the child to become alienated and psychologically withdraws from school (Burkam et al., 2007; David, 2008; Holmes, 1989; Holmes & Matthews, 1984; Hong & Raudenbuah, 2005; Jackson, 1975; Jimerson, 2001; Meisels & Liaw, 1993; Roderick & Nagaoka, 2005). Grade retention affects a student's educational progress and will eventually lead to the student dropping out of school (Bowman, 2005).

There is a substantial amount of research that has examined the effects of retention. Some studies have examined student performance before and after retention (Alexander et al., 2003). Other studies have examined student performance after retention (Gleason et al., 2007; Greene & Winters, 2004, 2007; Jacob & Lefgren, 2004; Karweit, 1999; Lorence & Dworkin, 2006; Peterson et al., 1987; Pierson & Connell, 1992; Roderick & Nagaoka, 2005). Also, there

have been studies conducted evaluating retention policies based on subjective criteria such as teachers' evaluations that students should be retained (Greene & Winters, 2004). Yet, no research studies are available that have examined the impact the teacher has on student achievement and grade retention. However, in North Carolina, teachers are required to obtain a teaching license which ensures that teachers are competent and qualified to teach (U.S. Department of Education, 2004). The next four sections focus on the effects of retention on student achievement. The research that supports retention suggests that retention may be more favorable for students in grades first through third as opposed to grades fourth through sixth (Alexander et al., 2003; Gleason et al., 2007; Greene & Winters, 2004, 2007; Jacob & Lefgren, 2004; Karweit, 1999; Lorence & Dworkin, 2006; Peterson et al., 1987; Pierson & Connell, 1992).

Seminal Research (1975-2001)

The strong beliefs about the ineffectiveness of grade retention are primarily derived from four seminal studies that are mentioned in the retention research. These include: Holmes (1989), Holmes and Matthews (1984), Jackson (1975), and Jimerson (2001). These studies include many retention studies spanning a large part of the twentieth century. These studies are frequently cited as definitive studies that demonstrate that requiring low-performing students to repeat a grade is a futile educational practice (Lorence & Dworkin, 2006).

Jackson (1975) provided the first systematic, comprehensive overview of the research evidence on the effects of grade retention. Jackson's research included 30 studies published between 1911 and 1973. Jackson set out to examine whether low-achieving students or those with social emotional maladjustment benefited from grade retention or promotion to the next

grade. The studies were divided into three groups based on their design type: naturalistic, pre-post, and experimental.

In his studies, Jackson (1975) acknowledged the results of the naturalistic students were clearly different from those of pre-post studies; however, they were not contradictory. Jackson concluded "There is no reliable body of evidence to indicate that grade retention is more beneficial than grade promotion for students with serious academic difficulties" (p. 627). Jackson further concluded that there was a lack of an empirical base supporting the use of either social promotion or retention because the flaws in the study designs could have contributed to the negative effects.

Holmes and Matthews (1984) conducted a meta-analysis on 44 studies that were published between 1929 and 1981. The analysis examined the effects of retention on elementary and junior high school students' achievement and social emotional outcomes. The participants included 4,208 retained students and 6,924 regularly promoted students. Studies selected to investigate the effects of retention contained sufficient data to estimate an effect size. An effect size is a statistic that encodes the critical quantitative information from each relevant study finding. In all, 575 effect sizes were calculated. The dependent variables measured consisted of academic achievement, personal adjustment, self-concept, attitude toward school, and attendance.

Of the 44 studies examined by Holmes and Matthews (1984), 31 measured the effect of nonpromotion on students' academic achievement. These studies yielded an overall effect size of -.37, indicating that the promoted students on the average had achieved .37 standard deviation units higher than the retained group, ($p < .001$). When effect sizes were calculated for specific content area, they continued to produce negative mean effect size values that were statistically

significant. The effect sizes translated into a loss of grade equivalent units ranging from a few months to a half of year. When the data were analyzed by grade level in which retention took place grades first through sixth, negative effects were found at all grade levels. Holmes and Matthews (1984) concluded that "those who continue to retain pupils at grade level do so despite cumulative research evidence showing that the potential for negative effects consistently outweighs positive outcomes" (p. 232).

Holmes (1989) conducted a subsequent meta-analysis including 19 more studies to the original meta-analysis conducted by Holmes and Matthews (1984). This meta-analysis consisted of 63 controlled studies published between 1925 and 1989 where retained students were followed and compared to promoted students.

Holmes (1989) reported that of the 63 studies, 54 studies reported overall negative effects associated with grade retention while nine studies yielded positive effects. The positive effects of retention appeared to diminish over time. Holmes concluded that when the overall findings were considered, greater negative effects were shown in studies where retained and promoted student controls were matched on IQ and past achievement scores.

More recently, Jimerson (2001) conducted a meta-analysis using 20 studies that examined the association between retention and academic achievement spanning from 1990-1999. Fourteen of the studies included students retained during kindergarten, first, second, and third grades. Six studies included students retained in kindergarten through eighth grade. The outcome of each analysis was coded with respect to its significance. An alpha level of - < .05 was established as the criterion for statistically significant outcomes.

The 20 studies yielded a total of 175 analyses exploring academic achievement outcomes of retained students relative to a comparison group of promoted students. Ninety-one of the

studies revealed statistically significant differences. Nine of the analyses favored the retained students relative to the comparison group of promoted students, whereas 82 favored the comparison group of promoted students relative to the retained students. Eighty-four yielded no statistically significant differences between the retained and comparison students. The results indicated that 47% of the analysis favored the matched comparison group of promoted students, 5% favored the retained students, and 48% indicated no significant differences between the two groups. It was noted that six of the nine analyses favoring the retained students reflected differences during the repeated year. Jimerson (2001) concluded that on average, retained students scored .31 standard deviations lower than students who were not retained on various academic achievement measures.

In summary, the seminal research conducted by Holmes (1989), Holmes and Matthews (1984), Jackson (1975), and Jimerson (2001) included 113 studies published between 1911 and 1999. The grade levels examined analyzing student achievement varied between the studies. The major conclusions cited in all four studies suggested that retention at any grade level was not academically beneficial to the students.

Kindergarten Retention (2005-2007)

In this study, two of the most recent kindergarten studies will be included but they are not pertinent to this study. Kindergarten retention was not included in this study because there is no way to monitor if students began kindergarten when age-eligible. Research indicates that some parents choose to delay kindergarten entry based on observable and unobservable factors (Holloway, 2003). However, the typical elementary school consists of grades kindergarten through fifth. Including kindergarten research will present a thorough overview of the effects of retention in the elementary grades.

One area of controversy is early grade retention. This is the practice that is based upon retaining children for immaturity or for literacy deficits (Karweit, 1999). Researchers advocate early grade retention because it does not carry the stigma of retention in the later grades (Pierson & Connell, 1992). However, recent studies analyzing the impact of retention on kindergarten achievement have yielded negative results.

Burkam et al. (2007) analyzed the effects of retention on kindergarten students' academic performance to determine if holding low-achieving kindergartners back a year would help them to catch up academically. The data used in this study were drawn from the Early Childhood Longitudinal Study-Kindergarten Class (ECLS-K) of 1998-1999 sponsored by the National Center for Education Statistics (NCES). The participants included 12,780 children, in 915 schools, from the 1998-1999 kindergarten cohort.

Student performance on tests in literacy and mathematics at the end of kindergarten and the end of first grade was examined. Burkam et al. (2007) analyzed the learning trajectories of three groups: (a) children who were repeating kindergarten during the 1998-1999 school year; (b) children who were enrolled in kindergarten for the first time during the 1998-1999 school year and who were promoted to first grade for the 1999-2000 school year; and (c) children who were enrolled in kindergarten for the first time in 1998-1999 school year but who repeated kindergarten in the 1999-2000 school year. No data were available on the performance of the ECLS-K kindergarten repeaters during their first year of kindergarten (the 1997-1998 school year).

The scores on literacy and mathematics tests were equated using Item Response Theory that allowed the tests to be used to measure change over time. The achievement scores were standardized into ES units based on the standard deviation of the achievement scores.

After examining the data after 2 years, Burkam et al. (2007), highlighted several important trends that arose from the data:

> First, all children were learning literacy and mathematics skills over the first 2 years of school. Second, kindergarten repeaters began their second year of kindergarten with a slight cognitive advantage in literacy and mathematics achievement, relative to the first timers who were promoted after 1 year in literacy 24.5 vs. 23.3, SD = .14. First time kindergartners who were retained entered kindergarten with significantly lower literacy and mathematics achievement than other first time kindergartners in literacy 17.6 vs. 23.3, SD = -.66. It is reasonable to suspect that the kindergarten repeaters in 1998-1999 entered school in 1997 with similarly lower achievement. Third, over the 2 years, the data indicated an increasing separation in achievement especially in literacy between the kindergarten repeaters and the first time kindergartners who were promoted, with the repeaters falling further behind in literacy, 51.2 vs. 57.2, SD = -.47. The first time kindergartners who were retained ended their initial kindergarten year with achievement levels comparable to the level that other first time kindergartners displayed at the beginning of the kindergarten year. Fourth, regardless of the group retained, promoted, or first time kindergarten retained, the achievement spread across all children increased over time. (p. 121)

Next, Burkam et al. (2007) analyzed the effects of repeating kindergarten on achievement at the end of kindergarten. The data suggested that children who repeated kindergarten ended the school year somewhat behind the first time kindergarteners in literacy achievement (ES = -.11). The children who would be repeating kindergarten the next year were substantially behind other first time kindergartners in literacy achievement (ES = -.88). After making adjustments for child

and school characteristics, the large cognitive difference between first timers who were promoted and first timers who were retained was reduced by over two-thirds. Mathematics achievement at the end of the kindergarten year indicated similar results. Kindergarten repeaters and first timers who were retained lagged behind the other first timers who were promoted (ES = -.10 and -.85, respectively).

The learning trajectories of the kindergarten repeaters were analyzed after they progressed through the first grade. The data indicated that the kindergarten repeaters exhibited lower levels of literacy achievement at the end of the first grade than did children who spent only 1 year in kindergarten (ES = -.47). After adjustments were made for child and school characteristics, a significant first grade literacy disadvantage still remained (ES = -.24). In mathematics, kindergarten repeaters demonstrated lower levels of achievement at the end of the first grade (ES = -.31), even after controlling for entering differences (ES = -.14). However, Burkam et al. (2007) did report that children with higher mathematics achievement at the end of kindergarten, who repeated kindergarten, outperformed their peers in mathematics at the end of the first grade year.

The data suggested that repeating kindergarten appeared to have a negative impact on early literacy and mathematics development. When same-grade comparisons were made over the 2 years of school, the children who repeated kindergarten lagged increasingly behind their grade level peers who did not repeat kindergarten. The research concluded that kindergarten students who were retained learned less than kindergarten students who were promoted in both reading and mathematics.

Hong and Raudenbush (2005) reported similar results when they analyzed the effects of retention using the same database as Burkam et al. (2007). The only difference in the two

studies was the number of students analyzed. From their research, Hong and Raudenbush (2005) concluded that retention in kindergarten left most retainees even further behind, and therefore impeded their cognitive development over the repeated year. At-risk students who were promoted to the first grade seemed to have a better chance of growth acceleration.

In summary, both kindergarten studies reached similar conclusions. The data showed that kindergarten students who were retained learned less than kindergarten students who were promoted; therefore, retention did not help the kindergarten students catch up academically. These studies contradict the notion that the earlier a child is retained the better.

Research Opposing Retention in Grades First Through Sixth (1993-2005)

Grade retention has been making a comeback and has been of interest to students, parents, educators, and school districts. Hundreds of individual studies, several meta-analyses, and various research summaries have demonstrated that grade retention is harmful for most low-achieving students. Meisels and Liaw (1993) conducted a longitudinal study using same age comparisons in grades kindergarten through eight to analyze retention. The subjects consisted of Black, White, and Hispanic students from the National Education Longitudinal Study of 1988. The three groups consisted of early retainees grades kindergarten through third (n = 2,075), later retainees grades four through eight (n = 1,128), and promoted students (n = 13,420).

Regression analyses were performed separately for early versus later retained students and retained versus never retained students. The data indicated that when retained versus promoted students were compared the retained students had lower standardized test scores and academic grades than promoted students. However, when early versus later retained students were analyzed the study indicated that students who experienced early retention grades kindergarten through third displayed significantly higher academic performance than students

retained in grades four through eight. The data suggested that retention in grades kindergarten through three was probably more beneficial to the students than retention in later grades.

In a more recent study, the Chicago Public Schools (CPS) took a stand to end social promotion and instituted promotion requirements based on students' scores on the Iowa Tests of Basic Skills (ITBS) in the third, sixth, and eighth grades (Roderick & Nagaoka, 2005). As a result of this policy, Chicago retained from 7,000 to 10,000 students per year. Given the renewed interest in grade retention related to the emphasis on educational standards and accountability, Roderick and Nagaoka analyzed the effects of grade retention on third and sixth grade students.

The data used in this analysis came from administrative student records, provided by CPS and maintained by the Consortium on Chicago School Research. The students' ITBS test scores in reading were examined from 1998 through 2000 to make comparisons. Roderick and Nagaoka (2005) limited their study to students who failed to make the promotional cutoff in reading after attending summer school. Test score cutoffs were measured in grade equivalent metric (GE). The third grade below-cutoff group consisted of students whose highest reading test score was between 2.4 and 2.7. The third grade above-cutoff comparison group was defined as third graders whose highest reading test score by summer was between 2.8 and 3.1 slightly above the cutoff, the majority of whom were promoted. In the sixth grade, the below-cutoff group consisted of students whose highest reading test score was between 4.9 and 5.2, and the above-cutoff group consisted of sixth graders who had reading test scores between 5.3 and 5.6.

The participants consisted of students who were in the grade for the first time and students who were retained. Students who did not have postgate grade test scores were excluded. Postgate grade test scores are test scores that measure academic achievement after the

students have been in grades third, sixth, or eighth. Tests scores for more than 90% of the students placed in special education after the gate grade were also included. The gate grades consisted of third, sixth, and eighth grade where the students must have met minimum test score standards in reading and mathematics on the ITBS to be promoted to the next grade. Analysis was restricted to third and sixth grade because of the lack of comparability of test scores for retained and promoted eighth graders. Students in the eighth grade take the ITBS, whereas students in the ninth grade take the Test of Achievement Proficiency (TAP), and the two tests were not comparable.

To measure achievement growth, the ITBS test scores were converted to a logit metric using Rasch models that are comparable across time and across test levels and forms. The reason the ITBS scores were converted to logits was because the scores were typically reported in terms of GEs. GEs present a number of shortcomings for comparison over time and across grades that consist of: (1) different forms of the tests are administered each year and could vary in difficulty; (2) GEs are not a liner metric, so a score of 5.3 on level 12 of the test does not represent the same thing as a score of 5.3 on a level 13; and (3) GEs are not linear within test level, because the scale spreads out more to the extremes of the score distribution. To address some of the concerns, an alternative metric derived from an item-response model was used. This model assumes that the probability that student *i* correctly answers question *j* correctly is a function of the student's ability and the item's difficulty. In practice, one estimates a simple logit model in which the outcome is whether or not student *i* correctly answers question *j*. The difficulty of the question is given by the coefficient on the appropriate indicator variable, and the student's ability is measured by the coefficient on the student indicator variable. The resulting metric is calibrated in terms of logits. Logit metrics provide an effective way to compare

students on different grade levels or who are taking different forms of the test (Jacob & Lefgren, 2004). During the time of this analysis, equated Rasch scores were only available for test forms used in Chicago through 2001. Therefore, the effects of retention were only examined 1 year after the gate grade.

There were three different comparisons made to estimate the effects of retention. Comparison 1 used across-cohort comparisons of third graders using the change in the retention policy in 2000 comparing the 1998 and 1999 predominantly retained below-cutoff group with the predominantly promoted 2000 below-and above-cut-off groups. In 1998 and 1999, Chicago had a strict promotion policy based on test scores. In 2000, the policy was revised to where a range around the cutoff scores could be used to determine promotion.

In Comparison 1, when examining the short-term effects of retention, students in the 1998 and 1999 below-cutoff group, most of who were retained, had slightly higher estimated learning gains .04 between second grade and the Postgate 1 than the 2000 above-cutoff group. They also had higher achievement gains than students in the same cohort who had test scores just above the cutoff. The data suggested that the third graders in the predominantly retained below-cutoff group experienced a slight boost in performance in the postgate year. The learning gains in the 2000 above-cutoff group between second grade and 1 year after the gate grade were 1.09 as compared to 1.13 from the predominantly retained group, a difference of less than 4%.

Comparison 2 used a within-cohort comparison of the postgate achievement growth 1 and 2 years after promotion/retention of the 1998 and 1999 predominantly retained below-cutoff group with the predominantly promoted above-cutoff groups for third and sixth graders.

In Comparison 2, Roderick and Nagaoka (2005) found that third graders in the below-cutoff group had slightly larger learning gains .05 logits between second grade and l year after

the gate grade than their counterparts in the above-cutoff group who attended summer school and who narrowly passed the test score cutoff in reading. The data indicated that the small gains were short lived. The achievement growth of the below-cutoff group between second grade and 2 years after the promotional gate, which was .02 on Postgate 2, was not statistically different than the above-cutoff group.

The results for the sixth grade were more negative. The Postgate 1 adjusted achievement growth of the students in the above-cutoff group between fifth grade and 1 year after promotion or retention was 1.13 as compared to 1.06 for the above-cutoff group. This -.07 translated into a 6% difference in learning gains over the 2 years. The gap between the achievement growths of both groups continued 2 years after promotion or retention.

Roderick and Nagaoka (2005) concluded that neither of the first two comparisons explicitly addressed the fact that the experience of retained students varied during their retained year and the following year. In Comparisons 3, they modified the basic model to estimate the achievement growth of promoted and retained students by their experience of retention.

Comparison 3 used a within-cohort comparison of the postgate achievement growth 1 and 2 years after promotion/retention in 1998 and 1999. Third and sixth graders in the below- and above-cutoff groups were compared by whether students: (a) were promoted, (b) experienced a full year of retention and remained one grade below their age-appropriate counterparts, (c) experienced two retentions, (d) were placed in special education, or (e) were initially retained but later rejoined their age appropriate groups.

In Comparison 3 in third grade, the effects of a full year of retention were similar to that found in the first two comparisons. The Level 2 coefficient on Postgate 1 was .02. From second grade to Postgate 2, achievement growth of third graders who experienced a full year of retention

was slightly but not significantly greater than the low-achieving students who were promoted at the end of the summer. At Postgate 2, there was no statistically significant difference in the learning achievement growth of third graders who experienced a full year of retention and those who were promoted. The data indicated that students who experienced a full year of retention demonstrated no negative or positive effects on achievement 2 years after retention.

The estimated growth in achievement of sixth graders who experienced a full year of retention was .37 logits lower than promoted low-achieving students in the sample between fifth grade and 1 year after the gate grade and .44 logits below the promoted students after 2 years. This was equivalent to a 31% difference in achievement growth over 2 years and a 24% difference over 3 years between sixth graders who experienced a full year of retention and students with similar performance who were promoted.

The data indicated that retained students experienced a high rate of special education placement. Eleven percent of third graders and 13% of sixth graders in the total sample were placed in special education within 2 years of the gate grade. Data from the third grade sample suggested that the achievement growth of these students was significantly lower at both Postgate 1 and Postgate 2. In the sixth grade, students placed in special education had an achievement growth nearly 28% lower than students with similar scores who were promoted (1.31 logits vs. 1.81 logits). The data indicated that even after accounting for differences in prior achievement, retained students who were placed in special education experienced a decrease in their achievement growth as compared to other low-achieving students who were promoted.

The students in third grade who were promoted in January half way through the school year after passing the promotional cutoff, 21% of the sample, had small but statistically greater

achievement growth than promoted students both 1 and 2 years after the gate year. In sixth grade, the data indicated there was no difference.

In summary, the research conducted by Meisels and Liaw (1993) and Roderick and Nagaoka (2005) was consistent with the results reported in the seminal studies. Retention did not proffer any academic benefits to the students who were retained and in some cases retention was associated with negative growth in achievement. However, Meisels and Liaw indicated that retention in the early grades kindergarten through third was probably less harmful to students then retention in the later grades.

Research Supporting Retention in Grades First Through Sixth (1987-2007)

Several effects regarding retention are both good and bad. Retention can have emotional, social, financial, and academic impacts on all stakeholders. This section provides an overview of some studies that are contrary to the belief that retention is an ineffective strategy to remediate underachieving students. The studies indicate that retention may be favorable for students in grades first through third.

Using the same databank and test design as Roderick and Nagaoka (2005), Jacob and Lefgren (2004) analyzed the impact of grade retention and found positive effects on academic performance for the third grade students but not for the sixth grade students. The participants consisted of third and sixth grade students from 1997 to 1999 whose June math score was above the cutoff, whose June reading score was below the cutoff reading, and whose August reading score was between 1 grade equivalent below and 0.5 grade equivalent above the cutoff on the Iowa Test of Basic Skills (ITBS).

After analyzing the data, it was concluded that being retained in the third grade increased performance the following year by .17 logits in reading and .23 logits in math. The retained

students demonstrated an increase in achievement of 41% and 33% of the average annual gain. By the second year following retention, the math effect had decreased substantially, but was still significant. The 2 year reading effect was not statistically different than 0.

When six grade students were compared, Jacob and Lefgren (2004) found no significant difference between the performances of retained and promoted students. However, when comparing the students after 2 years, the retained students scored .15 logits (27% of an annual learning gain) lower than the promoted students. There was no significant difference between the math achievement of retained and promoted students.

Karweit (1999) analyzed the correlates and consequences of grade retention on first grade students' academic progress. The first grade cohort consisted of 10,333 students who were tracked from 1991 to 1994. The cohort data came from Prospects, a nationally representative longitudinal databank.

Karweit's (1999) analysis examined reading and mathematics achievement as assessed in the fall of first grade fall 1991, spring 1992, spring 1993, and spring 1994. The analysis reported both same age and same grade comparisons from several perspectives (retained versus all promoted; retained versus all promoted after adjusting statistically for background differences and achievement; and retained versus a low performing promoted comparison group).

The data indicated that the same age comparisons showed no benefit or harm as a result of retention for all three groups. The achievement gap separating repeaters and promoted children were about the same after retention. However, same grade comparisons consistently favored retention. When comparisons were made between the never retained group and the retained group, on average, the gap between never retained and retained children was reduced from 1.25 standard deviation units prior to retention to .38 standard deviation units after

retention. The positive effects of retention on academic achievement were found in the other two same grade comparison groups as well.

Gleason et al. (2007) reached similar conclusions when they analyzed the short term effects of grade retention on academic performance of first grade students. The study's participants consisted of three hundred fifth (52.6% male) first grade students attending one of the three school districts (one urban, two suburbs) in central and southeast Texas. The ethnic composition consisted of 74 African American, 132 Hispanic, 130 White, 12 Asian or Pacific Islander, and 2 other. The students were recruited across two sequential cohorts in first grade during the fall of 2001 and 2002. To be eligible to participate, the students must have scored below the median score on a state approved district administered measure of literacy.

From November to March of year 1, when the students were in first grade, tests in reading and math achievement were administered. The tests were readministered 1 year later. The WJ-III Broad Reading scores (Letter-Word Identification, Reading Fluency, and Passage Comprehension subtests) and the WJ-III Broad Math scores (Calculation Skills subtests) were used to measure achievement. For students or their parents who spoke Spanish, the Bateria Woodcock-Murioz was used. The WJ-III Tests of Achievement are individually administered measures of academic achievement for individuals ages 2 to adulthood (Gleason et al., 2007).

Gleason et al. (2007) found a significant difference between retained and promoted students after the year of retention. Same age comparisons indicated that retained students scored significantly lower than the promoted students on both the Broad Reading and Math. The same grade comparisons showed that retained students scored significantly higher than the promoted students on both the Broad Reading and Math.

Gleason et al. (2007) concluded that when same grade comparisons were made the students succeeded academically and socially. The researchers suggested that the academic gains may have been the result of the peer support retained students received during the repeated year which increased their sense of school relatedness and academic motivation, resulting in more positive academic trajectories.

In the Beginning School Study (BSS), conducted by Alexander et al. (2003) positive gains were also found in student performance after grade retention. The BSS began in the fall of 1982, when the participants were just beginning first grade. The study analyzed academic performance and personal development of retainees and nonretainees before and after retention through the eighth grade.

A stratified random sample of 775 students was selected from the pool of children entering first grade in Baltimore City Public Schools (BCPS) in 1982. The 20 schools were chosen on a random basis from strata defined by the school's integration status and by community socioeconomic level.

The participants consisted of 55% African American, 45% White, two-thirds qualified for subsidized meals at school during the elementary years, 40% of their mothers lacked high school degrees, and 44% of the children were in solo-parent households as first graders. The first grade cohort consisted of a large, diverse, representative, nonvolunteer sample of typical children attending public schools in a high-poverty city (Alexander et al., 2003).

The strengths of the study consisted of: (1) the data and design provided a strong foundation for trying to understand the impact of retention; (2) data were longitudinal, so the students' histories were known before retention occurred; (3) the study began at the start of first grade, before anyone had been held back; (4) coverage of the participant's experience extended

several grades beyond the repeated year; and (5) an opportunity to compare the first graders against students held back in second grade or later, allowed the researchers to see whether first grade retention differed from retention in later years.

Some weaknesses of the study included: (1) detailed classroom observations were not conducted; (2) information from all sources was not available each year (e.g. interviewing the parents); (3) there was a gap in coverage where students transferred out of the BCPS each year; (4) some students had sketchy school record data on grade level progression; (5) some of the parents' economic standing was not as detailed as they would have liked; and 6. little was known about special services provided for students before, during, or after their retention year.

After the first year, 126 students were retained in first grade and two were assigned to special education classes. Six hundred twenty-six went to the second grade. The total retention rate for the first year was just over 16%. This was the highest grade-specific rate of first retention for any year; correspondingly, it yielded the largest single group of retainees.

Sixty-one students were identified as second grade repeaters in Year 3 and seven repeaters were assigned to special education classes at the end of the year. The 61 repeaters in second grade were joined in second grade by 103 first grade repeaters, promoted after 2 years in first grade. Other first grade repeaters 13.4% of the original 127 had been moved into special education classes. The connection between retention and special education was established early.

In the third year, the data indicated that two students that had been second graders in Year 2 skipped third grade and were identified as fourth graders in Year 3. In Year 3, most of the students were in second (164) or third (518) grade. The 518 students who remained on

schedule through the 3 years constituted 72% of the 715 children still in city schools at that point. Sixty children had left the city system by the fall of Year 3.

In the fourth year, only 415 children were fourth graders. Sixty-four percent of the 649 students were still in the BCPS through Year 4. Sixty-six more left the city schools between Years 3 and 4. Forty-five third grade repeaters not in special education were joined by 143 first and second grade retainees, who in the fourth year had made it to the third grade. Three students were still in the second grade and classified as double repeaters, and 13 more moved into special education classes.

By Year 5, the number of students still in BCPS ($N = 614$) was under 80% of the original cohort. Three hundred sixty-five were in fifth grade 60% total; 171 were in fourth grade, including 20 fourth grade repeaters. There were 22 double repeaters and 52 students placed in special education classes up 12 from the previous year. Two special education students were placed back in the regular classroom and one left the school system.

Alexander et al. (2003) found that during Year 6, 7, and 8, middle school years in most Baltimore schools, many students were still finishing elementary school. There were nine fifth grade repeaters in Year 6, 18 sixth grade repeaters in Year 7, and 22 seventh grade repeaters in Year 8.

There was a jump in the number of double retentions during the middle school period especially among sixth graders in Year 7. There were 151 children that year that had already repeated one grade. In the fall of Year 8, 25 of them were still in the sixth grade. According to Alexander et al. (2003), the jump in retentions toward the end of middle school probably reflected teachers' reluctance to passing students along to the next level that they deemed unready and the teachers were trying to shield struggling students from the pressures of high

school. However, many students who were still in the middle school reached the legal dropout age.

During the middle school years, a significant number of special education students were put back in regular education. Three special education students were put back in regular education classes at the end of Year 6, 15 in the spring of Year 7, and five in the spring of Year 8. Four were placed in eighth grade and were back on schedule. Four were placed in sixth grade, equivalent to repeating two grades; the remaining seven were behind l year. The reassignment of students probably reflected a desire to get students back into the regular classes before high school.

Just over half 238 of the 470 students who were still in city schools at this point were at grade level. Thirty-six percent or 169 were behind one or more grades, and 13% or 61 were in special education.

The data analyzed were collected on children's school performance over an 8 year period. Because the retainees spent 2 years in the repeated grade, the data on performance were plotted through seventh grade for practically everyone except the double repeaters. The number of retainees was largest in the first 3 years. There were 127 first-time repeaters in the first grade, 68 in the second, and 47 in the third. Because coverage of performance trends after retention was longest in these grades, profiles were presented separately for first, second, and third grade retainees. For fourth through seventh grade repeaters, the time line was abbreviated and sample sizes year by year were too small to analyze so these students were grouped together as late repeaters.

To determine student achievement, the California Achievement Test (CAT) scores were used. The CAT tests consisted of CAT Reading Comprehension (CAT-R) and the CAT Math

Applications and Concepts (CAT-M). The CAT tests were designed to measure a child's current educational achievement level. Repeaters' CAT performance was compared against that of all never-retained students and separately, the never-retained subset consisted of the poor-performing comparison group.

After examining the performance data of first grade students after the repeated year with same grade scores, the retained students pulled up their CAT-R scores from 72 points behind to just 18 points behind, on the CAT-M they moved up from 56 points behind to 16 points behind. When the comparisons were adjusted for risk factors that affect test performance apart from retention, they become more favorable. The retainees scored about 17 points above promoted children on the CAT-M. Retainees' scores also surpassed the poor-performing comparison groups in every instance, reaching almost 30 points on both CAT subtests. The data indicated that when first grade repeaters moved to second grade, they were performing at a level much closer to that of their classmates. However, the first grade repeaters began slipping as soon as they got beyond their repeated year and after just a few years they were far behind again.

Second grade repeaters, however, were found to be on par with CAT-R and above in CAT-M at the end of their repeated year. When changes did occur between the retainees and promoted students during the years up to seventh grade, the differences were too small to be significant. The second grade retainees' relative standing generally improved after retention.

For third grade retainees, trends were even more favorable. After adjusting for CAT scores and background factors, the third grade repeaters were significantly ahead of their classmates in the repeated year. Also, when retainees moved beyond third grade, their adjusted CAT averages never fail significantly below those of promoted students. The data suggested that

retention helped the third grade repeaters and they did better in fourth and fifth grades because they were retained.

The data indicated that retention helped students who were not to far behind the most. It allowed them more time to mobilize their resources and to master some of the skills they did not acquire the first time through the grade. The study showed that retention for the low achieving students in grades first through third was helpful and for many their level of performance stayed above the level projected for them on the basis of their earlier performance trajectories. Fourth through seventh grade repeaters were not discussed in detail but appeared to have less difficulty all along the way. However, the study indicated that students retained in grades four through seven were more likely to drop out of school than those retained in the early grades (Alexander et al., 2003).

Lorence and Dworkin (2006) reached similar conclusions after analyzing the effects of retention for third grade students in Texas from 1994 – 2002. Third grade students who failed the Texas Assessment of Academic Skills (TAAS) in reading in May 1994 and students who were socially promoted were the participants in the study. Only Non-Hispanic White, Hispanic, and African American students were analyzed. After exclusions were made for various reasons (e.g. missing data, retained previously, sever learning problems), 38,445 students failed the 1994 reading test. Three percent (n = 1,244) of the low performing students were required to repeat third grade while 97% were placed in the fourth grade.

Lorence and Dworkin (2006) utilized a nonequivalent control group design to analyze the data. Unlike the typical nonequivalent control group design, outcome means for eight grades were examined. Statistical controls were used in an attempt to equalize potential differences in the promoted and nonpromoted students.

After analyzing the mean scores, Lorence and Dworkin (2006) found that for the Non-Hispanic White students who repeated third grade the average reading score was 76.6 significantly higher than the third grade reading score of the promoted students that was 61.1. In fourth grade, the average reading score for the retained Non-Hispanic White students was 68.1 significantly higher than the fourth grade reading score of the socially promoted students that was 62.5. With the exception of grades eight and ten, the average reading scores for the retained third graders were significantly larger than those of the promoted students. Also, the Non-Hispanic White students who repeated third grade in 1994-1995, on average, began passing the TAAS reading test after being held back; socially promoted third graders, however, did not exceed the required level of proficiency until sixth grade.

A similar pattern occurred for both the Hispanic and African American students. Third grade repeaters markedly improved their TAAS reading scores by the end of their retention year. It was also observed that Hispanic and African American retainees in general correctly answered more questions than the socially promoted students and scored significantly higher on the tenth grade reading test than did the nonretainees.

The results after adjusting for differences between the retained and socially promoted students by taking into consideration initial test performance in 1994, gender, whether the student was enrolled for free or reduced price lunch, limited English proficiency (LEP), and special education status were consistent with the examination of the mean reading scores for the retained Non-Hispanic White students. The retained students had somewhat higher averages than the socially promoted students in all but grades 8 and 10. When Lorence and Dworkin (2006) calculated the effect size for the retained students after repeating third grade, the students scored almost two standard deviations higher than the promoted students. Also, the mean

adjusted reading score of the third grade repeaters exceeded the value of 75% needed to pass the reading test. In fourth grade while the mean scores were lower, the retained students obtained fourth grade reading scores that were .6 of a standard deviation higher than their socially promoted classmates. An effect size of .55 in fifth grade indicated superior reading performance among the retained students. According to Lorence and Dworkin, although the differences between the retained and socially promoted third grade students were not as pronounced in eighth and tenth grades, the retained students continued to correctly answer more reading items than the nonretainees.

The positive effects of grade retention were replicated within each of the minority groups. The findings revealed that making low-performing minority students repeat a grade was associated with higher reading scores. Retention allowed failing students the opportunity to learn material they missed. Not only did the retained students catch up with the socially promoted students, they showed evidence of higher levels of reading ability in the following grades. It was only in eighth grade that the African American repeaters did not obtain significantly higher reading scores. However, the average tenth grade reading score of the retained minority student was about four-tenths of a standard deviation greater than that of the nonretained students, indicating that the effects of retention persisted during the sophomore year of high school.

Lorence and Dworkin (2006) concluded:

> There is no evidence in the data that making academically challenged children repeat a grade harms their academic progress. Indeed, retention seemed to boost the ability of the initially low-performing minority readers over that of the socially promoted African American and Hispanic students who failed the state reading test. (p. 1,027)

Greene and Winters (2004) also examined grade retention for third grade students and found positive benefits. In Florida, the state legislature passed a law that third grade students had to score a Level 2 benchmark or above on the reading portion of the state's high-stakes test, the Florida Comprehensive Assessment Test (FCAT), in order to be promoted to the fourth grade. Students who failed to reach the benchmark were retained in third grade. This policy mandate began in the 2002-03 school year.

The law allowed for some exceptions to the retention policy and students could be promoted if they met one of the following criteria: (1) were a Limited English Proficiency student who had received less than two years of instruction in an English for Speakers of Other Languages program; (2) had a disability sufficiently severe that it was deemed inappropriate for the student to take the test; (3) demonstrated proficiency on another standardized test; (4) demonstrated proficiency through a performance portfolio; (5) had a disability and had received remediation for more than two years; or (6) had already been held back for two years (Greene & Winters, 2004).

The data consisted of the individual test scores of all third grade students who failed to reach the minimum benchmark on the FCAT reading test during the 2001-02 and 2002-03 school years. The first cohort of students consisted of the students who entered the third grade for the first time in 2002-03 and scored below the Level 2 threshold on the FCAT reading test in that year. Sixty percent were retained. Also, included in the study were the students who entered third grade for the first time in the 2001-02 school year who scored below Level 2 on the FCAT reading test who would have been retained under the new policy. Nine percent were retained. Greene and Winters (2004) pointed out that "both groups were very similar in all respects except

for the year in which they were born, making comparisons between their improvements particularly meaningful" (p. 6).

The one year test score gains that students made on the math and reading tests were analyzed. Each test used developmental scale scores to measure student achievement. Using developmental scale scores allowed Greene and Winters (2004) to compare the test score gains of all the students in the study even though they took different tests designed for different grade levels. Developmental scale scores are designed to measure academic proficiency on a single scale for students of any grade and in any year. This means that a third grader with a development scale score of 1,000 and a fourth grader with a developmental scale score of 1,000 have the same level of academic achievement: if a student gets a developmental scale score of 1,000 in 2001-02 and gets the same developmental scale score of 1,000 in 2002-03, the student has not made any academic progress in the intervening year. When interpreting the results from the data, it is important to understand that while the developmental scale scores are consistent on a test between grades and years, they are not consist between subjects (reading and math) or between two different standardized tests (the FCAT and the Stanford 9). Although there are differences, the differences have no effect on the analysis.

In this study, Greene and Winters (2004) analyzed the achievement made by students over one year in math and reading scores on the criterion referenced tests the FCAT, as well as the norm referenced version which was the Stanford 9. Both tests were given to all Florida students in grades three through ten each year in reading and math. The FCAT was the test that third grade students had to pass to be promoted to the next grade. The Stanford 9 was administered to help parents better understand their children's proficiency levels and to check the reliability of the results of the FCAT tests. According to Greene and Winters (2004), "if similar

results were found on both the FCAT and the Stanford 9, this would indicate that real learning had occurred" (p. 6).

Greene and Winters (2004) calculated the developmental scale score gains on the FCAT and Stanford 9 made in each student's first third grade year and the following year. For the students retained, the test score gains they made were measured between the 2002-03 and 2003-04 administrations of the tests. For the students who were not affected by the retention policy, their test score gains were measured between the 2001-02 and 2002-03 administrations of the tests.

For the first analysis, Greene and Winters (2004) wanted to find out the effect of Florida's retention policy. The state's policy was intended to be a treatment for every third grade student who scored below the necessary benchmark on the FCAT. To measure the effect of the program, a linear regression comparing the developmental scale score gains made by the treatment group, students who first entered third grade in 2002-03 and scored below the FCAT benchmark in that year, were compared with the control group, students who first entered third grade in 2001-02 and scored below the FCAT benchmark in that year. Controls were used for the student's race, free or reduced lunch status, and whether the student was identified as Limited English Proficient. Also, each student's test scores during their first third grade year were used as the baseline for test performance.

The results found after analyzing the effectiveness of the policy concluded that students who were subjected to the retention policy made gains of .06 standard deviation units in reading and between .14 and .15 standard deviation units in math relative to the students not subjected to the retention policy. The benefits translated into about two percentile points on reading and five percentile points on math over a one year period for the average student. Students who were

actually retained made gains of .11 to .13 standard deviation units on reading and .28 to .30 standard deviation units on math as compared to the students who were promoted. The benefits translated into about three or four percentile points in reading and about nine or ten percentile points in math over a one year period for the average student.

The results of the effects of the retention policy without accounting for whether the students were actually retained when examining the FCAT and Stanford 9 reading tests indicated that on the FCAT, students affected by the retention policy made reading test score improvements that were 16.66 developmental scale points greater than those students not affected by the policy. This translated into a gain of about .06 standard deviation units for the treatment group on the FCAT after one year. The results on the Stanford 9 reading test were similar. Students subjected to the retention policy made test score improvements that were 1.44 developmental scale points greater than those of the control group, which translated into a difference of about .06 standard deviation units. Both results were statistically significant at a very high level (p - values < .001).

The findings for the first analysis in math on the FCAT and Stanford 9 were even greater. On the FCAT in math, students subjected to the policy made improvements that were 41.67 developmental scale points greater than those in the control group, translating to an increase of about .15 standard deviation units. Similar results were found on the Stanford 9 math test. Students who were subjected to the promotion policy outperformed the control group by an average of 4.50 developmental scale points, translating to about .14 standard deviation units. Both results were statistically significant at a very high level (p - values < .001).

For the second analysis, the purpose was to evaluate the effects of actually retaining low performing students. For this analysis, a comparison was made between low scoring students

from either year who were actually retained with low scoring students from either year who were promoted. A two stage least squares regression analysis was conducted. The variable of interest for this analysis was whether a student was retained or promoted. In this model, student demographics and the year in which a student was born were used to predict whether students were retained. The prediction was then used to measure the relationship between retention and test score improvements. Controls were used for student race, free or reduced lunch status, Limited English Proficiency status, and baseline test scores.

The results for the second analysis on the effects of actually retaining students indicated that on the FCAT reading test students who were retained made improvements that were 32.48 developmental scale points or .13 standard deviation units greater than those students who were promoted. On the Stanford 9 reading test, students who were retained demonstrated an increase in achievement by 2.80 developmental scale score points as compared to the students who were promoted which translated to a gain of about .11 standard deviation units. Both findings were statistically significant at a very high level (p - values < .001).

On the FCAT math test, the data suggested that the retained students improved by 82.54 developmental scale points, an improvement of about .30 standard deviation units. On the Stanford 9 math test, the retained students improved by an average of 8.77 developmental scale points which translated to a difference of about .28 standard deviation units. Again, both results were statistically significant at a very high level (p - values < .001).

Greene and Winters (2004) extended the study by performing both analyses on all racial subgroups in the student population. They concluded that in both cases, very similar results were found for students in each racial group. The retention policy and actual retention of students had positive effects of about the same magnitude on students of all races.

In 2006, Greene and Winters further examined Florida's retention policy in a study called Getting Further Ahead by Staying Behind: A Second-Year Evaluation of Florida's Policy to End Social Promotion. While a copy of the study was obtained, the study could not be quoted or cited without the authors' permission. An attempt was made to contact the authors to use the study but no response was received. So, this study was not included in this literature review.

However, in 2007 Greene and Winters conducted a follow up study examining Florida's promotion policy. In this study, Greene and Winters (2007) analyzed the impact of grade retention on student performance in reading one and two years after the retention decision.

As mentioned in the previous study, third grade students had to meet at least the Level 2 benchmark in order to be promoted to the fourth grade. Students who scored a Level 2 were considered to have limited success with the challenging content on the test. The entering third grade class of 2002-03 was the first to be subjected to the mandate. At the end of the 2002-03 school year, 59% of the students were retained in third grade.

The data used consisted of test scores and demographic characteristics for the students enrolled in grades three through ten in a Florida public school from 2001-02 to 2004-05. As in the previous study, Greene and Winters (2007) analyzed the student level test scores on the Florida Comprehensive Assessment Test (FCAT) in reading that were reported as developmental scale scores.

In the first analysis, Greene and Winters (2007) use the across-year approach. The focus was only on the students in the third grade in 2001-02 or 2002-03 whose test scores were below the Level 2 benchmark on the FCAT reading test. A comparison was made between the academic achievement of students with low test scores who were in the first third grade class (subjected to the retention policy), referred to as the treatment group, with the test score gains of

students with the same low baseline scores but who entered the third grade in the year prior to the policy (who were not subjected to the policy) referred to as the control group.

The results of the across-year comparison on the test score gains made in reading indicated that retained students outperformed promoted students in both years. In 2002-03, the baseline third grade year for the treatment group, the mean developmental scale score on the FCAT reading test for all students was 1290.9 with a standard deviation of 381.2. The data showed that after one year the retained students outperformed promoted students by about 0.05 standard deviations. The reading benefits of retention after two years were 0.40 standard deviations.

Greene and Winters (2007) further analyzed the effects of retention using a regression discontinuity design. The test score gains of students whose reading score in 2002-03 was just below the threshold required for promotion were compared to students who were in the third grade that same year and whose scores were just above this threshold. Unlike the across-year approach, all students in the design were in the third grade in 2002-03 and were vulnerable to the retention policy if they did not score above the necessary threshold.

The results of the regression discontinuity comparison indicated that after one year, retained students made reading gains on the FCAT that were not statistically different from those made by the promoted students. However, the relative gains the retained students made in the second year grew to 176.90 developmental scale score points. In standard deviation terms, the data indicated that after two years, students who were retained outperformed promoted students by about 0.46 standard deviations in reading.

The results of both the across-year and regression discontinuity approaches indicated that the third grade students who were retained made significant and economically substantial gains

in reading relative to the promoted students. The fact that the retained third grade students' scores grew after two years is consistent with the idea that retained students will continue to gain ground in reading relative to promoted students in later years as academic material becomes more difficult (Greene & Winters, 2007).

The last two studies summarized are later studies. These studies are included because they provide data that examines retention over time. Some studies have indicated that if student performance is sustained over time, then retention was probably an effective remediation practice (Alexander et al., 2003; Greene & Winters, 2007; Jacob & Lefgren, 2004; Lorence & Dworkin, 2006).

In a 3-year study, Peterson et al. (1987) analyzed the effects of retention in grades first, second, and third for students in the Mesa Public schools. The research question in their study was whether retained children do better in reading, language, and math after 4 years of schooling than do children not retained. To create a basis of comparison, matched samples of promoted students were identified at each grade level.

Same age comparisons indicated that students made significant improvements in reading and math achievement in first and second grade during the year of retention but the benefits diminished significantly 2 to 3 years later. Findings of same grade comparisons indicated that first grade retainees had better performance in reading, math, and language at the end of the retention year and in reading and language 1 year later. Students retained in second and third grade significantly outperformed promoted students in reading, language, and math. The academic advantages gained by the second and third grade retainees were maintained over 2 years. The evidence suggested that retention in the early years of elementary school was not harmful, and in comparison with social promotion, was beneficial.

Pierson and Connell (1992) reported similar findings. They compared achievement scores and attitudes across four groups: repeaters (n = 74) retained in first through fourth grade who had passed at least a year since they completed their repeated grade, a matched-ability group (n = 69), a random sampling of students (n = 60), and a group of students who were socially promoted (n = 35). These subjects were chosen from grades three through six in two upstate New York school districts, one urban, one rural. The social promoted group consisted of eight students who were identified in school records as having an unearned promotion or were recommended for retention by the teacher but were promoted anyway.

Pierson and Connell (1992) concluded:

> It appears that whereas retention is not a cure-all for below grade-level academic performance, students whose academic performance suggests that they should be retained, and who are retained, perform better 2 or more years later than students with comparable performance who are promoted. Therefore, the findings support the use of retention as a potentially effective remediation for academic difficulty in the early elementary grades. (p. 305)

To conclude, the research on grade retention and its impact on student achievement indicated mixed results. However, there was evidence that suggested that retention for students in grades first through third was probably more favorable than for students in grades fourth through sixth. The data showed that retention significantly reduced the size of the gap between retained and promoted students that existed prior to retention. Retained students seemed to have mastered some of the skills they didn't acquire the first time through the grade. When the achievement of retained students was examined after several years, the data indicated that retained students continued to gain ground relative to the promoted students.

Summary

The increased emphasis on educational standards and accountability has rekindled the use of grade retention as an intervention to remedy academic deficits. The benefits and disadvantages of holding children back a year in school have been debated for years (Darling-Hammond, 1998; Position statement on student grade retention and social promotion, 2003). Despite the wealth of studies addressing this topic, little consensus has emerged on the effectiveness of grade retention as a practice.

The research on grade retention indicates mixed results. The strong beliefs about the ineffectiveness of grade retention are primarily derived from four seminal studies conducted by Holmes (1989), Holmes and Matthews (1984), Jackson (1975), and Jimerson (2001). These studies conclude that requiring low-performing students to repeat a grade is a futile educational practice (Lorence & Dworkin, 2006). However, other research studies indicate that retention is beneficial; helping students learn the material they missed the first time through the grade (Alexander et al., 2003; Gleason et al., 2007; Greene & Winters, 2004, 2007; Jacob & Lefgren, 2004; Karweit, 1999; Lorence & Dworkin, 2006; Peterson et al., 1987; Pierson & Connell, 1992). There is evidence in the research to suggest that retention of students in the early grades first through third is probably more favorable than for students in the later grades fourth through sixth (Alexander et al., 2003; Gleason et al., 2007; Greene & Winters, 2004, 2007; Jacob & Lefgren, 2004; Karweit, 1999; Lorence & Dworkin, 2006; Peterson et al., 1987; Pierson & Connell, 1992).

CHAPTER 3: METHODOLOGY

The purpose of this chapter is to explain the design and methodology that will be used in this study. The chapter includes the research questions, research design, setting and participants, instrumentations, data collection procedures, data analysis, and summary.

Research Questions

The research questions in this study are designed to compare the academic achievement in the areas of reading and mathematics of sixth grade students who have been retained in early grades first through third to those who have been retained in later grades fourth through sixth. The data collected will answer the following questions:

1. Is there a relationship in the growth performance as measured by the ABC's academic change score categories of within and/or above expectations versus below expectations between sixth grade students retained in grades first through third as compared to sixth grade students retained in grades fourth through sixth in reading in the Onslow County School System?
2. Is there a relationship in the growth performance as measured by the ABC's academic change score categories of within and/or above expectations versus below expectations between sixth grade students retained in grades first through third as compared to sixth grade students retained in grades fourth through sixth in mathematics in the Onslow County School System?
3. Is there a relationship in the growth performance as measured by the EVAAS predicted score categories of met and/or above versus below expectations between sixth grade students retained in grades first through third as compared to sixth grade

students retained in grades fourth through sixth in reading in the Onslow County School System?

4. Is there a relationship in the growth performance as measured by the EVAAS predicted score categories of met and/or above versus below expectations between sixth grade students retained in grades first through third as compared to sixth grade students retained in grades fourth through sixth in mathematics in the Onslow County School System?

Research Design

National, state, and local legislation and policies mandate an increase in academic standards and prohibit social promotion. Grade retention is used as a remediation practice to increase student achievement. Existing research examining the effects of grade retention on academic achievement indicates mixed results. The research articles in Chapter 2 provide both evidence that supports and opposes retention. Kindergarten retention was not included in this study because there is no way to monitor if students began kindergarten when age-eligible. Research indicates that some parents choose to delay kindergarten entry to allow their child an extra year to mature or because of developmental difficulties (Holloway, 2003). The research studies opposing retention suggest that retention does not demonstrate academic advantages for retained students at any grade level. This conclusion is cited in the three meta-analyses conducted by Holmes, (1989), Holmes and Matthews (1984), and Jimerson (2001), and also in data reported by Meisels and Liaw (1993), and Roderick and Nagaoka (2005).

The research studies in Chapter 2 supporting retention indicate that students retained in first, second, or third grade show positive gains in achievement as compared to students in grades four through six who show less gain in achievement (Alexander et al., 2003; Gleason et

al., 2007; Greene & Winters, 2004, 2007; Jacob & Lefgren, 2004; Lorence & Dworkin, 2006). When reading and math test score gains were analyzed, the data demonstrated that students retained in grades first through third significantly outperformed promoted students (Alexander et al., 2003; Gleason et al., 2007; Greene & Winters, 2004, 2007; Jacob & Lefgren, 2004; Lorence & Dworkin, 2006; Peterson et al., 1987; Pierson & Connell, 1992). Based on these studies retention in grades first through third might be an effective early intervention strategy to use to help underachieving students as opposed to retention in grades fourth through sixth.

Development of End-Of-Grade Tests

The NC Department of Public Instruction developed the End-of- Grade Reading and Mathematics tests as an achievement test to measure acquisition of specific subject area content and skills associated with particular grade levels in North Carolina Public Schools. The purpose of the tests is to find out how well students are performing and to hold each school and the school's personnel accountable for the education of students (NCDPI, 2006b). The North Carolina End-of-Grade reading and mathematics tests are multiple-choice tests for students in grades 3-8. The tests are specifically aligned to the North Carolina Standard Course of Study and include a variety of strategies to measure the achievement of North Carolina students (NCDPI, 2004a).

The NC End-of-Grade test of Reading Comprehension assesses a student's ability to read and comprehend written material that is appropriate for a particular grade level in terms of difficulty and content (Garland, 2009). The test assesses a student's ability to apply strategies such as using context clues to determine meaning, summarizing to include main points, and identifying the purpose of text features. Students are required to clarify, to explain the significance of, to extend, and/or to adapt ideas and concepts. Questions on the test are typically

at a broader level than just facts and concepts, for they also test major ideas that students are expected to know if they are to be considered literate (NCDPI, 2004a).

Student growth on the NCEOG reading test is measured by developmental scale scores. The number of questions a child answers correctly on the test is called a raw score. The raw score is converted to a developmental scale score. Under the State Accountability Model (NCDPI, 2006b):

> Student scores in reading from the end-of-grade test are used for computing school growth and performance composites as required by the state mandated ABCs Accountability Program and for determining adequate yearly progress (AYP) under Title I mandates of the No Child Left Behind Act of 2001. Student scores are also used in determining student progress and proficiency at grades 3, 5, and 8.

The developmental scale score measures growth in reading comprehension from year to year; allowing for the comparison of the student's EOG reading score from one grade to the next. On average, student scale scores are expected to go up every year.

The NC End-of-Grade Mathematics test assesses achievement in the five strands of the mathematics curriculum: (1) Number and Operations, (2) Measurement, (3) Geometry, (4) Data Analysis and Probability, and (5) Algebra. There are two parts of the 82-item math test calculator active and calculator inactive.

Like reading, student growth in mathematics achievement is measured by developmental scale scores. In North Carolina, student scores in mathematics from the end-of-grade test are also used for computing school growth and performance composites and for determining (AYP) under the accountability model. Student scores are also used in determining student progress and proficiency (NCDPI, 2006b).

Reliability of the NCEOG Reading Comprehension and Mathematics Tests

According to Garland (2009), tests developed by the North Carolina Department of Public Instruction's Test Development Section, when properly administered and interpreted, provide reliable and valid information. The ABCs of Public Education is North Carolina's statewide testing and accountability program and is considered to meet the reliability standard.

The North Carolina testing program develops its tests in six phases consisting of twenty-two steps over approximately four years. In the beginning stages, North Carolina educators are recruited and trained to write test item questions to be field tested. Once the test items are developed a new group of NC educators, test development staff members, and curriculum specialist review the test items. A field test is conducted using a randomly selected sample of students at each grade level. The data from the field tests are analyzed by the NCDPI psychometric staff and operational tests are constructed. The final item pool is based on approval by the (1) NCDPI Division of Instructional Services for curriculum purposes; and (2) NCDPI Division of Accountability Services/NC Testing Program from psychometrically sound item performance. Pilot tests are then conducted and the data analyzed. Standards are established and the tests are fully operational the following school year.

Validity of the NCEOG Reading Comprehension and Mathematics Tests

According to Garland (2009), "evidence of validity for the NCEOG Reading Comprehension and Mathematics tests are provided through content relevance, response processes, relationship of test scores to other external variables, and maintaining consistency in the testing environment" (p. 59). The end-of-grade assessments are given during the last three weeks of the school year. All students at the same grade level within a school are administered the appropriate end-of-grade test at the same time on the same day. The multiple choice test

questions in reading and mathematics are developed to measure the goals and objectives as specified in the NCSCS with particular focus on assessing students' ability to process information and engage in higher order thinking. Performance standards, called achievement levels, have been developed to identify levels of student performance. The achievement levels identify what is expected at various levels of performance. A description of each achievement level follows:

Level I: Students performing at this level do not have a sufficient mastery of knowledge and skills in the subject area to be successful at the next grade level.

Level II: Students performing at this level demonstrate inconsistent mastery of knowledge and skills in the subject area and are minimally prepared to be successful at the next grade level.

Level III: Students performing at this level consistently demonstrate mastery of the grade level subject matter and skills and are well prepared for the next grade level.

Level IV: Students performing at this level consistently perform in a superior manner clearly beyond that required to be proficient at grade level work.

Setting

The setting for this study will be the Onslow County School System located in Jacksonville, North Carolina. The school system is divided into 6 districts and is the eleventh largest school system in North Carolina serving 23,500 students at 36 learning sites which include 19 elementary schools, 8 middle schools, 7 high schools, an alternative learning center, and an early childhood center. At the time of this study, approximately 16,550 of these children were enrolled in grades kindergarten through eighth grade. Of the district's 23,500 students, 6.6% are Hispanic, 23.2% African American, 67.7% White, and 2.5% Asian/American Indian or

other. The school system serves 1% English Language Learners. Twelve percent of the students receive special education services. Approximately, 40% of the student population or 8,900 students receive free or reduced lunch. The percentage of economically disadvantaged students varies across schools ranging from 22.34% to 70.1%. It is estimated that one-third of the students move into or out of the school system or between schools during the school year. The transient student population is largely due to the large military community served by the system. Approximately 8,000 students enrolled in the school system have military connections.

Participants

The participants for this study will include students who attended Onslow County Schools in Jacksonville, NC in the sixth grade during the 2008-2009 school year that have been retained only one time in grades first through sixth. Students who have been retained in other grades will be deleted from the analyses. The number of students identified as being retained in grades first through third consists of (n=129). The number of students identified as being retained in grades fourth through sixth consists of (n=107).

The data will be managed by using archival End-of-Grade test data in reading and mathematics for sixth grade students that had been previously retained for the 2008-2009 school year obtained through the district. Also, EVAAS predicted scores for each student will be obtained from the district office. EVAAS is a customized software system provided to all NC LEAs by the Department of Public Instruction. EVAAS contains each student's historical test data that measures how much gain or growth an individual student or groups of student make over time.

Instruments

The instruments that will be used for data collection in this research study are the 2008-2009 NCEOG Reading and Mathematics tests. The EOG tests are designed and validated by the North Carolina Department of Instruction and are mandated for all students in grades three through eight. Standardized test scores are the logical choice for this research, as it is a study into relationships between grade retention and student achievement and academic change. In current studies examining the effects of retention, standardized tests are used to measure student achievement (Alexander et al., 2003; Gleason et al., 2007; Greene & Winters, 2004, 2007; Jacob & Lefgren, 2004; Lorence & Dworkin, 2006).

Measuring Student Growth

Since the implementation of the ABCs of Public Education in the 1996-97 school year, the formula for determining growth has changed. The new growth formula which began in 2006-2007 has shifted directions as a student's growth is based on previous as well as current student performance. The formula currently uses a standardized scale score, referred to as a c-scale (change scale), to measure relative student performance instead of the original developmental scale score (DSS). The c-scale score is similar to a z score in that it standardizes scores showing how far and in what direction the student's score is different from the expected score. A student's DSS is converted to a c-scale score. The DSS allows for the comparison of students across the state of North Carolina to each other and to their own expected growth by subject from one grade to the next. "A student's DSS is standardized and a student's performance is considered as a point on the c-scale (or change scale) relative to standard performance for that grade level in a standard setting year" (NCDPI Accountability Services, 2009).

Under this model, growth is based on academic change. As stated by the North Carolina Department of Education Accountability Services (2009):

> The academic change is expressed as the difference between a student's actual c-scale score for the current year and the student's average of two previous assessments with a correction for regression toward the mean. A positive academic change indicates a gain in academic achievement, while a negative academic change indicates a loss of academic achievement from the previous two years.
>
> The formula to determine academic change is:
>
> AC = CS – (0.92 X ATPA), where AC is academic change, CS is current score, and ATPA is the average of the two previous assessment scores.

In this study, each student's academic change, or growth, in reading and math will be used to determine performance. A positive or zero academic change will indicate the student is performing on and/or above expectations; while a negative academic change will indicate the student is performing below expectations.

SAS EVAAS

SAS EVAAS is a customized software system used by NC school districts to analyze student progress each year. EVAAS is based on more than ten years of research by Dr. William Sanders on value-added assessment. The Department of Public Instruction provided all NC LEAs with EVAAS in 2007-2008. Currently the tests included in EVAAS are EOG Reading and Math, all End of Course (EOC) subjects, Writing 4, 7, and 10, and Scholastic Aptitude Test (SAT). SAT is a standardized test taken for college admission (Ward, 2008).

EVAAS is populated with historical LEA test data, which follows the student through all NC schools and offers a precise measurement of how much gain or growth an individual student

or groups of students make over time. By linking each student's test records from grade to grade over subjects, it provides information about the overall effectiveness of an educational program. Much of the EVAAS research looks at the impact of the teacher.

EVAAS methodology extracts information from longitudinally linked student test scores. The closest measure to EVAAS is ABC growth. As stated previously, ABC growth examines a student's past performance on a maximum of two previous tests taking regression to the mean into account. EVAAS uses up to 5 years of test scores (at least 3 scores) to predict performance.

The School Diagnostic Reports list the current year's mean and previous years' mean for all subgroups. The means represent the difference between students' observed test performance and their predicted performance. A large negative mean indicates that students within a group made less progress than expected. A large positive mean indicates that students within a group made more progress than expected. A mean of approximately 0.0 indicates that a group is progressing at an average rate in the given subject and grade (NC Resource Guide for Value-Added Reporting, 2007).

As explained in the NC Resource Guide for Value-Added Reporting (2007):

> Students are assigned to Predicted Score Groups based on their predicted score in a given subject and grade. A student's predicted score is an expected score, based on his or her performance on previous tests, assuming the student is in an average school in the state.

EVAAS allows all tests to be used even though their scales are different. EVAAS accommodates all tests that are reliable, are highly correlated with curricular standards, and have sufficient stretch in the reporting scale to measure the achievement of both very low and very high achieving students. Students serve as their own control, creating a level playing field and

eliminating the need to adjust for race, poverty, or other socioeconomic factors (Ward, 2008). EVAAS, as stated previously, uses students' pervious scores to predict success probabilities.

In this study, to determine whether a student has met and/or is performing above expectations or is performing below expectations for reading and mathematics in EVAAS each student's obtained score will be compared to his/her predicted score. If the obtained score is equal to or greater than the predicted score, the student has met and/or is performing above expectations. If the obtained score is less than the predicted score, then the student is performing below expectations.

Collection Procedures

The following data collection procedures will be utilized for the study. There are no data available through the district data base identifying students who have been retained. Therefore, the researcher will examine each sixth grade student's yearly elementary and secondary school record from the 2009-2010 school year to determine retention status. The students retained in grades first through third will be assigned as group A and the students retained in grades fourth through sixth will be assigned as group B.

After the students have been identified, each sixth grade student's reading and math achievement growth scores generated by the state of North Carolina using the academic change formula for 2008-2009 will be sorted according to grade level of retention grades first through third or grades fourth through sixth. Each student's score will be analyzed and categorized as either performing on and/or above expectation or below expectations. A positive or zero academic change will indicate the student is performing on/and or above expectations; while a negative academic change will indicate the student is performing below expectations.

Next, each student's EVAAS predicted score and obtained score for 2008-2009 will be collected, sorted by grade level of retention, and calculated to determine growth. To calculate growth, each student's obtained score will be compared to his/her predicted score. If the obtained score is equal to or greater than the predicted score, the student has met and/or is performing above expectations. If the obtained score is less than the predicted score, then the student is performing below expectations. Once growth performance is determined each student's score will be categorized as either performing on and/or above expectations or performing below expectations for reading and mathematics.

Data Analysis

Separate Fisher's Exact Tests for reading and mathematics will be used to examine the frequency of distribution of student growth scores in these subjects as measured by the below expectation versus within or above expectations categories from the NC ABCs model as compared to students retained in grades first through third and grades fourth through sixth to see if the relationship of performance to category of retention differs more than expected by chance. The researcher will discuss patterns among variables and their percentage by categories.

Separate Fisher's Exact Tests for reading and mathematics will be used to examine the frequency of distribution of student growth scores in these subjects as measured by the below expectation versus within or above expectations categories from the EVAAS predicted scores as compared to students retained in grades first through third and grades fourth through sixth to see if the relationship of performance to category of retention differs more than expected by chance. The researcher will discuss patterns among variables and their percentage by categories.

For each two-tailed Fisher's Exact Test to be conducted in this study, the level of significance for the null hypotheses will be set at .05, or $p<0.5$. All statistical analyses will be

performed using the Predictive Analytics Software (PASW) (formerly SPSS) 17.0 quantitative software package.

The Fisher's Exact Test has been chosen as the statistic of analysis for this study because two-rows by two-columns contingency tables are required to analyze the relationships between students retained in early grades first through third versus students retained in later grades fourth through sixth and student growth performance in reading and mathematics. The Fisher Exact Test computes the exact probability of outcomes (Salkind, 2004). A series of two by two contingency tables will be constructed to express the relationship among variables.

Summary

The purpose of this study is to determine whether the timing of retention in early grades first through third as compared to later grades fourth through sixth demonstrates differences in sixth grade reading and mathematics growth. To determine growth using the state model, each student's academic change, or growth score will be analyzed and categorized as either performing within and/or above expectations versus below expectations. To determine whether a student has met and/or is performing above expectations or below expectations in EVAAS, each student's obtained score will be compared to their predicted score. Separate Fisher's Exacts Tests for reading and mathematics will be used to examine the frequency of distribution of student scores from both the NC ABCs model and EVAAS. In Chapter 4, the results of the Fisher's Exact Tests performed for this study will be presented.

CHAPTER 4: DATA ANALYSIS

The objective of this study was to determine the effects of grade retention on student achievement. The goal was to determine whether there was a significant difference in the growth performance scores and EVAAS predicted scores of sixth grade students who were retained in early grades first through third as compared to sixth grade students retained in later grades fourth through sixth utilizing the results from the NC End-of-Grade Tests in reading and mathematics. To accomplish this goal, separate Fisher's exact tests for reading and mathematics were used to determine if a relationship exists between the frequency of distribution of student growth scores for the two groups of repeaters in these subjects in the categories of within and/or above expectations versus below expectations from the NC ABC's model. As mentioned in Chapter 3, under the ABC's model growth is based on academic change. Academic change is expressed as the difference between a student's actual c-scale for the current year and the student's average of two previous assessments with a correction for regression toward the mean. A positive academic change indicates a gain in academic achievement, while a negative academic change indicates a loss of academic achievement from the previous two years. In this study, a positive or zero academic change indicated the student was performing on and/or above expectations; while a negative academic change indicated the student was performing below expectations.

Also, separate Fisher's exact tests for reading and mathematics were used to determine if a relationship exists between the frequency of distribution of student growth scores for the two groups of repeaters in these subjects in the categories of met and/or above expectations versus below expectations from the EVAAS predicted scores. To determine whether a student met and/or was performing above expectations or performing below expectations for reading and mathematics, each student's obtained score was compared to his/her EVAAS predicted score. If

the obtained score was equal to or greater than the predicted score, the student met and/or was performing above expectations. If the obtained score was less than the predicted score, the student was performing below expectations. In all, 4 two-tailed Fisher's exact tests were conducted with the significant level set at .05.

Demographics

The participants in this study were sixth grade students retained one time in grades first through sixth who attended the Onslow County School System in the 2008-2009 school year. The total number retained consisted of 236 students. Students were assigned to one of two groups based on grade of retention. Early retainees were students retained in grades first through third and later retainees were students retained in grades fourth through sixth. One hundred twenty-nine or 55% of the students were early retainees and 107 or 45% of the students were later retainees. Table 2 provides a summary by grade level of the number of early and later retainees.

While there were 236 students identified as being retained in the Onslow County School System in 2008-2009, due to missing data only 226 students were examined utilizing academic change scores and 198 students were examined utilizing EVAAS predicted scores. These data are reported in Tables 3 and 4.

When examining the distribution of the academic change scores for both groups of repeaters after dropping the missing scores, the distribution remained the same at 55/45. When examining the distribution of the EVAAS predicted scores for both groups of repeaters after dropping the missing scores, the distribution changed slightly 56/44.

Table 2

Number of Retainees per Grade

Grade Level	N	%
		Early Retainees (n=129, 55%)
First	68	53%
Second	40	31%
Third	21	16%
		Later Retainees (n=107, 45%)
Fourth	14	13%
Fifth	19	18%
Sixth	74	69%

Table 3

Early and Later Retainees per Grade with Reading & Math Academic Change Scores

Grade Level	N	%
		Early Retainees (n=125, 55%)
First	66	58%
Second	39	31%
Third	20	16%
		Later Retainees (n=101, 45%)
Fourth	14	14%
Fifth	17	17%
Sixth	70	69%

Table 4

Early and Later Retainees per Grade with Reading & Math EVAAS Predicted

Grade Level	N	%
		Early Retainees (n=111, 56%)
First	57	52%
Second	36	32%
Third	18	16%
		Later Retainees (n=87, 44%)
Fourth	12	14%
Fifth	14	16%
Sixth	61	70%

Research Questions

The primary research questions for this study were:

1. Is there a relationship in the growth performance as measured by the ABC's academic change score categories of within and/or above expectations versus below expectations between sixth grade students retained in early grades first through third as compared to sixth grade students retained in later grades fourth through sixth in reading in the Onslow County School System?
2. Is there a relationship in the growth performance as measured by the ABC's academic change score categories of within and/or above expectations versus below expectations between sixth grade students retained in early grades first through third as compared to sixth grade students retained in later grades fourth through sixth in mathematics in the Onslow County School System?
3. Is there a relationship in the growth performance as measured by the EVAAS predicted score categories of met and/or above versus below expectations between sixth grade students retained in early grades first through third as compared to sixth grade students retained in later grades fourth through sixth in reading in the Onslow County School System?
4. Is there a relationship in the growth performance as measured by the EVAAS predicted score categories of met and/or above versus below expectations between sixth grade students retained in early grades first through third as compared to sixth grade students retained in later grades fourth through sixth in mathematics in the Onslow County School System?

Findings

Academic Change Growth Performance Using ABC's Model for Reading and Math

The first research question examined whether there was a relationship in the growth performance as measured by the ABC's academic change score categories of within and/or above expectations versus below expectations between sixth grade students retained in early grades first through third as compared to sixth grade students retained in later grades fourth through sixth in reading. Throughout the remainder of this study, students retained in the early grades first through third will be referred to as early retainees and students retained in grades fourth through sixth will be referred to as later retainees. In examining the relationship between growth performance and timing of retention, a Fisher's exact test was performed. The analysis compared the total number of early retainees to the total number of later retainees and their performance as measured by the ABC's academic change score categories of within and/or above expectations versus below expectations.

The pattern of performance for the early retainees indicated that (n = 48, 38%) of the students scored within and/or above expectations while (n = 77, 62%) of the early retainees scored below expectations. The pattern of performance for the later retainees suggested that (n = 48, 48%) of the students scored within and/or above expectations while (n = 53, 52%) scored below expectations (see Table 5).

The results of the Fisher exact test when examining the academic change scores for reading indicated that a significant relationship does not exist between the growth performance of early retainees and later retainees χ^2 is .168, (p = 0.178) (see Table 6).

Table 5

Student Performance in Reading Utilizing Academic Change Scores

	N	Met/Above	Not Met/Below
Early Retainees	125	48 (38%)	77 (62%)
Later Retainees	101	48 (48%)	53 (52%)
Total	226	96 (42%)	130 (58%)

Table 6

Frequency Distribution and Results of Fisher's Exact Test for Reading Utilizing Academic Change Scores

	N	Met/Above	Not Met/Below
Early Retainees	125	48	77
Later Retainees	101	48	53
Total	226	96	130

Note. *p>.178.

The second research question examined whether there was a relationship in the growth performance as measured by the ABC's academic change score categories of within and/or above expectations versus below expectations between early retainees as compared to later retainees in math. In examining the relationship between growth performance and timing of retention, a Fisher's exact test was performed. The analysis compared the total number of early retainees to the total number of later retainees and their performance as measured by the ABC's academic change score categories of within and/or above expectations versus below expectations.

The pattern of performance for the early retainees indicated that (n = 68, 54%) of the students scored within and/or above expectations while (n = 57, 46%) of the early retainees scored below expectations. The pattern of performance for the later retainees suggested that (n = 55, 54%) of the students scored within and/or above expectations while (n = 46, 46%) scored below expectations (see Table 7).

The results of the Fisher's exact test when examining the academic change scores for math indicated that a significant relationship does not exist between the growth performance of early retainees and later retainees χ^2 is .993, (p = 1.000) (see Table 8).

EVAAS Predicated Scores Growth Performance for Reading and Math

The third research question examined whether there was a relationship in the growth performance as measured by EVAAS predicted score categories of met and/or above expectations versus below expectations between early retainees as compared to later retainees in reading. In examining the relationship between growth performance and timing of retention, a Fisher's exact test was performed. The analysis compared the total number of early retainees to

Table 7

Student Performance in Math Utilizing Math Academic Change Scores

	N	Met/Above	Not Met/Below
Early Retainees	125	68 (54%)	57 (46%)
Later Retainees	101	55 (54%)	46 (46%)
Total	226	123 (54%)	103 (46%)

Table 8

Frequency Distribution and Results of Fisher's Exact Test for Math Utilizing Academic Change Scores

	Met/Above	Not Met/Below	Total
Early Retainees	68	57	125
Later Retainees	55	46	101
Total	123	103	226

Note. *p>1.000.

the total number of later retainees and their performance as measured by the EVAAS predicted score categories of met and/or above expectations versus below expectations.

The pattern of performance for the early retainees indicated that (n = 42, 38%) of the students met and/or scored above expectations while (n = 69, 62%) of the early retainees scored below expectations. The pattern of performance for the later retainees suggested that (n = 44, 51%) of the students scored within and/or above expectations while (n = 43, 49%) scored below expectations (see Table 9).

The results of the Fisher exact test when examining the EVAAS predicted scores for reading indicated that a significant relationship does not exist between the growth performance of early retainees and later retainees χ^2 is .073, (p = 0.084) (see Table 10).

The fourth research question examined whether there was a relationship in the growth performance as measured by the EVAAS predicted score categories of met and/or above expectations versus below expectations between early retainees as compared to later retainees in math. In examining the relationship between growth performance and timing of retention, a Fisher's exact test was performed. The analysis compared the total number of early retainees to the total number of later retainees and their performance as measured by the EVAAS predicted score categories of met and/or above expectations versus below expectations.

The pattern of performance for the early retainees indicated that (n = 45, 41%) of the students met and/or scored above expectations while (n = 66, 59%) of the early retainees scored below expectations. The pattern of performance for the later retainees suggested that (n = 42, 48%) of the students scored within and/or above expectations while (n = 45, 52%) scored below expectations (see Table 11).

Table 9

Student Performance in Reading Utilizing EVAAS Predicted Scores

	N	Met/Above	Not Met/Below
Early Retainees	111	42 (38%)	69 (62%)
Later Retainees	87	44 (51%)	43 (49%)
Total	198	86 (43%)	112 (57%)

Table 10

Frequency Distribution and Results of Fisher's Exact Test for Reading Utilizing EVAAS Predicted Scores

	Met/Above	Not Met/Below	Total
Early Retainees	42	69	111
Later Retainees	44	43	87
Total	86	112	198

Note. *p>.084.

Table 11

Student Performance in Math Utilizing EVAAS Predicted Scores

	N	Met/Above	Not Met/Below
Early Retainees	111	45 (41%)	66 (59%)
Later Retainees	87	42 (48%)	45 (52%)
Total	198	87 (44%)	111 (56%)

The results of the Fisher exact test when examining the EVAAS predicted scores for math indicated that a significant relationship does not exist between the growth performance of early retainees and later retainees χ^2 is .276, ($p = 0.314$) (see Table 12).

Summary

This chapter included a detailed accounting for the results of the data analyses conducted for this study. The percentage of students retained in early grades first through third and later grades fourth through sixth was examined as well as the percentage of early retainees versus later retainees that met and/or performed above expectations or below expectations in reading and math utilizing both the ABC's academic change scores and the EVAAS predicted scores. The two Fisher's exact tests examining academic change using the ABC model for reading and math indicated that there were no statistically significant association between early retainees and later retainees. The two Fisher's exact tests examining the EVAAS predicted scores for reading and math also indicated that that there were no statistically significant association between early retainees and later retainees. Discussion and implications related to the findings of this study, as well as recommendations for further research, are included in Chapter 5.

Table 12

Frequency Distribution and Results of Fisher's Exact Test for Math Utilizing EVAAS Predicted Scores

	Met/Above	Not Met/Below	Total
Early Retainees	45	66	111
Later Retainees	42	45	87
Total	87	111	198

Note.*p >.314.

CHAPTER 5: SUMMARY AND DISCUSSION

The final chapter restates the research problem, reviews the methodology, and summarizes the findings from the study. The researcher drew conclusions and made connections between the current research and prior research. In addition, recommendations for educators, as well as, the researcher's insights, and recommendations for further research were provided. The chapter concludes with a summary of the study.

Statement of the Problem

In an era of increased accountability for student learning, administrators and teachers are forced to address the needs of struggling students. Grade retention is used as a practice to help low-performing students overcome their academic deficits. The objective of this study was to determine the effects of grade retention on student achievement. The goal of this study was to determine whether there was a significant difference in the growth performance scores and EVAAS predicted scores of sixth grade students retained in early grades first through third as compared to students retained in later grades fourth through sixth utilizing the NC End-of-Grade Tests in reading and mathematics in the Onslow County School System.

Review of the Methodology

The research design for this study used data from four separate sources to measure student achievement. First, the reading and mathematics growth performance scores for sixth grade students retained in early grades first through third and students retained in later grades fourth through sixth were compared. Separate Fisher's exact tests for reading and mathematics were used to determine the frequency of distribution of student growth performance scores in these subjects as measured by the within and/or above expectations versus below expectations categories from the NC ABCs model. This was done in order to compare early retainees and

later retainees to determine if the relationship of performance to category of retention differed more than expected by chance. Two hundred twenty-six students were examined utilizing the academic change scores. Of the 226 students, 125 were early retainees and 101 were later retainees.

Next, the reading and mathematics EVAAS predicted scores for sixth grade students retained in early grades first through third and students retained in later grades fourth through sixth were compared. To calculate growth, each student's obtained score was compared to his/her predicted score. If the obtained score was equal to or greater than the predicted score, the student met and/or was performing above expectations. If the obtained score was less than the predicted score, then the student was performing below expectations. Separate Fisher's exact tests for reading and mathematics were conducted to examine the frequency of distribution of student growth scores in these subjects as measured by the below expectation versus within or above expectation categories. The EVAAS predicted scores compared early retainees and later retainees to ascertain if the relationship of performance to category of retention differed more than expected by chance. One hundred ninety-eight students were examined utilizing EVAAS predicted scores. Of the 198 students, 111 were early retainees and 87 were later retainees.

Summary of the Results

First Research Question: The first research question examined whether there was a relationship in the growth performance as measured by the ABC's academic change score categories of early retainees as compared to later retainees in reading. The results of the Fisher exact test when examining the academic change scores for reading indicated that a significant relationship does not exist between the growth performance of early retainees and later retainees. When examining the reading academic change scores by percentage of early retainees and later

retainees who met and/or scored above expectations, the data demonstrated that the later retainees outperformed the early retainees 48% to 38%. The data also demonstrated that 62% of the early retainees scored below expectations and 52% of the later retainees scored below expectations.

Second Research Question: The second research question examined whether there was a relationship in the growth performance as measured by the ABC's academic change score categories between early retainees as compared to later retainees in math. The results of the Fisher's exact test when examining the academic change scores for mathematics indicated that a significant relationship does not exist between the growth performance of early retainees and later retainees. When examining the mathematics academic change scores by percentage of early retainees as compared to later retainees who met and/or scored above expectations, the data suggested there was no difference. Fifty-four percent of the early retainees met and/or scored above expectations and 54% of the later retainees met and/or scored above expectations. The findings also demonstrated that 46% of the early and later retainees scored below expectations.

Third Research Question: The third research question examined whether there was a relationship in the growth performance as measured by EVAAS predicted score categories between early retainees as compared to later retainees in reading. The results of the Fisher exact test when examining the EVAAS predicted scores for reading indicated that a significant relationship does not exist between the growth performance of early retainees and later retainees. When examining the reading growth performance as measured by EVAAS predicted scores by percentage of early retainees as compared to later retainees, the data indicated that the later retainees outperformed the early retainees 51% to 38%. The data also indicated that 62% of the

early retainees scored below expectations and 49% of the later retainees scored below expectations.

Fourth Research Question: The fourth research question examined whether there was a relationship in the growth performance as measured by the EVAAS predicted score categories between early retainees as compared to later retainees in math. The results of the Fisher exact test when examining the EVAAS predicted scores for math indicated that a significant relationship does not exist between the growth performance of early retainees and later retainees. When examining the mathematics growth performance as measured by EVAAS predicted scores by percentage of early retainees as compared to later retainees who met and/or scored above expectations, the data suggested that the later retainees slightly outperformed the early retainees 48% to 41%. The findings also demonstrated that 59% of the early retainees scored below expectations and 52% of the later retainees scored below expectations.

Conclusions

There are 3 major conclusions drawn based on the results.

1. Although results from the Fisher's exact tests indicated no significant relationship, the percentage of students who were later retainees demonstrated greater academic benefits in reading from retention when compared to early retainees. The data when analyzing academic change scores indicated that 48% (n = 48) of the later retainees met or scored above expectations as compared to 38% (n = 48) of the early retainees. Similar results were demonstrated utilizing the EVAAS predicted scores; 51% (n = 44) of the later retainees met or scored above expectations as compared to 38% (n = 42) of the early retainees.

2. After analyzing student performance in math, this study suggested that there was no significant difference in math performance between early and later retainees. Fifty-four percent of both groups of repeaters met or scored above expectations.
3. In this study, the data did not support the notion that the earlier a child is retained the better.

Relationship of the Current Study to Prior Research

Previous studies on grade retention have suggested that retention of students in grades first through third is probably more favorable than for students in grades fourth through sixth (Gleason et al., 2007; Greene & Winters, 2004, 2007; Jacob & Lefgren, 2004; Karweit, 1999; Lorence & Dworkin, 2006; Peterson et al., 1987; Pierson & Connell, 1992). There are two areas in this study that do not yield the same conclusions as prior research.

First, the study conducted by Meisels and Liaw (1993) examined retention comparing students retained in kindergarten through third grade with students retained in grades fourth through eighth. Although the researchers concluded that retention at any point was associated with less optimal academic outcomes as compared to non-retention, the data suggested that retention in grades kindergarten through third displayed significantly higher academic performance than students retained in grades fourth through eighth. Also, Jacob and Lefgren (2004) reached similar conclusions when they analyzed the impact of grade retention on third and sixth grade students. The data showed that the third grade students demonstrated positive academic benefits from retention while the sixth grade students did not. This is a distinct contrast to the results found in this study. In the present study when examining academic change scores in reading, the later retainees demonstrated greater academic benefits when compared to the early retainees 48/38. The same results were found when examining EVAAS predicated

scores in reading. The later retainees demonstrated greater academic benefits as compared to the early retainees 48/38.

Many educators believe that early retention kindergarten or first grade will more positively impact student performance. They believe that giving students and additional year to learn the material will provide them with the foundation needed to proceed successfully through the remainder of their education (David, 2008; Greene & Winters, 2004; Lorence & Dworkin, 2006). It is often assumed that students will be unable to learn the more advanced material in the next grade if they do not understand the subject matter of the current grade (Lorence & Dworkin, 2006). The findings from this study, however, clearly implied that students retained in the later grades demonstrated higher achievement than students retained in the earlier grades. These findings contradict the notion that the earlier a child is retained the better.

Similar findings were demonstrated in the Beginning School Study (BSS) conducted by Alexander et al. (2003). The BSS followed students in Baltimore city schools for eight years. The researchers found that students retained in first grade improved their achievement test scores the year they were retained. However, the first grade retainees began slipping as soon as they go beyond their repeated year and after just a few years they were far behind again. When first through third grade data were analyzed, the data indicated that the first grade students were helped the least by retention and the third grade students benefited the most. Students retained in the later grades fourth through seventh had less difficulty in school than the students retained in the early grades (Alexander et al., 2003).

The conclusions reached by this study and the BSS demonstrate similar results. The early retainees did not perform as well as the later retainees. While it is difficult to determine why the results were similar, there are three reasons why this may have occurred. First, research

indicates that students demonstrate initial academic gains during the year they are retained. However, numerous studies show that the achievement gains decline two to three years after retention (Lorence & Dworkin, 2006).

Second, students retained the earliest are usually the ones that struggle the most. Often times these students end up having more severe learning problems. Whereas, students retained in the later grades are generally not as far behind as the early retainees and perform better in school.

Finally, when students are retained early not only do they have academic difficulties; they generally have behavioral issues as well. Early retainees are usually less cooperative, less invested in classroom activities, more restless, and more easily distracted (Alexander et al., 2003). They have a difficult time adjusting to the classroom environment. By the time later retainees are retained, they have had time to adjust to school and the majority of time learning is the only factor.

Recommendations for Practice for Educators

The current study had no statistically significant findings. Since this study analyzed data from only one district, the recommendations for current practice are confined to the Onslow County School System:

1. Based on the number of students retained in the early grades kindergarten through third grade over a five year period (n = 1,858) in the OCS, high impact interventions both to prevent and respond to learning problems should be clearly communicated and implemented before retention occurs in grades first through third. Some of the alternatives to consider consist of: (a) provide double-dose periods for reading or math for students who are behind, (b) establish multiage classrooms for at risk

students in the primary grades, (c) provide supplemental instructional time through after-school, weekend, or summer programs, and (d) develop student support teams with appropriate professionals to assess and identify specific learning or behavior problems then design interventions to address the problems, and evaluate the efficacy of the interventions.

2. A yearly collection system needs to be implemented to identify retained students. At the current time, retained students are not identified and cannot receive services. By initiating a yearly collection system these students could be identified and receive additional services.
3. A cost analysis needs be conducted each year to find out how much money is being spent to educate retained students. The cost analysis should be used to help the school system decide how to most effectively spend school funds. Other alternatives to retention may be more cost effective.
4. Ongoing professional development needs to be provided that is designed to help teachers diversify their instructional approaches in ways that meet the instructional needs of the lowest performing students. The professional development should include training on multiple assessment measures that can be used to identify at risk students. As stated in Chapter 1, teacher quality is the most important school-related factor influencing student achievement (Rice, 2003). By providing ongoing professional development, Onslow County Schools can ensure that all teachers are qualified to teach all students.
5. The district office should develop clear policies regarding retention and promotion. The policies should include factors such as the number of times the student can be

retained, the maximum number of years behind grade level allowed, and cutoff scores on district mandated standardized tests. Student progress should be aligned with the Standard Course of Study's goals and objectives and retention decisions should also consider the same criteria.

Researcher's Insights

The conclusion from this study was surprising. The literature reviewed in Chapter 2 indicated that students retained in grades first through third demonstrated academic benefits from retention as compared to students in grades fourth through sixth. This was not the conclusion reached when examining the data obtained from the Onslow County School System. Retention in the earlier grades demonstrated lower achievement growth.

Educators believe that giving students an additional year to learn the material will help build a better foundation so they can proceed successfully through the remainder of their education (David, 2008; Greene & Winters, 2004; Lorence & Dworking, 2006). This study indicated the opposite. It suggested that later retention provided students with a greater opportunity for success.

It is the belief of this researcher based upon the results of this study that students in the earlier grades would be better served with intensive interventions that would accelerate learning. For example, double-dose periods of reading and math where students receive additional instruction. Also, carefully structured experiences should be provided such as one-on-one tutoring, after school tutoring, or summer school to help students catch up. A support team consisting of the teacher, parent, student, and other necessary personnel should develop a personalized educational plan for the student where research based strategies are implemented.

While implementing the strategies, constant progress monitoring should be conducted to determine any action corrections that might be necessary.

It would be unwise to suggest that students should never be retained. For some students, retention works. However, it is important for educational personnel to be familiar with the accumulated research on retention and its alternatives so that more informed decisions can be made regarding student educational placement.

Recommendations for Further Research

While this study's results did not indicate significant differences in performance between early and later retainees, the data did indicate that a higher percentage of later retainees met and/or scored above expectation in reading when compared to the early retainees. Based on these findings further research is recommended in the following areas.

The current study only considered student achievement after retention. Further research needs to be conducted measuring student achievement before retention so exact performance growth can be measured before and after retention.

This study should be expanded to include early and later retainees' performance at the eighth, tenth, and twelfth grades. The students should be tracked to determine their educational paths after high school.

Also, in this study academic achievement was the only factor considered. The study should be expanded to include parental perceptions of the practice of retention and the impact on the children. The study could also include perceptions from students previously retained and the educational outcomes experienced by these students. This would further clarify the retention experience.

Summary

Chapter 5 included a restatement of the research problem and a review of the study's methodology and findings. Conclusions, connections between current and prior research, recommendations for educators, researcher's insight, and recommendations for further research were also included. Three major conclusions were drawn from the study's findings (a) in this school system, later retainees demonstrated greater academic benefits in reading from retention when compared to early retainees, (b) there was no significant difference in math performance between early and later retainees, and (c) the study did not support the notion of the earlier a child is retained the better the child will be academically.

While the results from this study may not be generalized to another school district, the results from this study will provide educators of the Onslow County School District with valuable and usable information in regards to student placement.

REFERENCES

Alexander, K. L., Entwisle, D. R., & Dauber, S. L. (2003). *On the success of school failure: A reassessment of the effects of retention in the primary grades* (2nd ed.). New York: Cambridge University Press.

Anderson, G. E., Whipple, A. D., & Jimerson, S. R. (2002). *Grade retention: Achievement and mental outcomes*. Retrieved from http://www.cdl.org/resources-library/articles/grade_retention.php

Atherton, J. S. (2005). *Learning and teaching: Piaget's developmental theory*. Retrieved from http://www.learningandteaching.info/learning/piaget.htm

Bowman, L. J. (2005). Grade retention: Is it a help or hindrance to student academic success? *Preventing School Failure, 49*(3), 42-46.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Burkam, D. T., LoGerfo, L., Ready, D., & Lee, V. E. (2007). The differential effects of repeating kindergarten. *Journal of Education for Students Placed at Risk, 12*(2), 103-136.

Darling-Hammond, L. (1998). Alternatives to grade retention. *The School Administrator, 55*(7), 18-21.

David, J. L. (2008). What research says about grade retention. *Educational Leadership, 65*(6), 83-84.

Fager, J., & Richen, R. (1999). When students don't succeed: Shedding light on grade retention. (Report No. RJ96006501). Office of Educational Research and Improvement, Department of Education.

Frey, N. (2005). Retention, social promotion, and academic redshirting: What do we know and need to know? *Remedial and Special Education, 26*(6), 332-346.

Garland, R. (2009). North Carolina reading comprehension tests: Technical report. North Carolina: Department of Public Instruction, Academic Services and Instructional Support.

Gleason, K. A., Kwok, O., & Hughes, J. N. (2007). The short-term effect of grade retention on peer relations and academic performance of at-risk first graders. *The Elementary School Journal, 107*(4), 327-340.

Grant, J., & Richardson, I. (1998). *The retention /promotions checklist*. Peterborough, NH: Crystal Springs Books.

Greene, J. P., & Winters, M. A. (2004). An evaluation of Florida's program to end social promotion (Education Working Paper No. 7). New York, NY: Center for Civic Innovation at the Manhattan Institute.

Greene, J. P., & Winters, M. A. (2007). Revisiting grade retention: An evaluation of Florida's test-based promotion policy. *Education Finance and Policy*, 319-338.

Heubert, J. P., & Hauser, R. M. (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Academy Press.

Holloway, J. H. (2003). When children aren't ready for kindergarten. *Educational Leadership,* 60(7), 89-90.

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: Falmer Press.

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.

Hong, G., & Raudenbush, S. W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis, 27*(3), 205-224.

Jackson, G. B. (1975). The research evidence on the effects of grade retention. *Review of Educational Research, 45*(4), 613-635.

Jacob, B. A., & Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *Review of Economics and Statistics, 86* (1), 226-44.

Jerald, C. D., & Boser, U. (1999). Taking stock - quality counts'99. *Education Week, 18*(17), 81-82. Retrieved from Academic Search Premier database.

Jimerson, S. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *Journal of School Psychology, 30*(3), 420-437.

Karweit, N. L. (1999). Grade retention: Prevalence, timing, and effects (Tech. Rep. No. 33). Baltimore, MD: John Hopkins University and Howard University, Center for Research on the Education of Students Placed at Risk.

Kelly, K. (1999). Retention vs. social promotion: Schools search for alternatives. *Harvard Education Letter, 15*(1), 1-3.

Kubiszyn, T., & Borich, G. (1987). *Educational testing and measurement classroom application and practice* (2nd ed.). Glenview, IL: Scott, Foresman and Company.

Lorence, J., & Dworkin, A. G. (2006). Elementary grade retention in Texas and reading achievement among racial groups: 1994-2002. *Review of Policy Research, 23*(5), 999-1033.

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

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: Department of Education. Raleigh, NC: Department of Public Instruction.

North Carolina Department of Education Accountability Services. (2009). The abc's of public education academic change for schools 2008-09. Retrieved from www.ncpublicschools.org/docs/accountability/reporting/abc/2008-09/academicchange.Pdf

North Carolina Department of Public Instruction. (2001). Statistical profile 2001. Retrieved from www.ncpublicschools.org/fbs/resource/data/

North Carolina Department of Public Instruction. (2002). Statistical profile 2002. Retrieved from www.ncpublicschools.org/fbs/resource/data/

North Carolina Department of Public Instruction. (2003). Statistical profile 2003. Retrieved from www.ncpublicschools.org/fbs/resource/data/

North Carolina Department of Public Instruction. (2004a). NC standard course of study. Retrieved from www.ncpublicschools.org/curriculum.languagearts/scos/2004/25grade8

North Carolina Department of Public Instruction. (2004b). Statistical profile 2004. Retrieved from www.ncpublicschools.org/fbs/resource/data/

North Carolina Department of Public Instruction. (2005). Statistical profile 2005. Retrieved from www.ncpublicschools.org/fbs/resource/data/

North Carolina Department of Public Instruction. (2006a). Statistical profile 2006. Retrieved from www.ncpublicschools.org/fbs/resource/data/

North Carolina Department of Public Instruction. (2006b). Understanding the individual student report for North Carolina-grades 6, 7 and 8. Retrieved from www.education.com/print/Ref_Individual_Report/

North Carolina Department of Public Instruction. (2008). Evolution of the abc's. Retrieved from www.dpi.state.nc.us/docs/accountability/reporting/abc/2007-08/abcevolution.pdf

North Carolina Department of Public Instruction. (2010). Licensure. Retrieved from www.dpi.state.nc.us/licensure/

North Carolina Resource Guide for Value-Added Reporting. (2007). SAS Institute Incorporated. Cary, NC: USA.

North Carolina School Psychology Association. (2005). Position statement: Grade retention and the student accountability standards in North Carolina. Retrieved from www.ncschoolspy.org

Owings, W., & Kaplan, L. (2001). Standards, retention, and social promotion. *National Association of Secondary School Principals, NASSP Bulletin, 85*(629), 57-67.

Peterson, S., DeGracie, J., & Aybe, C. (1987). A longitudinal study of the effects of retention/ promotion on academic achievement. *American Educational Research Journal, 24*(1), 107-18

Pierson, L. H., & Connell, J. P. (1992). Effects of grade retention on self-system processes, school engagement, and academic performance. *Journal of Educational Psychology, 84*(3), 300-307.

Position statement on student grade retention and social promotion. (2003). Retrieved from www.nasponline.org/information/pospaper_graderetent.html

Powers and duties of principal GS 115C-288(a). Retrieved from http://ncga.state.nc.us/EnactedLegislation/Statues/HTML/BySection/Chapter_115c/GS_115

Rice, J. K. (2003). Teacher quality: Understanding the effectiveness of teacher attributes. *Educational Evaluation and Policy Analysis*, *22*(2), 129-145.

Roderick, M., & Nagaoka, J. (2005). Retention under Chicago's high-stakes testing program: Helpful, harmful, or harmless? *Educational Evaluation and Policy Analysis, 27*(4), 309-340.

Salkind, N. (2004). *Statistics for people who hate think they hate statistics*. Thousand Oaks, California: Sage Publications.

Smith, M. L., & Shepard, L. A. (1987). That doesn't work: Explaining policies of retention in the early grades. *Phi Delta Kappan, 66*(2), 129-134.

Steiner, K. (1986). Grade retention and promotion. Urbana, IL: Elementary Childhood Education. (ERIC Document Reproduction Service No. ED267899)

U. S. Department of Education. (1995). Dropout rates in the United States: 1995. Retrieved from http://nces.ed.gov/ pubs/dp95/97473-5.asp

U. S. Department of Education. (1999). Taking responsibility for ending social promotion: A guide for educators and state and local officials. Retrieved from http://www.ed.gov/ppubs/socialpromotion/

U. S. Department of Education. (2004). *No child left behind: A toolkit for teachers*. Jessup, MD: U.S. Department of Education.

U. S. Department of Education. (2006). Student effort and educational progress elementary/secondary persistence and progress. Retrieved from http://nces.ed.gov/programs/coe/2006/section3/indicators25.asp

U.S. Department of Education. (2008). The no child left behind act 2001. Retrieved from http://ed.gov/policy/elsec/leg/esea02/index.html

U.S. Department of Education, National Center for Education Statistics. (2006). The condition of education 2006, NCES 2006-071. Washington, DC: U.S. Government Printing Office

Ward, M. (2008). SAS EVAAS: A way of measuring schooling influence. NCARE 2008 annual conference proceedings.

APPENDIX: INSTITUTIONAL REVIEW BOARD APPROVAL LETTER

EAST CAROLINA UNIVERSITY
University & Medical Center Institutional Review Board Office
IL-09 Brody Medical Sciences Building• 600 Moye Boulevard • Greenville, NC 27834
Office 252-744-2914 • Fax 252-744-2284 • www.ecu.edu/irb

Date: November 15, 2010

Principal Investigator: Gail Pylant
Dept./Ctr./Institute: 202 Linda Loop
Mailstop or Address: Jacksonville, NC 28546

RE: **Exempt Certification** KK
UMCIRB# 10-0629
Funding Source: Unfunded

Title: "The Effects of Fairly Versus Later Grade Retention on Student Achievement"

Dear Ms. Pylant:

On 11.10.10, the University & Medical Center Institutional Review Board (UMCIRB) determined that your research meets ECU requirements and federal exemption criterion #4 which includes 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. ***NOTE: 1) This information must be existing on the date this IRB application is submitted. 2) The data collection tool may not have an identifier or code that links data to the source of the information.***

It is your responsibility to ensure that this research is conducted in the manner reported in your Internal Processing Form and Protocol, as well as being consistent with the ethical principles of the Belmont Report and your profession.

This research study does not require any additional interaction with the UMCIRB unless there are proposed changes to this study. Any change, prior to implementing that change, must be submitted to the UMCIRB for review and approval. The UMCIRB will determine if the change impacts the eligibility of the research for exempt status. If more substantive review is required, you will be notified within five business days.

The UMCIRB Office will hold your exemption application for a period of five years from the date of this letter. If you wish to continue this protocol beyond this period, you will need to submit an Exemption Certification Request at least 30 days before the end of the five year period

Sincerely,

Chairperson, University & Medical Center Institutional Review Board

Cc: Dr. Jim McDowelle