

A META-ANALYSIS OF THE EFFECTS OF GRADE RETENTION
OF K-6 STUDENTS ON STUDENT ACHIEVEMENT,
1990-2010

by

ANTHONY DEWAYNE BRIGHT

DAISY ARREDONDO RUCINSKI, COMMITTEE CHAIR
SARA TOMEK (TEMPLIN)
PHILIP WESTBROOK
BRENDA MENDIOLA
DAVID DAGLEY

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Education
in the Department of Educational Leadership,
Policy, and Technology Studies
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2011

Copyright Anthony Dewayne Bright 2011
ALL RIGHTS RESERVED

ABSTRACT

This dissertation investigates the relationship between grade retention and students' academic achievement for K-6 students. A meta-analysis was conducted from studies published between 1990 and 2010 that reported data on the effects of elementary grade retention and students' academic achievement. The primary hypothesis for this dissertation was that there was a positive relationship between grade retention and students' academic performance. An extensive systematic review of the literature was conducted using bibliographic databases and other sources, resulting in the review of hundreds of abstracts and articles. Initially, this review resulted in the identification of approximately 120 articles, from which, 68 were identified as potential studies for inclusion in this meta-analysis.

As data were abstracted from each potential study and evaluated, 43 studies remained for inclusion in this meta-analysis. Of these 43 studies, 31 either reported effect sizes in the results section or included sufficient data to calculate the effect sizes. After testing each study for statistical significance and eliminating insignificant studies, 26 studies remained. Effect sizes for these 26 studies were averaged and an effect size of medium strength was revealed ($ES = 0.50$). This effect size indicated that retained students scored 0.50 standard deviations lower than promoted students on academic outcome measures. Sixteen studies included in this meta-analysis had never been included in any prior meta-analysis, thus adding to the existing literature.

This study found that there was not a positive relationship between grade retention and students' academic performance. Results support the findings of most prior studies on grade retention concluding that grade retention is not an effective intervention. Major findings are

provided for the four research questions examined in this study. In addition, implications for practitioners and implications for researchers are included, as well as, suggestions for future research.

LIST OF ABBREVIATIONS AND SYMBOLS

d	Cohen's measure of sample effect size for comparing two sample means
df	Degrees of freedom
ES	Effect size
F	Fisher's F ratio
n	Number of cases in a subsample
N	Total number of cases
r	Pearson product-moment correlation
r^2	Estimate of the Pearson product-moment correlation squared
R^2	Multiple correlation squared; measure of strength of association
t	The sample value of the t -test statistic
η^2	Measure of strength of relationship (eta squared)

ACKNOWLEDGMENTS

Pursuing my doctoral degree has been a tremendous blessing to me. The entire process has helped strengthen my faith in God for without Him I would not have been able to complete my degree. I thank God for placing many great people in my path to lead, motivate, and assist me during this endeavor. From my parents, my high school teachers who molded my early life, my friends and family, and all of my professors at The University of Alabama, all of you are why I am where I am today. I thank God for all of you.

Specifically, I want to thank my wife, Lucy, and my daughters, Jena and Ashley, for always believing in me and supporting me. Not one time, over the last four years, did I hear any one of them complain to me about pursuing this degree. To the contrary, they were my biggest encouragers. I want to thank my dissertation chair, Dr. Daisy Arredondo-Rucinski, for her direct assistance, expertise, and willingness to walk me through the dissertation process step by step. Without her, achieving this accomplishment would have been much more difficult and much less enjoyable. Also, I want to extend my appreciation to Dr. Sara Tomek (Templin) for her assistance with the statistics used in this dissertation. In addition, I thank my other dissertation committee members, Dr. David Dagley, Dr. Philip Westbrook, and Dr. Brenda Mendiola for their contributions and outstanding leadership. I am very fortunate to have had the opportunity to work with each one of you. This degree is a great milestone in my life, but the friendships I have made along the way are far greater.

TABLE OF CONTENTS

ABSTRACT.....	ii
LIST OF ABBREVIATIONS AND SYMBOLS	iv
ACKNOWLEDGMENTS	v
LISTS OF TABLES.....	ix
INTRODUCTION TO THE PROBLEM	1
Introduction	1
Statement of the Problem	1
Significance of the problem	2
Purpose of the Study	3
Research Questions	3
Definitions of Terms	4
Limitations of the Study	5
Summary	5
REVIEW OF THE LITERATURE	7
Introduction	7
A Brief History of Grade Retention	7
Reform and Retention	7
Prior Meta-Analyses	8
Grade Retention Research 1990-1999	12
Effects of Early Grade Retention	12

Longitudinal Studies	15
Retention of Minority Students	21
Retention and Learning Disabilities	23
Teachers' Beliefs	25
Grade Retention Research 2000-2010	26
Longitudinal Studies	26
Longitudinal Studies Outside United States	32
Characteristics of Retained Students	35
Retention and Minority Students	37
Teachers' Beliefs	38
Alternatives to Retention	40
Accountability and Policy	43
Summary	44
METHODOLOGY	46
Research Design and Methodology	46
Problem and Purposes Overview	46
Research Questions and Hypotheses	47
General Research Design	49
Data Collection	50
Coding of Studies	51
Summary	53
RESULTS	54
Introduction.....	54

Data Set for Meta-Analysis.....	54
Study Procedures	55
Study Findings	57
1990-1999 Studies	58
2000-2010 Studies	67
Summary	80
DISCUSSION AND RECOMMENDATIONS.....	81
Introduction.....	81
Statement of the Problem.....	81
Review of the Methodology.....	82
Analysis of the Findings	83
Conclusions.....	89
Implications for Practitioners.....	90
Implications for Researchers.....	91
Summary	92
REFERENCES	94
APPENDICES	
Appendix A IRB Exclusion Letter.....	101
Appendix B Initial Screening Instrument	103

LIST OF TABLES

1. Studies Included in the Meta-Analysis	52
2. Homogeneity Test.....	57
3. Average Effect Size and Confidence Interval (1990-1999 Studies).....	57
4. Average Effect Size and Confidence Interval (2000-2010 Studies).....	66
5. Studies Related to Research Question 2	88
6. Studies Related to Research Question 3	88
7. Studies Related to Research Question 4	89

CHAPTER I

INTRODUCTION TO THE PROBLEM

Introduction

Schools across the United States are being held more accountable for their students' academic success. For the past several years, the standards movement has been influencing the nation's educational program due to the fear that U.S. students are not learning the information and skills necessary for college or full-time employment. Based on international assessments (Owings & Kaplan, 2001), U.S. students appear to be performing well below students from other countries, negatively affecting students, their communities, and the nation. No longer do U.S. students have to contend just with other U.S. students in the job market, they have to be able to compete globally. Therefore, the push for academic reform in the United States, which includes more accountability and higher standards, has led to the retention of more students, a very costly and ineffective intervention strategy (Owings & Kaplan, 2001).

Statement of the Problem

One type of intervention that begins early in many students' educational career is grade retention (Burkam, LoGerfo, Ready & Lee, 2007; Byrd & Weitzman, 1994; Cosden, Zimmer, & Tuss, 1993; Dennebaum & Kulberg, 1994; Dong, 2010; Johnson, Merrell, & Stover, 1990; Mantzicopoulos & Morrison, 1992). Even as early as kindergarten, teachers begin contemplating on whether or not to retain or promote students who have not quite mastered the basic skills that teachers believe to be necessary for the students' future success. According to Tomchin and

Impara (1992), teachers often choose to retain slow performing students in hopes of giving them an extra year to catch up academically and/or mature socially. However, what teachers believe contradicts much of the published research on grade retention (Tomchin & Impara, 1992). Furthermore, researchers do not agree on the effectiveness/ineffectiveness of grade retention, and Tanner and Galis (1997) have suggested that the reporting of the results from grade retention research studies sometimes appears to be biased and misleading in the literature. Therefore, it is important to properly examine the results of the most current and best-designed grade retention research reports that have studied retention and students' academic achievement.

Significance of the Problem

Fifteen to twenty percent of all students repeat at least one grade between the ages of 6 and 17, and it is estimated that 7 million elementary and secondary students will be retained at least once (Smink, 2001). As long as teachers at all grade levels believe retention is a suitable school practice that prohibits students from facing daily disappointment and inspires them to work harder in the classroom (Tomchin & Impara, 1992), retention will continue to be used as an intervention strategy. Due to grade retention, students devote an entire year repeating a grade, and U.S. taxpayers spend billions of dollars yearly providing for an extra year of schooling that generates nominal benefits and/or sometimes harmful consequences for the students (Shepard & Smith, 1990). Therefore, if retention is an ineffective educational practice for improving students' academic achievement as suggested by several contributors to the field (Holmes, 1989; Holmes & Matthews, 1984; Jackson, 1975; Jimerson, 2001), or if it is an effective educational practice for improving students' academic achievement as others argued (Allen, Chen, Willson, & Hughes, 2009; Lorence, 2006), educational practitioners and policy makers need to be

knowledgeable of the available research in order to make informed decisions about retaining or not retaining students.

Purpose of the Study

The purpose of this study is to conduct a comprehensive meta-analysis of available scientific studies published between 1990 and 2010 that have examined the effects of elementary school grade retention on students' academic achievement. Retention has a long history of support from many educational professionals, parents, and the general public even though it is viewed as having a negative effect on academic achievement in much of the research literature (Holmes, 1989; Jimerson, 2001). The intent of this study is to report the results of the meta-analysis so that teachers, parents, school administrators, and school policy makers can make better informed decisions to retain or not to retain students based on empirical evidence. In addition, this study proposes to resolve some of the differences in the literature on grade retention by examining current literature that has not been included in previous meta-analyses.

Research Questions

The research questions posed in this study are as follows:

1. What is the research consensus on the effects of grade retention on academic performance of students?
2. If initial gains are evident in academic achievement from repeating a grade in elementary school, does the research show that they persist in high school?
3. Does the research show that retaining a student in kindergarten improves the academic performance of the student?
4. Does the research show that retaining students early (Grades 1-2) benefits students' academic performance more than retaining them later (Grades 3-6)?

Definitions of Terms

At-Risk—a student who is performing poorly academically and/or in danger of dropping out of school.

Dropout—person who is at least 17 years of age, has not graduated from high school, and is not enrolled in any type of school.

Early retention—retaining a student in first or second grade.

Effect size—measure indicating the degree of departure from the null hypothesis in standard units (Cohen, 1977).

High school—implies grades 7-12 in this study.

Kindergarten retention—requiring a student to repeat another year of kindergarten.

Late retention—retaining a student in third, fourth, fifth, or sixth grade.

Meta-analysis—refers to the analysis of analyses, the statistical analysis of a large collection of analyses results from individual studies for the purpose of integrating the overall findings (Glass, 1976).

Retention—practice of requiring a student who has been in a particular grade for a full school year to repeat that grade the succeeding school year (Jackson, 1975).

Same-age comparison—comparison of pupils at different grade levels but with the same number of attended school years and normally the same age (Bonvin, Bless, & Schuepbach, 2008).

Same-grade comparison—comparison of pupils at the same grade level who are of different ages due to retention (Bonvin et al., 2008).

Social promotion—allowing a student to go on to the next grade although he/she did not meet all academic criteria of previous grade.

Limitations of the Study

The limitations for this study are as follows:

1. This study will be limited to an examination of the literature regarding grade retention and the effects on K-6 students.
2. All reviews will come from peer-reviewed journals or books published between 1990 and 2010.
3. Some studies may have small sample sizes.
4. Some statistical data may not be available.
5. Analysis is constrained by the quality of the studies that are available.

Assumptions

1. The author(s) will not knowingly express any biases when reporting data.
2. The data reported will be accurate.

Summary

This chapter provided an introduction to grade retention and some of the issues related to its use as an intervention strategy for students' academic achievement. Grade retention is an expensive intervention, and there is a disconnect between practitioners' personal beliefs and the majority of researchers' findings. The continuous and perhaps increased use of grade retention in the United States is driven by the push for higher standards and more accountability. This study is a meta-analysis of scientific studies between 1990 and 2010. A meta-analysis is a systematic approach used to evaluate the data supporting claims made to explicate the overall results, and to identify where additional information and research are needed.

Chapter II includes a review of the literature in the topic areas that were included in the meta-analysis. The chapter describes review articles and studies that did not meet criteria for

inclusion in the meta-analysis, but which provide contextual information on grade retention and students' academic achievement. Chapter III describes the methodology used in this study and explains the rationale for undertaking a meta-analysis. Chapter IV explains the data set and study procedures for the meta-analysis. The chapter also examines the specific findings from the comprehensive review of the studies included in the meta-analysis. Chapter V describes the implications of the research findings and discusses recommendations for future research.

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

A review of the literature relevant for this study is presented in this chapter. The review is presented as follows: (a) a brief history of grade retention, (b) grade retention research 1990-1999, and (c) grade retention research 2000-2010.

A Brief History of Grade Retention

Reform and Retention

Grade retention became an issue in public education around 1860 after school systems began to categorize their students into grade levels in order to allow teachers a more uniform grouping in which instruction could be addressed (Holmes & Matthews, 1984; Owings & Magliaro, 1998). As a result, students failing to make adequate progress in their academic achievement were often held back or retained.

Although the use of grade retention as an educational intervention for low-achieving students has varied over the last 40 years, it remains a common educational practice. In 2004, U. S. Census data revealed that 9.6% of U.S. youths between the ages of 16 and 19 had been retained in grade one or more times, suggesting the pendulum has again moved toward higher rates of grade retention (Bali, Anagnostopoulos, & Roberts, 2005; Wu, West, & Hughes, 2008b). Owings and Kaplan (2001) credit the rise in grade retention rates from the 1980s to the mid-1990s to the intensification of the criterion-based reform movement in education that followed the publication of *A Nation at Risk: The Imperative for Educational Reform* (National

Commission on Excellence in Education, 1983) which emphasized setting competency standards for students at each grade level and holding both schools and students accountable for meeting them. This document rallied the nation against declining educational standards, called for an end to social promotion, and directed public sentiment back to retention (Owings & Kaplan, 2001). Social promotion gained prominence in the 1970s arising as a consequence of the movement away from grade retention. Social promotion is motivated by a desire to protect the social adjustment and school motivation of struggling students. It is believed by the proponents of social promotion that students who are promoted will learn more from exposure to new content at the next grade than they would from repeating their current grade. However, due to the declines in student achievement test scores as evidenced by *A Nation at Risk*, lenient policies like social promotion were believed to have caused a dilution of standards and a decline in the quality of American education (Roderick, 1995). President William Clinton in his 1998 and 1999 State of the Union addresses urged an end to social promotion and stated that scores on standardized tests should be the basis for promotion. The No Child Left Behind (2002) federal legislation passed in 2001 required assessments, aligned with state standards, be used to measure the achievement of all children at grade level. No Child Left Behind (NCLB) recommended grade retention as an intervention for low performing students, resulting in increased numbers of retained students in the past decade (Lazarus & Ortega, 2007)

Prior Meta-Analyses

Grade retention has been a source of debate in American public education since its inception. Beginning in the early 1900s, there have been many studies published on the subject of grade retention. However, most studies on grade retention published since 1975 have consistently referenced one major review (Jackson, 1975) and/or three meta-analyses (Holmes,

1989; Holmes & Matthews, 1984; Jimerson, 2001). Recently, Allen et al. (2009) published a fourth meta-analysis on grade retention challenging the view that grade retention is negative, contradicting the findings of the previously mentioned studies. Therefore, a review of the aforementioned studies will be described in order to put this study into context.

In 1975, Jackson reviewed 30 published studies between 1911 and 1973. From his analysis, Jackson concluded there were not enough adequate studies to make broad generalizations about the effects of grade retention on students' academic achievement. Flaws in the studies he reviewed included: (a) a lack of representative samples of the nation's schools and students, (b) the most recent research in the study was over 30 years old, and (c) the studies failed to investigate the long-term effects of grade retention. Thus, Jackson (1975) found no reliable body of evidence to indicate that repeating a grade is more beneficial, or less beneficial, than being promoted for students with academic or adjustment difficulties.

In 1984, Holmes and Matthews conducted a meta-analysis that included 44 studies published between 1929 and 1981, including 13 studies from Jackson (1975). The 44 studies consisted of 18 published studies, 14 dissertations, and 12 master's theses. To have been included in Holmes and Matthews' final list, the study must have (a) presented the results of original research of the effects on pupils of retention in the elementary or junior high school grades, (b) contained sufficient data to allow for the calculation or estimation of an effect size, and (c) compared a group of retained pupils with a group of promoted pupils. In all, 575 individual effect sizes (*ES*) were calculated and averaged, resulting in a mean *ES* of - 0.37. Simply stated, retained students scored .37 standard deviations lower on the various outcome measures than the promoted students. Hence, Holmes and Matthews' research found evidence to

support that the potential for negative effects regarding retention outweighs the positive outcomes. Therefore, those who retain students do so against the evidence.

Holmes (1989) conducted a second meta-analysis using a total of 63 studies published between 1925 and 1989. The 63 studies consisted of 20 published studies, 22 dissertations, 18 master's theses, and 3 unpublished papers. Forty-four of these studies had been included in the Holmes and Matthews' (1984) meta-analysis. In fact, the criteria used to narrow the list of studies in the 1989 meta-analysis were the same as the 1984 study. Holmes, hoping to extend the synthesis of his previous work, included 19 new studies. Of the 63 studies, 54 studies indicated that retention produced a negative result, while nine studies indicated that retention produced a positive result. A total of 861 individual effect sizes were calculated and the mean *ES* was - 0.15. Again, the group of retained pupils scored lower than the promoted comparison group on the various outcome measures. The overall effect from Holmes' 1989 meta-analysis was somewhat less negative than that reported in Holmes and Matthews' 1984 meta-analysis. The difference in the two meta-analyses seemed to come from the 19 new studies, seven of which reported results in favor of retention. However, Holmes (1989) concluded that the empirical evidence argues against retention.

A third meta-analysis conducted by Jimerson (2001) included 20 articles published between 1990 and 1999. The criteria Jimerson used for inclusion in his review were: (a) research must have been presented in a journal article or book, (b) results must have addressed the efficacy of grade retention, (c) study must have included an identifiable comparison group of promoted students, and (d) research must have been published during the 1990-1999 timeframe. A mean *ES* of - 0.31 was calculated from a total of 246 individual effect sizes favoring the promoted comparison group. Of the 20 studies in Jimerson's (2001) meta-analysis, only four

studies favored grade retention. It should be noted that the authors concluding in favor of grade retention proposed that additional remedial strategies are important to facilitate the educational success of students. The results of his meta-analysis also reinforced the importance of considering the longitudinal outcomes of retained students.

Recently, Allen, Chen, Willson, and Hughes (2009) conducted a fourth meta-analysis on grade retention. Allen et al. reviewed 22 studies published between 1990 and 2007. Fifteen of the studies were included in Jimerson's (2001) meta-analysis along with seven new studies published between 2003 and 2007. To be included in the analysis, studies had to meet the following criteria: (a) be conducted in North America and published in English between 1990 and 2007 in a scholarly book or peer-reviewed journal, (b) report outcomes of retention at the student level with quantifiable measures of academic achievement, (c) use a quasi-experimental or experimental design with at least one comparison group of similarly achieving students or employed covariates statistically to control for pre-retention differences between retained and promoted children, and (d) not duplicate previously reported achievement outcomes. The mean weighted *ES* calculated from the 207 individual and study level effect sizes was - 0.11 indicating that the promoted group benefited slightly more than the retained group. Nonetheless, the results of this meta-analysis challenge the widely held view that retention has a negative impact on achievement. The authors believed the empirical evidence on grade retention is inconsistent but is generally characterized in published literature as negative. Allen et al. (2009) challenged the methodological quality of the meta-analyses conducted by Holmes (1989) and Jimerson (2001) and asserted that their study used a more careful evaluation of design quality from the existing recent literature, thus allowing them to make better conclusions regarding grade retention.

Grade Retention Research 1990-1999

Effects of Early Grade Retention

Johnson, Merrell, and Stover (1990) examined the academic effects of early grade retention on a group of fourth-grade students who were retained in either kindergarten or first grade. The major inquiry in their research was whether fourth-grade children who were retained at the kindergarten or first-grade level displayed significantly greater academic achievement when matched to their grade peers who were selected for retention but did not repeat. A total of 57 students were selected from a group of 604 fourth-grade students from four school districts in a western state.

Johnson et al. (1990) concluded that the use of early grade retention was not a successful academic intervention for the retained students based on the fact that their findings showed no significant differences in the achievement results between the retained and non-retained groups. On the other hand, the data did not support the practice of social promotion as being effective either. Therefore, educational professionals need to know what the data reveal concerning the practice of grade retention and look for other interventions to both retention and social promotion alone.

In 1992, McCombs-Thomas, Armistead, Kempton, Lynch, Forehand, Nousiainen, Neighbors, and Tannenbaum examined whether retention of kindergarten and first-grade children was linked to any long-term favorable effects. Participants were 62 students from the fourth and fifth grades of a rural elementary school. Thirty-one students had been retained in kindergarten or first grade and were compared to 31 children who had similar grades but had never been retained.

Results from the study indicated that retention was not associated with long-term favorable effects but rather, especially for White children, appeared to be associated with poorer academic and social functioning. Retained students, regardless of race, were viewed as having more anxiety and withdrawal problems than those students who were promoted. The authors stated that a major strength of their study was the ability to examine the long-term effects of retention from the data that were collected over a four-year period after retention. Something, at the time, that had rarely been done (McCombs-Thomas et al., 1992).

Mantzicopoulos and Morrison (1992) examined the effect of retention at kindergarten on academic achievement and behavior on 53 students from two school districts in California. The retained students were matched to a group of 53 promoted peers on demographic characteristics, a measure of school readiness, and pre-academic achievement in reading and mathematics. Both same-age and same-grade comparisons were used when evaluating the data.

Mantzicopoulos and Morrison's (1992) results indicated that students retained in kindergarten benefited academically during their second year of kindergarten, but the advantage was not maintained past kindergarten. Academic gains during the second year of kindergarten were attributed to taking the same end-of-the-year test as taken the previous year, as well as, being a year older at the time of the test. In addition, the retained students in the study exhibited a decline in attention problems during their second year of kindergarten. However, despite these findings, Mantzicopoulos and Morrison (1992) do not support kindergarten retention for the young at-risk student.

Kindergarten plays an important part in a child's socialization to the school system. However, due to increased academic accountability at all grade levels, the kindergarten curriculum has intensified. As a result, many parents hold children out of school until they

believe their children are academically prepared. Retaining children in kindergarten who are not prepared for first grade has become a frequent practice. Cosden, Zimmer, and Tuss (1993) analyzed the impact of the age of entry to kindergarten and the use of retention as functions of a child's sex and ethnicity in three school districts in southern California.

Cosden et al. (1993) ascertained that school entry and kindergarten retention and promotion decisions were related to some extent to a student's age, sex, and ethnicity. The authors suggested that retention policies may have a greater impact on some groups of students than others. For example, if younger students and students with less preparation for formal schooling are more likely to be retained, then minority students, particularly those from lower socioeconomic backgrounds, may also be more vulnerable to retention.

Byrd and Weitzman (1994) reached a similar conclusion that early grade retention is common and that a number of factors are independently associated with the increased risk of retention. Predictors associated with early grade retention in their study included poverty, male gender, low maternal education, deafness, speech defects, low birth weight, and exposure to household smoking. To the contrary, high maternal education and living with both biological parents at age six were independently linked with a decreased risk of retention. Understanding the factors that suggest a child is at-risk is important, but in some circumstances, efforts should be aimed at referring such children and families to services that might boost the child's preparedness for school (Byrd & Weitzman, 1994). Comprehending why children fail is important. The larger issue is how to help children succeed.

So, what happens to children who are perceived as not ready and who fail to meet the high expectations for kindergarten and first-grade? Sometimes, these children are placed in some form of extra-year program, including repeating kindergarten or placement in a transitional

classroom before first grade. Dennebaum and Kulberg (1994) analyzed the relationship of such programs and later school achievement. Ninety-five fourth- and fifth-grade students, from a Rhode Island school district, were identified as either retained in kindergarten, placed in a transition classroom, recommended for an extra-year program but went on to first grade, or went from kindergarten to first grade without hesitation.

Results revealed that children retained in kindergarten performed significantly lower on a standardized achievement test than did children in the other comparison groups. Despite an extra-year of schooling, students placed in transition classrooms did not differ significantly in their performance from students who were recommended for an extra-year, but went on to first grade, and students in the control group. The study does not support the notion that holding a student back will help later achievement. According to Dennebaum and Kulberg (1994), it seems more beneficial to move the student along and promote him/her to first grade.

Longitudinal Studies

Longitudinal research is an important methodology for studying retention because it allows the researcher to understand how the effects of retention change or persist over time (Tanner & Galis, 1997). Pierson and Connell (1992) conducted a study with third- through sixth-grade students from a suburban/rural school district and a suburban/urban school district in upstate New York. The authors compared 74 retained students with 60 random and 69 matched-ability non-retained students from their present classroom, as well as, 35 socially promoted students from earlier classrooms when the retained students were held back.

Results of Pierson and Connell's (1992) study indicated that students whose academic performance suggested that they be retained, and who were retained, performed better two or more years later than students with comparable performance who were promoted. However,

children who were retained did not perform as well academically as students randomly selected from their current classroom who had not been retained. Retained students performed as well as the matched ability sample and better than the socially promoted sample. Overall, the research evidence suggested that retention in the early years of elementary school was not harmful to general self-worth or to perceived relatedness to peers, and in comparison with social promotion, was beneficial to academic performance. Therefore, the findings supported the use of retention as a potentially effective remediation strategy for academic difficulty in the early elementary grades.

Most of the students in the study were retained in first and/or second grade, were retained only once, and outcomes were assessed only in elementary and middle school. The authors stated that no conclusions can be drawn about the effects of retaining older students, the effects of retaining students more than once, or the longer term effects of retention in high school or beyond. They also noted that without attending to the origin of students' academic difficulties and without motivational support by parents and teachers, holding children back may not work (Pierson & Connell, 1992).

In another longitudinal study, Phelps, Dowdell, Rizzo, Ehrlich, and Wilczenski (1992) examined the long-term efficiency of retention and pre-grade transition five to ten years post-placement. Seventy White students from a small suburban school district east of Buffalo, New York, participated in the study. Sixty-five percent of the sample had been retained or transitioned at the conclusion of their kindergarten or first-grade year, and the students were completing either their seventh-, eighth-, or ninth-grade year at the time of the study. Twenty-four students in the study had been retained, 22 had received transitional placement (smaller group instruction), and 24 had been advanced into the next grade in the usual manner.

Results from the study concluded that students placed in the transitional program did not benefit from the intervention even though they received one full year of compensatory educational programming. In fact, the students performed three-quarters of a standard deviation below their peers in reading achievement. In addition, students in the transitioned group and the retained group required more special education/remedial services than the matched sample group. Simply repeating the same curriculum or providing smaller group instruction did not sufficiently benefit the transitioned or retained students in the study (Phelps et al., 1992).

Reynolds (1992) tested the effect of early grade retention on reading achievement, mathematics achievement, teacher ratings, and perceived competence in fourth grade for 1,255 low-income, mostly Black children. Participants were part of The Longitudinal Study of Children at Risk in the Chicago Public Schools, who enrolled in kindergarten in the fall of 1985 and remained in the school system in fourth grade, allowing effects to be traced up to three years post-retention. Also, a more restrictive comparison group of 200 promoted children were tested.

Longitudinal data revealed that retention had mixed effects on students' school adjustment in fourth grade. Retention had substantially negative effects on cognitive achievement in reading and mathematics, and retention was unrelated to teacher ratings of school adjustment. However, retention positively affected students' perceived school competence, especially for early-retained children. Analyses of the comparison group of 200 promoted children produced similar results. Reynolds (1992) concluded that grade retention was not an effective educational practice with children at-risk of underachievement.

Meisels and Liaw (1993) examined the occurrence of retention from kindergarten to eighth grade using data from the National Educational Longitudinal Study of 1988 (NELS: 88). Participants in the analytic sample consisted of 16,623 eighth-grade students of which 2,075 had

been retained in kindergarten through third grade (early retainees), 1,128 had been retained in grades four through eight (later retainees), and 13,420 who had never been retained. In general, students who were retained showed a disadvantage on later school performance outcomes when eighth-grade outcomes of those who had been retained were compared with those who had never been retained. In addition, the results suggested that retention at any point is associated with less optimal academic outcomes and strengthens the argument against retention policies.

Meisels and Liaw's (1993) study provided a national picture of retention and confirmed the negative outcomes associated with failure in grade. Until this study, no one had studied retention in a sample as large and as representative of the public school population as these authors examined. The data on the 16,623 White, Black, and Hispanic students showed that boys, minorities, and students from lower socioeconomic status (SES) were more likely to be retained. Furthermore, retention was associated with more negative outcomes for female, White, and higher SES students. The authors stated that retention had substantially more negative than positive associations, and despite the popularity of retention among practitioners, it is apparently useless to students and should be used only on rare occasions.

Dauber, Alexander, and Entwisle (1993) conducted a longitudinal study of inner-city children in Baltimore City Public Schools. Their study examined various predictors of retention in 728 first- through fourth-grade students participating in the Beginning School Study (BSS). Socio-demographic variables (e.g., race, gender, family's education, and economic status), test scores, teacher and parent judgments of children's ability, and adjustment to school were used to separate retained from promoted children, as well as, children retained in first grade from those retained in later grades. The children were randomly selected from the cohort that entered first grade in the fall of 1982. Of the 728 students in the study, 202 had been retained once. From the

202 retained students, 87 had been retained in first grade, 52 in second, 44 in third, and 19 in fourth. An additional 24 students had been retained at least twice, 20 of whom were initially retained in first grade.

Dauber et al. (1993) concluded that children who are retained in early elementary school typically fit the at-risk profile. Retained children experience serious academic difficulties that become evident very early in their school careers. In fact, academic difficulties are the most important factor predicting retention, according to the authors. Low marks in reading and math and low scores on standardized tests are powerful predictors of grade retention.

Other results from the study (Dauber et al., 1993) included that boys were more often held back than girls. This finding has been ascribed to a variety of possible factors, such as boys developing later than girls, being more insistent and disorderly, having trouble adapting to the predominantly female environment of the elementary school, and being unfavorably affected by a father's absence. Also, retained students were more often Black, had less educated parents, came from poverty-level families, demonstrated poorer test performance, rated by their parents as less able to complete school lessons, received lower first-quarter conduct grades, and were more likely to have transferred schools between kindergarten and first grade.

Kundert, May, and Brent (1995) explored whether there were differences in IQ and achievement (at second, fifth, and seventh grade) between students who delayed school entry and those who were retained in later elementary grades. Participants in the study were middle-class students enrolled in a suburban district that served 3,238 students, all of whom were White. Two groups of students, retained and delayed entry, currently enrolled in Grades 3-12 were compared in the study. The retained group consisted of 314 students who had been retained during Grades K-5, and the delayed entry group consisted of 157 students who had been held out of

kindergarten one year beyond the time they were eligible to enter school. There was a significant 6-point difference in IQ, favoring the delayed-entry group.

Results of the study found, when achievement test scores were analyzed, that there were no significant differences identified for either group, despite the initial difference in IQ. The authors reached the conclusion that instructional reform, not retention or delayed school entry, may be a more appropriate and effective intervention for improving a student's success in school (Kundert et al., 1995).

In 1997, Mantzicopoulos re-examined whether or not early (kindergarten) retention resulted in positive long-term academic and behavioral outcomes for a subgroup of kindergarten children with attention problems. Forty children (25 retained and 15 promoted) with high inattention, selected from a sample of 62 children (32 retained and 30 promoted) who had taken part in an earlier study by Mantzicopoulos and Morrison (1992) participated in the study.

The results of the investigation did not support the notion that kindergarten retention was a beneficial educational intervention for children with academic and/or behavior difficulties. At the completion of second grade, there were no consistent academic benefits for retained children with attention problems in the study. On the other hand, there was no evidence to conclude that early retention produced more negative effects than promotion. However, the author concluded that being overage for grade was one confirmable result of retention (Mantzicopoulos, 1997).

Jimerson, Carlson, Rotert, Egeland, and Sroufe (1997) examined the characteristics of children retained in early elementary school and the effects of retention on achievement and adjustment throughout the elementary years. The participants were selected from 190 children taking part in the Minnesota Mother-Child Interaction Project, a longitudinal study of children at-risk for problems in social and emotional development, and placed in either the retained

group, the low-achieving promoted group, or the control group. The retained group consisted of 29 children who were retained only once in Grades K-3. The low-achieving promoted group consisted of 50 children, and the control group consisted of 100 children, 25 participants per group in Grades K-3. The participants attended more than 120 different schools from 1982 to 1986, the period during which the students were retained in Grades K-3.

The findings from the study found little evidence to support retention as an intervention for improving educational outcomes, despite controlling for initial levels of achievement and adjustment. It was concluded that elementary grade retention was an ineffective intervention for both achievement and adjustment. Retention may appear to help early academic performance, but it is suggested by Jimerson et al. (1997) that because the retained subjects continued to display comparable achievement to their low-achieving peers at age 16, retaining students did not benefit them academically and was harmful in terms of their self-esteem and behavior problems.

Retention of Minority Students

A few studies have found that students from ethnic minority groups are more likely to be retained or identified with early school problems. For example, Cosden et al. (1993) found that a significantly greater proportion of Latino than White students were identified as having school problems in two of three districts studied. The rising academic demands within kindergarten programs, along with the lack of knowledge concerning the early schooling experiences of Latino students and the uncertain impact these experiences have on later school outcomes, have resulted in some parents holding their children out of school until they are older and better prepared to succeed. At the same time, schools have increasingly relied on retention in

kindergarten to prevent children who were not prepared from going to first grade (Cosden, Zimmer, Reyes, & Gutierrez, 1995).

Cosden et al. (1995) assessed the impact of gender, ethnicity/home language, holding out, retention and social promotion on first-grade achievement of 616 children across 11 elementary schools in a school district in southern central California. The authors' findings suggested that neither holding out, nor retention, nor social promotion had their projected results on students' achievement. With regard to holding out, there was no evidence that younger children who entered on schedule were at a disadvantage comparative to their classmates who delayed a year to enter school. This is important, given the widely held belief that younger children will be more successful if they enter school a year older. In addition, the study found that children who were retained or promoted performed more poorly than their classmates at the end of first grade, suggesting that neither intervention had a positive impact on the students' achievement. These results confirm Mantzicopoulos and Morrison's (1992) findings of retention in kindergarten and speak against social promotion, even in kindergarten. Results of the study also revealed that Latino students were more likely than White students to be promoted, with Spanish-speaking Latino students being the most vulnerable to advancement. The study indicated that the needs of Latino students must be addressed early in their schooling.

McCoy and Reynolds (1999) conducted a follow-up study of the predictors and consequences of grade retention up to age 14 on 1,164 low-income, minority (95% Black, 5% Hispanic) children from the Chicago Longitudinal Study previously conducted by Reynolds (1992). About 20% of the study sample was retained between kindergarten and third grade with 50% of those students retained in first grade.

The strongest predictors of retention identified by the study were academic performance, parental participation in school, school mobility, and gender. Grade retention was significantly related to lower reading and mathematics achievement at age 14. Results based on same-age comparison groups yielded larger effects of retention on school achievement than results based on same-grade comparison groups. However, both methods indicated that grade retention was associated with significantly lower reading achievement. Like the earlier study (Reynolds, 1992), these findings suggested that intervention tactics other than grade retention are needed to better advance school achievement and adjustment (McCoy & Reynolds, 1999).

Retention and Learning Disabilities

Several studies on grade retention have looked at the retention of average students within the public schools, but very little research has focused on retention among students with learning disabilities (LD). McLeskey and Grizzle (1992) designed their study to address this void in the literature. McLeskey and Grizzle (1992) collected data on 689 students referred and identified as having learning disabilities during the 1987-88 school year in Indiana. Data were collected for students in Grades K-12 with 45% of the students being referred and labeled while in Grades K-2, 34% being labeled in Grades 3-5, and the remaining 22% being labeled in Grades 6-12.

Results of McLeskey and Grizzle's (1992) investigation revealed that approximately 58% of all students with learning disabilities in Indiana were retained before they were labeled. Approximately 54% of the third-grade students and 61% of the sixth-grade students had been retained. The study found that twice as many students with learning disabilities were being retained in Indiana as students without disabilities, suggesting that retention was being used in Indiana as a remedial measure before labeling a student with a learning disability. Furthermore, retention decisions seemed only to delay the identification of students as learning disabled for

approximately one year. Also, it appeared from the study that many students with learning disabilities, who are least likely to benefit from retention, were being retained. Thus, the authors concluded that students with learning disabilities do not benefit from retention.

Barnett, Clarizio, and Payette (1996) advanced the work of McLeskey and Grizzle (1992) by investigating the relationship of gender, socioeconomic status, and ethnicity among retained learning disabled (LD) children, as well as, the frequency and effects of repeat retentions among retained learning disabled children. Barnett et al. (1996) collected data on 344 Michigan students who were referred for special education due to learning problems during the 1990-91 school year. The participants were enrolled in Grades K-12 in six school districts. Approximately 50% of the participants were enrolled in urban schools and the other 50% were enrolled in suburban or rural schools. Of the 344 students participating in the study, 229 were retained at least once before being referred to special education, 114 had never been retained, 166 had been retained once, 60 had been retained twice, and 3 students had been held back more than three times before being referred to special education. In addition, from the sample of 344 students, 201 were classified as LD. Of the 201 LD students, 144 were retained at least once before being identified as LD, while 57 LD students were never retained. These data reiterated McLeskey and Grizzle's (1992) findings that students with learning disabilities were being retained before being referred to special education. Also, Barnett et al. (1996) found that although retention among LD students did not appear to be related to gender or grade placement, retained LD students usually exhibited lower levels of intelligence and weaker skills in reading comprehension, writing, and math at the time of their referral, strengthening McLeskey and Grizzle's (1992) argument that students with learning disabilities were unlikely to profit from retention and should not be retained.

Teachers' Beliefs

Students who are promoted from one year to the next are largely determined by the school system's promotional policies and by the attitudes and beliefs of teachers and administrators regarding the benefits of retention (Roderick, 1994). For this reason, Tomchin and Impara (1992) designed a study to examine teachers' beliefs about retention in an effort to better understand retention practices. Teachers commonly believe that retaining children in early grades is not as harmful as retaining them in later grades (Okpala, 2007; Witmer, Hoffman, & Nottis, 2004). Therefore, Tomchin and Impara (1992) focused on why teachers retain students in Grades K-7, emphasizing patterns and distinctions between retentions in Grades K-3 and Grades 4-7. The participants in the study were 135 classroom teachers from one rural school system in a southern state.

Results of a 20-item retention belief questionnaire indicated that teachers at all grade levels believed retention to be an acceptable school practice that prevents students from facing daily failure and motivates them to work harder. Teachers agreed that retention was not harmful in Grades K-3, but disagreed about the impact on students in Grades 4-7. Teachers of Grades 4-7 were less likely to retain students due to their belief that retention was more harmful to a child's self-concept, was more likely to permanently label a child, and was less likely to give immature students a chance to catch up. Results identified student academic performance, maturity, ability, gender, and age as factors that influenced retention decisions. However, the importance of factors differs among teachers. Teachers' beliefs about their students, as well as, beliefs about their own responsibilities for student progress provide the context for understanding the retention decisions of teachers (Tomchin & Impara, 1992).

Grade Retention Research 2000-2010

Longitudinal Studies

Ferguson, Jimerson, and Dalton (2001) investigated reasons connected with longitudinal academic and behavioral outcomes of 106 students followed from kindergarten through eleventh grade. Students were originally classified into one of four groups reflecting their educational experience concerning grade retention, but these groups were eventually collapsed into two groups referred to as retained and promoted. The study examined the family and individual characteristics of successful and unsuccessful retained students by including both between-group and with-in group effects of academic and behavior outcomes.

Results of the study indicated that background characteristics of retained students (i.e., lower SES, lower levels of mother's education, lower parental value of education, poorer initial personal-social skills, students' older age) were found to be risk factors associated with their lack of early school readiness, as well as, their subsequent poor academic progress. These same factors appeared to be associated with long-term behavioral problems following early academic failure and grade retention (Ferguson et al., 2001).

Silberglitt, Appleton, Burns, and Jimerson (2006), in their longitudinal study, examined the reading growth trajectories of 147 students from first through eighth grade from five districts in rural and suburban Minnesota. The study compared short-term and long-term effects between children retained one year, a matched group of children not retained, and a third group of randomly selected children. Data were compiled over the school years from 1996-97 to 2003-04. Silberglitt, Appleton, et al. (2006) speculated that (a) children retained would exhibit learning growth rates in reading during their second year in the same grade level (retained year) that was equivalent to the initial growth rate in that grade level, (b) children retained would exhibit

longitudinal reading growth rates that were similar to a matched group of promoted students, and (c) children retained would exhibit longitudinal reading growth rates that fell below a group of randomly selected students. Moreover, the study attempted to address several previously noted inadequacies of retention research such as lack of a control/comparison group (Holmes, 1989), poorly defined variables (Jimerson, 2001), and lack of random selection or assignment (Jackson, 1975).

The results revealed that retained students did not experience a benefit in their growth rate (relative to either the preceding year, or to the similarly performing but promoted students), and made less progress compared to the randomly selected group of students. The results of the study were consistent with prior research examining the effectiveness of grade retention as an academic intervention. The authors noted that the empirical evidence did not support the use of grade retention, and considering the expense of students repeating a grade, the lack of positive effects found in the study, and the negative long-term outcomes reported in related research, it is disturbing that the practice of retention continues (Silberglitt, Appleton, et al., 2006).

While research examining grade retention has provided evidence that it is an ineffective and harmful practice, supporters of grade retention often advocate that retention in the early elementary grades (K-2) is the defensible exception. For this reason, Silberglitt, Jimerson, Burns, and Appleton (2006) examined the reading growth trajectories of 49 retained students from first through eighth grade. The students were assigned to one of two groups based on grade of retention, with 27 students who were early retained (Grades K-2) and 22 who were later retained (Grades 3-5).

Analyses of the data in the study revealed that the growth trajectories of students retained early (K-2) were comparable to those retained later (3-5). Thus, retention did not produce

advantages in reading from first to eighth grade. Although the idea of early retention is appealing, the data failed to support retention, whether it was early or later in elementary school. The authors stated, “Proponents of grade retention base their position on common sense that repeating the year will enhance student achievement, but research reveals that the practice of grade retention is nonsense” (Silberglitt, Jimerson, et al., 2006, p.140).

Burkam, LoGerfo, Ready, and Lee (2007) used the Early Childhood Longitudinal Study-Kindergarten cohort (ECLS-K) to investigate nationwide patterns addressing who repeats kindergarten, and the subsequent cognitive effects of the event. The authors investigated first-time kindergarteners who were promoted, first-time kindergarteners who were retained, and children who were already repeating kindergarten. The analytic sample in the study consisted of 12,780 children in 915 schools, from the 1998-1999 kindergarten cohort.

The study suggested that boys, children from low socio-economical backgrounds, and children who entered kindergarten younger than average age are consistently at risk for repeating kindergarten. The evidence also suggested that repeating kindergarten seldom leads to cognitive benefits in literacy or mathematics performance. On average, kindergarten repeaters continue to perform below their peers in terms of literacy skills both at the end of kindergarten and at the end of first grade. Burkam et al. (2007) concluded that most children appeared to receive little or no cognitive benefit from repeating kindergarten, suggesting the need for a careful review of present retention practices.

In another longitudinal review of the ECLS-K, Hong and Yu (2007) researched two questions. First, will kindergarten retainees recover their lost ground and excel in the long run? Second, what are the effects of first-grade retention? The authors, using recently released waves of data investigated these questions. According to the results of the study, kindergarten retention,

relative to promotion, showed negative effects in both reading and mathematics at the end of the treatment year, but faded considerably by fifth grade. Throughout the elementary years, the kindergarten retainees and the first-grade retainees never achieved more on average than they would have if they had been promoted instead. Evidence indicated that most children primarily at risk for repeating kindergarten would have been competent of learning first-grade reading and math content if promoted. In addition, first-grade retention shows negative effects that stay almost constant from one year after treatment to three years later. Many children retained in first grade, if promoted to second grade instead, would have benefited from second-grade reading and math instruction and exhibited growth similar with those of their same-cohort promoted peers throughout the elementary years. The authors found no evidence that early-grade retention benefits retainees' reading and math scholarship toward the end of the elementary years (Hong & Yu, 2007).

Jimerson and Ferguson (2007) explored factors with longitudinal academic and behavioral outcomes of 137 students identified in kindergarten through second grade, of which 72 were followed through eleventh grade. All students were from a single school district in a western region community. Students were classified into one of four categories reflecting their educational experience regarding grade retention (i.e., early grade retainees; students retained in a transition classroom; students recommended for transitional placement but promoted; regularly promoted students, through eleventh grade). All children in the study were White, received no English as a Second Language (ESL) services, had no prior handicapping condition, had no prior retentions, no prior delays in enrollment in kindergarten, and received no special education services before first grade. The 12-year study examined the following research questions:

1. What is the association of grade retention and academic achievement during adolescence?
2. What is the association of grade retention and aggression during adolescence?
3. What is the association between grade retention and dropping out of high school?

Results of the study failed to demonstrate the effectiveness of grade retention on academic achievement. The effect sizes for the analyses of achievement were each within the small to moderate range (0.11-0.22). Analyses of behavior suggested that retained students displayed more aggression during adolescence. Effect sizes for the analysis of aggression were in the small range (0.09). In addition, the group of students recommended for transitional placement but promoted, were equal to the control group on all achievement and behavior measures during high school. The study revealed that 19% of the retained students dropped out of high school by the end of the eleventh grade (Jimerson & Ferguson, 2007).

Gleason, Kwok, and Hughes (2007) conducted a study for the purpose of investigating the one-year longitudinal effect of retention in first grade on peer acceptance. The authors predicted that retained children would be perceived by peers and teachers as exhibiting higher academic competence in year two, relative to promoted children. In addition, Gleason et al. (2007) expected that the positive effect of grade retention on peer acceptance would be reflected by children's ability to meet academic challenges in their classrooms. Participants were 350 ethnically diverse and academically at-risk students who attended school in Texas. Data were collected in first grade and one year later, at which time 63 students were repeating first grade and 287 were in second grade.

Findings from the study revealed that children who repeated first grade improved more in peer acceptance the repeated year than did children who were promoted to second grade. First-

grade teachers rated children subsequently promoted as more engaged and as achieving more than retained children. However, repeaters made substantial progress the following year. In addition, the retained child was viewed by classmates and the teacher as more academically capable during the repeated year. The study's results suggested that repeaters benefited from the extra year in grade in terms of classroom competence and peer acceptance (Gleason et al., 2007).

The purpose of kindergarten is not primarily education, but to introduce children to a school's social environment and acclimate them to all the activities involved in attending and learning in an institutional setting. For this reason, retention is likely to differ from retention in other grades in a couple of ways. First, kindergarten retention targets children who are socially immature or having difficulty acquiring basic academic skills (Mantzicopoulos & Morrison, 1992), while upper grade retention is based solely or largely on academic performance. Second, kindergarten retention is administered before any real failure occurs, making it more likely to have positive effects.

Dong (2010), based on data from the Early Childhood Longitudinal Study-Kindergarten cohort (ECLS-K) 1998-1999, estimated the underlying effect of repeating kindergarten on the retained children's academic performance in later grades. Children in the study were either first-time kindergarteners or kindergarten retainees in the 1998-1999 school year. The sample consisted of 8,672 children, of which 281 were retained. The major research question in the study was whether the retained children actually performed better than they would have done, had they been socially promoted.

Findings from the study showed that repeating kindergarten had positive, but diminishing, effects on the retained children's later academic performance. Math test scores were still significant up to third grade while the effect on the reading test scores were not. Results

should encourage researchers, education professionals, and legislators to take a more positive attitude regarding kindergarten retention. Kindergarten retention may give lagging children a chance to make up, if not catch up (Dong, 2010).

Using National Education Longitudinal Study (NELS) data, Griffith, Lloyd, Lane, and Tankersley (2010) examined the reading growth rates of secondary students who were retained between kindergarten and eighth grade and those of a matched comparison group who had never been retained. Each group contained 878 students.

Griffith et al. (2010) found a significant difference between the scores of the retained and promoted groups on reading achievement in the eighth grade, with the discrepancy favoring the promoted group. In addition, the results of the study revealed that the promoted students could expect to correctly answer 3.33% more reading questions at tenth and twelfth grades related to 2.44% for the retained group. However, the study's findings cannot be used to argue that grade retention was the cause of inadequate reading performance, or that retention was an ineffective practice, because the reasons for the grade retention and information on the quality of instruction that the students received were unavailable.

Longitudinal Studies Outside United States

Pagani, Tremblay, Vitaro, Boulerice, and McDuff (2001) examined the effects of grade retention on children's academic performance and behavioral development. They also addressed whether early grade retention was beneficial for children. Data for the study were obtained from the Quebec Longitudinal Study of Kindergarten Children. A sample of 1,830 children who attended kindergarten in French-speaking public elementary schools in all 11 regions of the Canadian province of Quebec, who met the criteria for the study, participated. During the course of data collection, children were followed from the end of kindergarten to the beginning of

adolescence. Methodologically, retained children's academic performance and behavior adjustment were rated by their teachers from year-to-year and compared to their same-grade peers. Statistically, retained children were compared to their same-age promoted peers.

Results of the study indicated both a short- and long-term negative influence on academic performance for boys and girls. Also, children's fearful, distracted, and disorderly behaviors persisted and, in some cases, worsened after grade retention. These forthcoming associations were long lasting and more prominent when grade retention occurred early in primary school. Boys were more susceptible to the negative impact of grade retention on academic performance and disruptive behavior in the classroom. To the contrary, unruly behavior in girls was comparatively less associated with long-term consequences in boys. Nevertheless, girls experienced both short- and long-term academic performance problems in the aftermath of grade retention. The authors recommended the search for strategies, other than grade retention, in order to battle children's deteriorating school performance since retention does not seem to profit children (Pagani et al., 2001).

In another Canadian study, participants from Manitoba were examined over the period from 1997-1998 to 2002-2003. Guèvremont, Roos, and Brownell (2007) addressed the three following research questions:

1. What are the characteristics of retained students?
2. How do students retained in third grade perform on a standards test in two consecutive years?
3. How does retention influence high school withdrawal?

Data from the study showed that students who were male, young for grade, in the first, second, seventh, or eighth grade, lived in a low or mid-low SES area, changed schools

frequently, had younger mothers, and their family had received income assistance were most likely to be retained. Also, data from third grade standards tests taken in successive years indicated that only 25% of the retained students improved their score after repeating third grade. In addition, retention was a strong forecaster of high school withdrawal rates, with odds of student departure three times higher for once retained students and eight times higher for students retained more than once. Retention does not solve the problems of children struggling in school. Students who are not achieving need extra support and merely having them repeat the same grade does not offer the support they need (Guèvremont et al., 2007).

In order to assess the determinants and effects of grade retention, Bonvin, Bless, and Schuepbach (2008) conducted an investigation based on a matched-samples longitudinal design involving second-grade primary school children. The authors chose second grade because it was the school level where retention was applied the most frequently. Participants were composed of 83 pairs of matched pupils from German- and French- speaking Swiss schools. Forty-two matched pairs were German-speaking, and 41 matched pairs were French-speaking.

Results showed that the decision for grade retention did not rest only on the students' actual academic performance but also on the teachers' attitudes and evaluations. With regard to improvement in learning, the study produced contrasting short- and medium-term results. They were positive in the case of same-grade comparisons and negative when same-age comparisons were applied. The authors suggested that grade retention be avoided at the primary school level. Worldwide, grade retention effectiveness is unacceptable when one considers its long-term consequences (Bonvin et al., 2008).

Martin (2009) examined the relative importance of age within cohort, grade retention, and delayed school entry in 3,684 high school students' academic motivation, engagement, and

performance from seven Australian high schools. Five percent of the participants in the study reported having to repeat a grade at some phase in their education with 47% of the retentions occurring in the first three years of elementary school and another 42% occurring in the final four years of elementary school.

The study's findings showed more significant effects such that clearly older-for-cohort students were higher in disengagement, lower in positive intentions, lower in homework completion, and lower in performance scores. Meanwhile, younger-for-cohort students scored higher in valuing school, higher in positive intentions, higher in attendance, higher in homework completion, and higher in performance. Data suggested that there was little to no academic advantage to being noticeably older for cohort (delaying entry) or to being retained in a grade. The effects of grade retention were also found to be negative regardless of the grade in which a student was retained. The author concluded that even if a child was young for his or her cohort or was male, it was best for the child to remain with the cohort rather than be retained in that grade for another year (Martin, 2009).

Characteristics of Retained Students

According to Beebe-Frankenberger, Bocian, MacMillan, and Gresham (2004), little is known about the kind of child who tends to be retained relative to another child with comparable low rates of academic achievement but who is promoted or identified with high-occurrence disabilities and receives special education support services. Therefore, the authors sought to clarify who is retained by asking two questions. First, do retained students differ on academic and social-behavioral factors from peers who are in special education, placed at risk for retention, or normally promoted to the next grade? Second, how do students who were retained at the end of second grade differ from all other students in the sample, but received alternative

educational treatments (i.e., special education, notification of being at risk)? The 224 participants in the study were enrolled in second grade during the 1999-2000 school year, at which time a decision was made to promote or retain each student. When first-year data were collected in 2000-2001, 64 participants were second graders (retained group) and 160 were third graders (promoted group).

Beebe-Frankenberger et al. (2004) suggested that most children with academic shortfalls were identified by schools early and organized into educational treatments varying in intensity based on ability. Results indicated that students who were retained or received special education services scored lower than promoted students on measures of IQ, overall academic competence, and reading skills. However, finding out that retained students and special education students closely resembled each other, particularly on IQ, was unexpected. In addition, results showed reliable differences between promoted students and all students experiencing academic difficulties, regardless of treatment. Moreover, students in the retained group were absent twice as often as the promoted students, had attended more elementary schools by the end of second grade, and were more likely to be from a low income family. Also, the data do not suggest that retained students were significantly different in ability (IQ) when compared to other students with academic problems who received other treatments.

Willson and Hughes (2009) examined child, classroom, and family variables in order to determine which ones contributed to decisions to retain a student in first grade. A sample of 784 children with low literacy performance in kindergarten, or at the beginning of first grade, was assessed in academic competence, socio-demographic characteristics, social/emotional characteristics, school context, and home environment. However, only 769 participants were

included in the analyses due to 15 students leaving the school system before information on retention or promotion could be attained. Of the 769 students, 165 were retained in first grade.

Results found that the academic variables were significantly related to retention, and all indicated low performance for retained children. Economic disadvantage and age (socio-demographic characteristics) were both significantly related to retention, but gender was not. The ratio of retained to promoted students was larger for the economically disadvantaged, and retained children were younger than promoted students. Retained children exhibited low ego resilience and were rated more hyperactive (social/emotional characteristics). Teacher support (school context) was negatively related to retention. Also, parent communication with school was related to retention, as were parent positive perceptions of school and parental aspirations for the educational achievement of their child (home environment). The 165 students retained in first grade were found to differ from promoted students on reading and mathematics achievement test scores, teacher-rated engagement and achievement, and intelligence as individual predictors of academic competence (Willson & Hughes, 2009).

Retention and Minority Students

Willson and Hughes (2006) investigated predictors of being retained following first grade in an academically at-risk sample of Hispanic/Latino youth. Two hundred eighty-three Hispanic children with below the median literacy performance at entrance to first grade participated in the study. Participants attended one of three school districts in Texas. Fifty-one of the participants were retained in first grade with low literacy skills, being young at entrance to first grade, low ego resilience, low support in the teacher-student relationship, and parents' low sense of responsibility for their children's educational outcomes predicting retention decisions.

Results indicated that controlling for literacy, only being young for grade, and parents' low sense of responsibility for their children's adjustment to school made a direct contribution to retention. In addition, early literacy skills were higher for children enrolled in bilingual classrooms than for children in non-bilingual classrooms (Willson & Hughes, 2006).

To assess whether making academically challenged students repeat a grade was an effective remediation practice, Lorence and Dworkin (2006) followed students who failed the Texas mandatory third-grade reading test through their sophomore year in high school. The study only examined non-Hispanic White, Hispanic, and Black students. After these exclusions, students failing the 1994 state reading exam totaled 38,445, of which, 1,244 were required to repeat third grade.

The findings of the study revealed that requiring low-performing Black and Hispanic students to repeat third grade was associated with increased reading performance that persisted over time. In addition, retention in third grade benefited low-performing readers regardless of race. After controlling for initial difference between the retained and socially promoted students, the advantage in academic performance became more prominent among the retained children. Making students repeat a grade, when enhanced with additional educational assistance can benefit academically challenged children (Lorence & Dworkin, 2006).

Teachers' Beliefs

Witmer, Hoffman, and Nottis (2004) explored elementary teachers' beliefs, knowledge, and practice related to grade retention using an adapted version of the Teacher Retention Beliefs Questionnaire (Tomchin & Impara, 1992). Witmer et al. (2004) added a knowledge section to the original questionnaire to measure teachers' propositional knowledge of retention. Thirty-five, K-4 teachers from a rural school district in the northeastern part of the United States took part in the

study. Twenty-one of the teachers were assigned in the K-2 group, and 14 teachers were assigned in the 3-4 group. Thirty-seven percent of the teachers had 1 to 10 years of teaching experience, 17% between 11 and 20 years, 37% between 21 and 29 years, and 8% of the teachers had more than 30 years of teaching experience. Sixty percent of the teachers reported that they held a Master's degree in education.

Results indicated that K-4 teachers in this study believed that retention was an acceptable practice for preventing failure. For example, 77% of the teachers believed retention was effective. In fact, teachers said that a student's academic performance was the most influential factor in their decision to retain. Significant differences between kindergarten through second grade and third through fourth grade teachers were found on responses to four of the 20 questions on the questionnaire. More kindergarten through second grade teachers than third through fourth grade teachers disagreed with the statements that retention is necessary to maintain grade level standards, pupils who do not apply themselves in their studies should be retained, and over-age children in kindergarten through second grade cause more behavioral problems than other children. On the other hand, more kindergarten through second grade teachers than third through fourth grade teachers agreed that students should only be retained once in elementary school. Results also indicated that teachers' knowledge about the effects and outcomes of retention were low, regardless of grade taught. Forty-four percent of the teachers indicated that their source of knowledge about retention came from personal experiences while 22% gained their knowledge by talking to other teachers. Nine percent of the teachers reported that their knowledge about retention came from reading journals or attending workshops on retention. However, no significant correlation between teachers' knowledge and retention practice was found (Witmer et al., 2004).

In kindergarten, teachers are the ones who recommend students for retention. Therefore, it is important to better understand their beliefs and perceptions on kindergarten retention as an intervention. Okpala (2007) conducted such a study examining the perceptions of 37 kindergarten teachers who were randomly selected from public elementary schools in a southeastern school district in North Carolina. The major goals of the study were to: (1) identify teachers' perceptions of retention, (2) examine differences in reported perceptions of retention by certified versus non-certified teachers, and (3) investigate teachers' perceptions of retention as a function of teaching experience.

Okpala's (2007) study found that kindergarten teachers perceived retention as a necessary intervention tool in the larger context of increased accountability and education reform. The participants in the study perceived academic ability (87%), attendance (78%), social maturity (74%), emotional maturity (73%), and physical maturity (67%) to be the major reasons for kindergarten retention. Also, results showed the kindergarten teachers' perceptions of retention appeared to vary as a function of teaching experience and between certified and non-certified teachers. Teachers with less than five years of experience perceived the benefits of kindergarten retention to be greater than did teachers with more than five years of experience while non-certified teachers placed a higher value on the benefit of retention than did certified teachers (Okpala, 2007).

Alternatives to Retention

Participation in a transitional grade is thought to prepare unprepared children for first grade. Therefore, Mantzicopoulos (2003) examined the effectiveness of a developmental first-grade transition program by following the progress from kindergarten through third grade of 34

children placed in a transition room program and 37 transition-room eligible (but not placed) children.

By the end of the three-year study, results indicated that participating in a developmental program before first grade seemed to have some academic and socio-emotional advantages for the recommended-and-enrolled group over peers who were recommended for the program but not enrolled due to parent refusal. However, effect sizes declined as the study progressed. It is not clear if the findings of the study supported the transition room policy. The program was child centered, individualized, provided opportunities for discovery, and included parents as partners in their children's education. Therefore, the positive effects seen by the recommended-and-enrolled group may have been the result of the adaptations that were made to the educational environment, rather than the extra year of schooling (Mantzicopoulos, 2003).

As school districts impose tougher standards on students and increasingly hold schools accountable for their performance, there will be a growing need to find effective remedial education programs to help underachieving students. Jacobs and Lefgren (2004) found evidence in their study that summer school and grade retention had a small but positive overall impact on student achievement scores for third-grade students. This overall effect was due to a combination of benefits from both summer school and grade retention. Unlike many studies, the authors found that retention may actually increase academic achievement for low-achieving third graders. However, summer school and retention appeared to have little, if any, effects for sixth-grade students. Thus, remedial summer school and retention programs under favorable circumstances can improve the performance of young disadvantaged students.

The best solution to the retention versus social promotion debate is to prevent academic failure before it occurs. Thus, the purpose of Picklo and Christenson's (2005) study was to

determine whether differences existed in the availability of instructional options for at-risk students and to determine differences in the use of these options as a function of retention practices, awards and sanctions, and school level. Participants in the study were 242 teachers and school psychologists from 99 elementary and middle or junior high schools across 19 states. Of the respondents, 86% indicated that their school retained students.

The survey results from Picklo and Christenson (2005) revealed that cooperative learning strategies and group work were used more often to assist struggling students, regardless of school characteristics. Also, elementary schools implemented a greater variety of instructional options with more frequency than middle or junior high schools. In order for struggling students to improve, teachers must increase their use of instructional strategies to assist these students. The study indicated that instructional options were not being used consistently with struggling students. In addition, teachers needed to increase the variety of interventions used to achieve a better instructional match for struggling students. The authors pointed out that it was important for instructional options to be available, but it is more important how those options are delivered. School personnel may not be implementing evidence-based interventions (i.e., those that demonstrate effectiveness).

Abbott, Wills, Greenwood, Kamps, Heitzman-Powell, and Selig (2010) examined the effect of the amount of literacy instruction on the academic literacy outcome effects of retention among kindergarten and first-grade students versus their promoted low-achieving peers. The study matched 15 kindergarten and first-grade retained students in seven schools with their promoted peers on grade-level literacy performance. There were a total of 30 students in the sample.

Retained kindergarten students received less intervention and did not benefit academically from retention. On the other hand, promoted first-grade students who received additional small-group interventions showed a nearly significant effect. The findings of the study suggested that, even with good-quality instruction, in order for retained students to accelerate reading progress beyond the levels of their low-achieving promoted peers they must have increased levels of academic time and intensity of literacy instruction. Approximately two and one-half hours per school day of general education and small-group intervention literacy instruction is needed to bring students within average range. Thus, it is unwise to return students to the same insufficient academic environment that failed them in the first place (Abbott et al., 2010).

Accountability and Policy

In the era of No Child Left Behind, many school systems in the United States are struggling with motivating students to achieve, while at the same time addressing the needs of students who persistently struggle. In the mid-1990s, the Chicago Public Schools declared an end to social promotion and instituted promotional requirements based on standardized test scores in the third, sixth, and eighth grades. Roderick and Nagaoka (2005) found that students who were retained under Chicago's high-stakes testing policy continued to struggle during their retained year and faced significantly increased rates of special education placement. Among third graders, there was no evidence that retention led to greater achievement growth two years after the promotional decision. Among sixth graders, there was no evidence that retention was associated with lower achievement growth. Therefore, the authors concluded that retaining low-achieving students did not help.

By analyzing data from the Early Childhood Longitudinal Study-Kindergarten cohort (ESLS-K) with the technique of multi-level propensity score stratification, Hong and Raudenbush (2005) informed policy makers and educational practitioners with new evidence regarding the effects of school-level retention policies at the kindergarten level. They suggested that the richness of the ECLS-K data and the propensity score-based adjustment method enabled them to achieve more comprehensive control of the pre-treatment differences than previously possible. The total sample included more than 20,000 kindergarteners. However, the analytic sample included 471 kindergarten retainees and 10,255 promoted students in 1,080 retention schools, and 1,117 promoted students in 141 non-retention schools.

According to the Hong and Raudenbush (2005) analytic results, the average effect of the kindergarten retention policy, as compared to a policy that banned retention, was null or very small. On the other hand, no evidence was found that a policy banning retention would benefit those children who would be promoted if such a policy were adopted, casting doubt that a policy on grade retention in kindergarten would improve instruction. However, the study found evidence, in both reading and math, that children who were retained would have learned more had they been promoted. After being retained for one year, retainees experienced about two-thirds of a standard deviation loss in academic growth in both subject areas. The authors concluded that kindergarten retention left most retainees even further behind and hindered their cognitive development over the repetition year. At-risk children promoted to the next grade level seemed to have a better chance of growth acceleration.

Summary

This chapter provided an overview of the research on grade retention and student academic achievement for elementary students. The chapter gave a brief history of grade

retention by examining one early review of grade retention research and four prior meta-analyses. Next, the chapter focused on grade retention research between 1990-1999, then between 2000-2010. Reviews were discussed in chronological order to better understand the development of the field. Chapter III provides an overview of the research design and describes the steps that were taken to conduct this study.

CHAPTER III

METHODOLOGY

Research Design and Methodology

This study examined the evidence between grade retention and students' academic achievement in the elementary grades. The study provided a quantitative summary (meta-analysis) of findings from available studies published between 1990 and 2010 that examined the effects of grade retention of K-6 students on student achievement.

Problem and Purposes Overview

Grade retention is one of the few topics in education where it is challenging to find research studies showing a positive effect, and the few research studies that do exist have very small effect sizes. In general, there are negative results for students who are retained, and there are more positive long-term results for promoted students than for retained students. This is accurate even when students are matched for achievement at the time of the decision to retain or promote. Retention has been found to have negative consequences on academic achievement in language arts, reading, mathematics, social studies, and grade point average. In addition, promoted students scored higher than retained students on social and emotional adjustment, behavior, self-concept, and attitude towards school (Hattie, 2009).

Research Questions and Hypotheses

Research question 1: What is the research consensus on the effects of grade retention on academic performance of students?

Research question 2: If initial gains are evident in academic achievement from repeating a grade in elementary school, does the research show that they persist in high school?

Research question 3: Does the research show that retaining a student in kindergarten improves the academic performance of the student?

Research question 4: Does the research show that retaining students early (Grades 1-2) benefits students' academic performance more than retaining them later (Grades 3-6)?

The general expression of the null hypotheses is to test the significance of an average effect size. To achieve the objectives of finding out the magnitude of the effect of grade retention on students' academic achievement, four research questions and two generally expressed null hypotheses were examined.

Since different criteria have been used in the various studies to measure the effect of grade retention on academic achievement, the data set was expected to be heterogeneous. The data from the studies were grouped by the number of years post-retention (2 years or less, 3 to 5 years, over 5 years) to reach homogeneity and produce an average effect size (*ES*) for each group.

Cohen (1977) provided standards against which to evaluate an effect size. He used $d = 0.20$ (small effect), $d = 0.50$ (medium effect), and $d = 0.80$ (large effect) as rough guidelines. However, Cohen cautioned that it is better to obtain standards for comparison from the professional literature, rather than use arbitrary guidelines. Effect sizes of 0.20 or higher, in absolute value, have been interpreted as educationally meaningful in the literature on grade

retention (McCoy & Reynolds, 1999; Reynolds, 1992). However, Hattie (2009) suggested that $d = 0.20$, $d = 0.40$, and $d = 0.60$ represented small, medium, and large effects, respectively, when judging educational outcomes. Therefore, these guidelines will be used to evaluate effect sizes in this meta-analysis. In addition, Hattie stated that the average effect, or hinge-point, of 0.40 should be used as the benchmark to judge effects in education. The hinge-point does not mean that all effects below 0.40 are not worth having, it just means that these effects need more consideration (i.e., cost of intervention, intervention effects). The effect size of 0.40 sets a level where the effects of innovation enhance achievement in such a way that real-world differences can be noticed and student change can be seen. Thus, influences above the hinge-point ($d = 0.40$) fall into the “zone of desired effects” and have the greatest impact on student achievement outcomes (Hattie, 2009). It should be noted that the previously mentioned effect sizes are not magic numbers or cut-off points. They are only guidelines to begin discussion about whether the effects of an intervention were better for students than the alternative innovation.

A mean effect size and confidence interval were calculated for each study to test the null hypotheses to determine if the effect size was likely to be a true difference or whether it could have occurred by chance.

The first null hypothesis is: there is no significant ($\alpha = .05$) relationship between grade retention and academic performance of students.

The second null hypothesis is that there is no significant ($\alpha = .05$) difference in academic performance, over time, for students who receive grade retention as an intervention compared to those students who did not.

The study hypothesis is that there is a positive relationship between grade retention and students' academic performance. That is, grade retention is an effective intervention and produces its intended result.

General Research Design

A meta-analysis is a form of survey research in which research reports are surveyed for the purpose of integrating and summarizing the findings. In this study, a coding form was used to extract information consistently across studies, which were then analyzed and reported. In the 1970s, there was a paramount change in the method researchers used to examine literature. This change offered a means to curb the enormous amount of research data so that it could offer useful information for teachers. The prevalent method had always been to write a synthesis of several published studies in the form of a composed literature review. However, in 1976, Glass introduced the concept of meta-analysis. In a meta-analysis, the effects in each study, where suitable, are converted to a common measure (effect size), so that the overall effects can be quantified, interpreted, and compared. Meta-analysis quickly became popular and by the mid 1980s more than 100 meta-analyses in education were accessible (Hattie, 2009).

Despite their growing popularity, meta-analyses have not been free from criticism, nor have they been a panacea for resolving all the problems associated with building reliable and valid scientific knowledge and theory (Wolf, 1986). Glass, McGaw, and Smith (1981) grouped criticisms of meta-analyses into the following categories:

1. Logical conclusions cannot be drawn by comparing studies that include different measuring techniques, variables, and subjects (comparing apples to oranges).
2. Results of meta-analyses are un-interpretable because results of poorly designed studies are included with results from good studies.

3. Published research is biased in favor of significant findings because non-significant findings are rarely published.

4. Multiple results from the same study are often used making the results appear more reliable than they are.

It is very important for researchers to be aware of these potential problems and issues when reading published meta-analyses or designing their own.

Data Collection

For this study, a literature search was conducted using relevant computerized databases (e.g., Education Full Text, Educational Resources Information Center (ERIC), and PsychINFO) using subject terms such as *grade retention*, *grade repetition*, *retained*, *held back*, *failure*, *flunking*, and *non-promotion*. Searches were limited to studies written in English. Relevant articles published in peer-reviewed journals relating to grade retention and the effects on the academic achievement of K-6 students were identified, as well as a few articles from non-peer reviewed journals and books, which were used for background information. Published and unpublished conference papers and doctoral dissertations were excluded from this study. References in the articles retrieved, including studies, were also reviewed to identify relevant articles for inclusion in the meta-analysis. Since all the data for this study came from published research, this study was not considered to be research involving the use of human subjects. Therefore, The University of Alabama's Internal Review Board (IRB) determined that IRB approval for this study was excluded from their review (see Appendix A).

Hundreds of abstracts or citations were reviewed to identify articles that might meet the screening criteria for inclusion in the meta-analysis or which would provide relevant background information. This screening included several articles that were identified through review articles

or other studies relevant to grade retention and academic achievement. The initial screening of abstracts and reference lists from available articles resulted in identification of approximately 120 articles reviewed in greater depth for this study. A total of 68 articles were identified as potential studies for inclusion in the meta-analysis. The following criteria were used in screening articles for inclusion in the meta-analysis:

1. Studies reported at least one quantitative outcome measure of grade retention.
2. Studies involved K-6 students at time of retention.
3. Studies reported sufficient statistical information to allow calculation of effect sizes (F or t values, means, standard deviations, or some information regarding significance).
4. Studies included a comparison group.
5. Studies had a sample size of ten or more.
6. Studies included original data.
7. Studies did not duplicate previously reported outcomes. No “double-counting.”
8. Studies were relevant to this study’s research questions.

The random assignment of students to the treatment of retention is not feasible or ethical in many social contexts. Therefore, studies cannot use a randomized experimental design.

Coding of Studies

Each study that met the inclusion criteria was systematically coded for multiple variables. Coding typically identifies the characteristics of participants, research design and measurement procedures, and findings. An initial screening instrument was developed by the researcher to capture key information from each study (see Appendix B). Statistical data from all relevant studies were then entered into an EXCEL® spreadsheet which was used to calculate effect sizes.

A total of 43 studies were screened and coded. Coding procedures included the name(s) of each study's author(s), the year of publication, the study design, the grade(s) in which the participants were retained, the outcome grade(s), the dependent variables (academic achievement outcomes), the number of participants in each sample, various statistical data for every academic outcome measure in each study (converted to Cohen's *d*), and an average effect size for each study. Of these 43 studies, 31 studies had sufficient data to calculate effect size. Table 1 lists the 31 studies that were analyzed in the meta-analysis.

Table 1

Studies Included in the Meta-Analysis

Author(s)	Study Design	# Academic Outcomes
Beebe-Frankenberger et al. (2004) +	Independent	10
Bonvin et al. (2008) +	Paired	16
Burkam et al. (2007) +	Independent	18
Cosden et al. (1995) +	Independent	6
Dennebaum & Kulberg (1994)	Independent	14
Dong (2010) +	Independent	16
Gleason et al. (2007)	Independent	4
Griffith et al. (2010) +	Matched	1
Hong & Raudenbush (2005)	Independent	28
Hong & Yu (2007) +	Independent	10
Jacob & Lefgren (2004)	Independent	8
Jimerson (1999) +	Independent	3
Jimerson et al. (1997)	Independent	15
Jimerson & Ferguson (2007) +	Independent	7
Johnson et al. (1990)	Independent	7
Kundert et al. (1995) +	Independent	3
Lorence & Dworkin (2006)	Independent	21
Mantzicopoulos (2003)	Independent	9
McCoy & Reynolds (1999)	Independent	10
McLeskey & Grizzle (1992) +	Independent	6
Meisels & Liaw (1993)	Independent	8
Moller et al. (2006)	Independent	2
Phelps et al. (1992)	Independent/Matched	5
Pierson & Connell (1992)	Independent	10

(Continued)

Table 1 (Continued)

Studies Included in the Meta-Analysis

Author(s)	Study Design	# Academic Outcomes
Reynolds (1992)	Independent	10
Roderick & Nagaoka (2005)	Independent	12
Silberglitt, Appleton, Burns, & Jimerson (2006) +	Independent/Matched	2
Silberglitt, Jimerson, Burns, & Appleton (2006) +	Independent	1
Willson & Hughes (2009) +	Independent	2
Wu et al. (2008a) +	Paired	8
Wu et al. (2008b) +	Paired	2

Note. + denotes studies not included in any prior meta-analysis. Total outcomes equal 274.

Summary

This chapter provided an overview of the methodology used in this meta-analysis. An explanation of the study's research questions, the null hypotheses, and the general research design of the study were explored. The section on data collection described the methods and search criteria used for conducting the systematic literature search. In addition, an explanation was provided regarding the initial and secondary screening processes. The chapter concluded with a list of all 31 studies included in the meta-analysis indicating study design and number of academic outcomes.

The next chapter, Chapter IV, describes the specific findings of the literature synthesis and meta-analytic calculations. Depending on the data provided in each study, various equations were used to calculate effect sizes. Therefore, the equations were presented and explained in Chapter IV. In addition, calculations of significance (confidence intervals) were also reported.

CHAPTER IV

RESULTS

Introduction

This study examined the most current research between grade retention and academic achievement for kindergarten through sixth-grade elementary students. A comprehensive meta-analysis of available scientific studies published in peer-reviewed journals between 1990 and 2010 was conducted. This chapter examined the specific findings from the comprehensive review of the studies included in the meta-analysis.

Data Set for Meta-Analysis

The systematic review of the literature identified 68 articles as potential studies for inclusion in the meta-analysis. These 68 studies were narrowed to 43 studies, eliminating studies that did not involve K-6 students and/or some measure of academic achievement. The 43 remaining studies were then narrowed to 35 studies because eight of the studies (Abbott et al., 2010; Ferguson et al., 2001; Guèvremont et al., 2007; Hong & Raudenbush, 2006; Martin, 2009; Pianta et al., 1997; Willson & Hughes, 2006; Wu et al., 2010) either lacked the necessary data to calculate effect size, or the studies did not specifically relate to this study's research questions. From these 35 studies, four more studies (Mantzicopoulos, 1997; Mantzicopoulos & Morrison, 1992; McCombs-Thomas et al., 1992; Pagani et al., 2001) were omitted because they only reported the significant effects in their results and did not include sufficient data to calculate the non-significant effects. Since one of the criticisms of conducting a meta-analysis concerns the publication of only significant findings (Glass et al., 1981), studies that did not publish

non-significant effects, or include data to calculate non-significant effects, were not considered for this meta-analysis. Therefore, a total of 31 studies remained for inclusion in the meta-analysis, of which, 16 studies have never been used in any prior meta-analysis.

Study Procedures

A meta-analysis focuses on the magnitude of the effect observed in a study by using a metric called an effect size. An effect size permits meaningful numerical comparisons and analysis to be made across similar studies. An effect size measure is used to determine the amount of departure from the null hypothesis in standard units. In this meta-analysis, S-PLUS® version 6.2 Professional Edition for Windows® was used to calculate the effect size (Cohen's d) for each dependent variable related to student achievement in each study. Depending on the type of data provided in the study, S-PLUS® transformed the statistical data into an effect size using one of the following equations:

F-test

$$d = \frac{2\sqrt{F}}{\sqrt{df_{error}}} . \quad (1)$$

Where the degree of freedom error is the second degree of freedom listed.

t-test

$$d = \frac{2t}{\sqrt{df}} . \quad (2)$$

Where the degree of freedom is equal to $n - 2$.

r

$$d = \frac{2r}{\sqrt{1-r^2}} . \quad (3)$$

Chi-square

$$r = \sqrt{\frac{\chi^2}{n}} . \quad (4)$$

In addition, S-PLUS® totaled all d values and calculated an average effect size for each study using the following formula:

$$\bar{d} = \frac{\sum d}{n}. \quad (5)$$

For studies that included effect sizes in their findings, effect sizes were totaled and averaged to acquire an overall effect size for each of those studies by the researcher.

Once effect sizes were calculated for all studies, confidence intervals were calculated for each study to determine significance. The formula given by Hedges and Olkin (1985) for calculating the confidence interval or “margin of error” for an effect size is as follows:

$$\sigma[d] = \sqrt{\frac{N_E + N_C}{N_E \times N_C} + \frac{d^2}{2(N_E + N_C)}}. \quad (6)$$

Where N_E and N_C are the numbers in the experimental and control groups, respectively. The study’s effect size is represented by d .

After determining the “margin of error,” the confidence interval was calculated by multiplying the “margin of error,” also called the standard error, by the desired confidence level and adding the result to the effect size for the upper limit and subtracting the result from the effect size for the lower limit. A 95% confidence interval was used in this study.

To test for homogeneity and produce an average effect size for each group, the data from the studies were grouped by the number of years post-retention (2 years or less, 3 to 5 years, over 5 years), and an ANOVA analysis was conducted on the effect sizes for the three groups. A test of homogeneity of variance (Levene Statistic) revealed a significance of .736 (see Table 2). Since .736 is greater than .01, there was homogeneity of variance. Therefore, the data in this meta-analysis reached similar conclusions and can be combined with reasonable confidence.

Table 2

Homogeneity Test

Years Post-Retention	<i>N</i>	Average Effect Size	Levene Statistic
2 Years or Less	9	.45	
3 to 5 Years	5	.58	
Over 5 years	12	.51	
Total	26	.50	.310*

Note. * Levene statistic had a significance of .736.

Study Findings

Thirty-one studies reported sufficient data to compute effect sizes. The following section provides data on each of these studies and includes a discussion of the major findings. Table 3 indicates the studies published between 1990 and 1999 included in this meta-analysis. Also, the table includes each study's average effect size and confidence interval.

Table 3

Average Effect Size and Confidence Interval (1990-1999 Studies)

Study/Year	Average Effect Size	Standard Error	95% CI	
			LL	UL
Cosden et al. (1995)	0.77	0.17	0.09	0.75
Dennebaum & Kulberg (1994)	0.55	0.20	0.16	0.94
Jimerson (1999)	0.48	0.20	0.09	0.87
Jimerson et al. (1997) +	0.28	0.20	- 0.11	0.67
Johnson et al. (1990)	0.77	0.29	0.20	1.34
Kundert et al. (1995) +	0.05	0.18	- 0.13	0.23
McCoy & Reynolds (1999)	1.17	0.07	1.03	1.31
McLeskey & Grizzle (1992)	0.79	0.08	0.63	0.95
Meisels & Liaw (1993)	0.27	0.02	0.23	0.31
Phelps et al. (1992)	0.59	0.25	0.10	1.08
Pierson & Connell (1992)	0.60	0.14	0.33	0.87
Reynolds (1992)	0.64	0.07	0.50	0.78

Note. CI = confidence interval; LL = lower limit, UL = upper limit. + denotes study results were insignificant at 95% confidence interval.

1990-1999 Studies

Johnson et al. (1990) compared Metropolitan Achievement Test (MAT) scores from seven subtests of academic achievement for 20 retained students to 17 recommended, but not retained students, and 20 students who made normal progress. When the F values of all subtests were converted to Cohen's d (Equation 1) and averaged (Equation 5), an ES of 0.77 (large effect) was found with the largest effects coming from vocabulary, reading comprehension, total reading, and language arts (1.08, 0.91, 0.852, and 0.85 respectively). The math calculations subtest had a medium, yet "visible" effect size of 0.42 according to Hattie (2009).

Johnson et al. (1990) concluded from their investigation that the use of early grade retention was not an effective academic intervention for the retained students. However, the data did not support the practice of social promotion as being effective either. Both the retained and not retained students were significantly lower in overall academic achievement than the norm group of students.

McLeskey and Grizzle (1992) used either the Woodcock Johnson Test of Achievement (WJ) or the Wide Range Achievement Test-Revised (WRAT-R) to determine the academic achievement levels of 689 LD retained and LD not retained students. Using F values and other information available, an average ES of 0.79 was calculated. Achievement scores contributing to this effect size included WJ-Math (1.11), WRAT-Math (0.77), and WRAT-Reading (0.81). WJ-Reading had the lowest effect size (0.65) of all six achievement scores.

After covarying for IQ differences between the two groups, McLeskey and Grizzle (1992) found that the LD not retained students scored significantly higher on each of the reading, mathematics, and spelling/written language measures than the LD retained students. They also found that retention was being used as a remediation before labeling students with learning

disabilities, even though, students with learning disabilities are the least likely to benefit from retention.

In a longitudinal study, Pierson and Connell (1992) compared 74 retained students with 60 random and 69 matched-ability non-retained students, as well as, 35 socially promoted students. The variable used to measure academic performance was the mean of the students' standardized test performance scores and report grades, which were computed by the authors. *F* values of planned comparisons were computed between the retained group and each non-retained group on academic performance. These *F*-test results were used to calculate effect sizes for each of the following comparison groups: retained vs. matched ability, retained vs. socially promoted, and retained vs. random. The largest effect was found to be between the retained and random group with a very large effect size of 0.997. An effect size of 0.46 was found between the retained and socially promoted group, and the retained and matched-ability comparison revealed an effect size of 0.35. An average *ES* of 0.60 was calculated for the study.

Results of Pierson and Connell's (1992) study revealed that retained students showed significantly lower levels of academic performance than did the random sample of their classmates who had not been retained. In addition, retained students did not differ significantly from their classmates with the same academic ability, but the retained students did perform better than their socially promoted peers. Pierson and Connell suggested that retention in the early years of elementary school, in comparison with social promotion, was beneficial to academic performance and supported the use of retention as a potentially effective strategy for academic difficulty in the early elementary grades.

Reynolds (1992) investigated the effects of early grade retention on reading and mathematics achievement in Grade 4 for an analytic sample of 1,231 at-risk children.

Participants were children who enrolled in the school system in kindergarten and remained in the school system in fourth grade, thus, allowing the effects of retention to be traced up to three years post-retention. Analyses of a matched control group of 200 promoted children were also conducted. Reading achievement results revealed an *ES* of 0.75 for the total sample group and an *ES* of 0.80 for the matched control group. These almost identical results indicated that retained students performed 8 months lower than similar and equally low-scoring promoted children. In addition, mathematics achievement results revealed an *ES* of 0.68 for the total sample group and an *ES* of 0.74 for the matched control group, indicating a 6-month and a 7-month difference in mathematics achievement for the retained children, respectively.

Reynolds' (1992) findings indicated that retained children declined in reading performance relative to promoted children after being retained. However, effects on math achievement exhibited no such pattern of decline over time. The study found that retention had substantially negative effects on cognitive achievement in reading and mathematics and does not support grade retention as an educational practice with at-risk children. An average *ES* of 0.64 was calculated for the study.

In a longitudinal study conducted by Phelps et al. (1992), the long-term effects of retention and transition placement on academic achievement were examined. The majority of the participants had previously been retained/transitioned at the end of their kindergarten or first-grade year and had completed their seventh-, eighth-, or ninth-grade year at the time of the study. Therefore, student data were collected 5 to 10 years post-placement. California Achievement Test (CAT) reading and math standard scores reported at both pre-placement and current levels were used to compare academic achievement among the 24 retained students, the 22 transitioned students, and the matched control group of 24 students in the study.

Using academic achievement data from the study, effects were computed by transforming F values into effect sizes. Pre-placement reading and pre-placement math showed large effects ($ES = 0.90$ and 0.72 , respectively) while current reading and current math showed medium effects ($ES = 0.36$ and 0.48 , respectively). Phelps et al. (1992) suggested that neither retention nor transition placement resolved the academic deficiencies of the students in the study. An average ES of 0.59 was computed for the study.

Meisels and Liaw (1993) examined the effects of grade retention on 16,623 students in kindergarten through eighth grade using data from the National Educational Longitudinal Study of 1988. Participants in the study were 2,075 students retained early (Grades K-3), 1,128 students retained later (Grades 4-8), and 13,420 students never retained. Two sub-analyses of the data were conducted in the study. In the first sub-analysis, students who were retained in kindergarten through third grade were compared with those students who were retained in fourth through eighth grade. The second analysis compared students retained in kindergarten through eighth grade with the total sample of non-retainees. Results of these analyses suggested that the timing of retention was not consistently associated with improved performance, and retention at any point was associated with less favorable academic results.

In order to compute the effect sizes for Meisels and Liaw's (1993) study, the reported data were used to calculate t values if t values were not specifically given. The t values were then transformed into effect sizes (Equation 2). A descriptive analysis of grades and cognitive test scores for the retained vs. the never retained students revealed the largest effect sizes in the study (0.66 and 0.74 respectively) while a multivariate analysis of grades (0.21) and test scores (0.26) for the retained vs. never retained students found much lower effect sizes. An average ES of 0.27 was computed for the study.

Dennebaum and Kulberg (1994) examined later school achievement of 95 children that were retained in kindergarten, placed in a transition classroom, or recommended for retention but not retained. A control group of students who went straight from kindergarten to first grade was also used in the examination. The Metropolitan Achievement Test, Survey Battery, Sixth Edition (MAT-6) results over three grade levels were used to measure school achievement. Test scores in reading, mathematics, language, and the complete battery for each group and year were examined. Findings from the study were reported in *F* values and were converted to effect sizes using data reported in the study. An average *ES* of 0.55 was calculated for the study.

Results from Dennebaum and Kulberg's (1994) study indicated that children retained in kindergarten performed significantly lower on standardized achievement tests than did children in the other groups. In addition, the children who were recommended for retention, but went on to first grade, did not perform differently than the control group. Thus, the promoted children did not suffer the anticipated learning difficulties, and the notion that holding a child back cultivates later academic achievement was not supported.

Using achievement scores from either the Stanford Achievement Test (SAT) or the comparable Spanish version the Aprenda, Cosden et al. (1995) assessed the impact of holding out, retention, and social promotion on first-grade achievement. The practice of parents "holding out" their children until they were older and more ready to succeed in school was found not to have any significant effect on achievement scores on the SAT or the Aprenda. However, significant differences were found between the retained and non-retained students who took the SAT, with the students who had been retained in kindergarten scoring lower than the students not retained. For students who had taken the Aprenda, retention was not a significant factor in estimating achievement. Likewise, the advanced students (socially promoted to first grade

despite concerns about their performance) taking the SAT had significantly lower achievement scores than the promoted matched students, while no significant difference of achievement was found for the advanced students who had taken the Aprenda. In addition, a comparison between the students who had been advanced and the students who had been retained found no significant differences in SAT or Aprenda test scores. There was no difference in achievement at the end of first grade between the students who had been advanced to first grade after a year of kindergarten and those who had repeated kindergarten. However, students in both groups performed worse than the students in their matched group. Thus, neither holding out, nor retention, nor advancement had their expected impact on achievement.

Cosden et al. (1995) reported results of the study in F values which were transformed into effect sizes. A large effect size (0.75) was computed for the SAT achievement scores between the retained and the promoted matched group of students. The smallest effect size (0.26) was found for the Aprenda achievement scores between the retained and the advanced students. An average ES of 0.42 was computed for the study.

Kundert et al. (1995) compared scores from the Comprehensive Tests of Basic Skills (CTBS) for 314 students retained during Grades K-5 and 157 students who had been held out of kindergarten one year beyond the time they would have been eligible to enter school (delayed entry). Both groups of students were enrolled in Grades 3-12 during the study, and data from second-, fifth-, and seventh-grade CTBS scores were used to determine if there were any differences in achievement for the two groups. When achievement test scores were analyzed using IQ as a covariate, no significant differences were identified for either group. The authors determined that delayed entry does not lead to achievement advantages. Computed effect sizes for the study revealed very small effects with an average ES of 0.05. The confidence interval

calculation (Equation 6) for the study included zero (-0.13 to 0.23). According to Lipsey and Wilson (2001), if the confidence interval does not include zero, then the mean effect size is statistically significant at the level specified by the confidence interval (i.e., $\alpha = .05$ for a 95% confidence interval). Therefore, the mean effect size was found not to be significant.

The most common reason given by teachers advocating retention is low academic achievement. Thus, Jimerson et al. (1997) examined the effects of early elementary school retention on academic achievement throughout elementary school and again at 16 years of age. Twenty-nine retained children in Grades K-3 were compared to a randomly assigned comparison group of 50 low-achieving promoted children and a randomly assigned control group of 100 children. Assessment results from the Peabody Individual Achievement Test (PIAT) and the Woodcock-Johnson Achievement Test (WJ) were used to measure both short- and long-term effects of retention on academic achievement.

After converting the F values from 15 total effects into effect sizes, the short-term effects for kindergarten retainees upon completion of first grade revealed an effect size of 0.67 in reading comprehension, and an effect size of 0.59 in total battery. The largest short-term effect for first and second graders upon completion of second and third grade was found in math, with an effect size of 0.37. Reading comprehension ranked lowest among first and second grade retainees with an effect size of 0.14. The long-term effects for these students in the sixth grade revealed an effect size of 0.20 for both reading recognition and reading comprehension. Math achievement was found to be the lowest among sixth graders with an effect size of 0.09. In addition, the long-term effects for students at age 16 revealed small effects for both vocabulary and math achievement (0.10 and 0.17, respectively). An average ES of 0.28 was calculated for

the study. In addition, the confidence interval results found the study to be statistically non-significant (-0.11 to 0.67).

Jimerson et al. (1997) found little evidence to support retention as an intervention in elementary school for improving academic achievement. The authors found no significant differences on PIAT performance between the retained and the low-achieving promoted group despite the extra year of schooling for the retained students. Any gains found in academic achievement for the retained students disappeared in time as they faced new material.

McCoy and Reynolds (1999) extended an earlier study (Reynolds, 1992) by investigating the effects of retention on student achievement up to age 14. Both same-age and same-grade comparisons of reading and math achievement were used to examine the effects of retention. The variance (R^2) provided in the study for both reading and math achievement was converted to r by taking the square root of the variance. Then, r was transformed to an effect size measure of d (Equation 3). This resulted in rather large effects for both reading and math achievement. Same-age comparisons revealed an effect size of 2.30 for math, and an effect size of 1.88 for reading. Likewise, same-grade comparisons revealed very large effect sizes for math (1.96) and reading (1.74). Both approaches were consistent in finding that retained students scored significantly lower than their promoted peers.

McCoy and Reynolds (1999) also reported effect sizes on the timing of grade retention on reading and math achievement. The authors examined any retention (Grades 1-7), early retention (Grades 1-3), and later retention (Grades 4-7). The effect sizes reported in the study were calculated by dividing the mean difference between groups by the within-group standard deviation. Effect sizes revealed that the relationship between retention and lower reading achievement was greater for early retention ($ES = 0.40$) than for later retention ($ES = 0.19$). For

math, both early retention ($ES = 0.39$) and later retention ($ES = 0.36$) were associated with lower achievement. In addition, students who were retained during first to seventh grade had significantly lower scores in reading and math achievement with effect sizes of 0.42 and 0.48 respectively. The average effect size calculated for the study was 1.17.

In order to examine the association between grade retention and academic adjustment in high school, Jimerson (1999) conducted a study from data obtained from a 21-year longitudinal study that included retained students, low-achieving but promoted students, and a control group. Contrasts for the three groups were calculated on levels of high school adjustment in 11th grade and reported in F values, which were converted to effect sizes. The results of the retained vs. control group and the retained vs. comparison group revealed an effect size of 0.66 for the former, and an effect size of 0.58 for the latter. These results indicated that the retained group of students exhibited lower academic adjustment than either the promoted or the control group. No significant difference was found between the comparison and the control group ($ES = 0.21$). Jimerson's study provides evidence that retained students have a greater probability of poorer educational outcomes during late adolescence. An average effect size of 0.48 was calculated for the study.

Table 4 indicates the studies published between 2000 and 2010 that were included in this meta-analysis. The average effect size and confidence interval are included for each study.

Table 4

Average Effect Size and Confidence Interval (2000-2010 Studies)

Study/Year	Average Effect Size	Standard Error	95% CI	
			LL	UL
Beebe-Frankenberger et al. (2004)	0.67	0.14	0.40	0.94
Bonvin et al. (2008)	0.55	0.14	0.28	0.82
Burkam et al. (2007)	0.24	0.03	0.18	0.30

(Continued)

Table 4 (Continued)

Average Effect Size and Confidence Interval (2000-2010 Studies)

Study/Year	Average Effect Size	Standard Error	95% CI	
			LL	UL
Dong (2010)	0.27	0.06	0.15	0.39
Gleason et al. (2007)	0.94	0.14	0.67	1.21
Griffith et al. (2010)	0.31	0.04	0.23	0.39
Hong & Raudenbush (2005)	0.54	0.04	0.46	0.62
Hong & Yu (2007)	0.11	0.04	0.03	0.19
Jacob & Lefgren (2004)	0.11	0.02	0.07	0.15
Jimerson & Ferguson (2007)	0.50	0.17	0.17	0.83
Lorence & Dworkin (2006)	0.57	0.03	0.51	0.63
Mantzicopoulos (2003)	0.67	0.24	0.20	1.14
Moller et al. (2006)	0.34	0.03	0.28	0.40
Roderick & Nagaoka (2005)	0.16	0.03	0.10	0.22
Silberglitt, Appleton, et al. (2006)	0.34	0.17	0.01	0.67
Silberglitt, Jimerson, et al. (2006) +	0.34	0.28	- 0.21	0.89
Willson & Hughes (2009)	0.40	0.09	0.22	0.58
Wu et al. (2008a) +	0.21	0.14	- 0.06	0.48
Wu et al. (2008b) +	0.21	0.14	- 0.06	0.48

Note. CI = confidence interval; LL = lower limit, UL = upper limit. + denotes study results were insignificant at 95% confidence interval.

2000-2010 Studies

One alternative to grade retention that has been used in schools across the United States is the transition room program. Mantzicopoulos (2003) examined the effectiveness of such a program by following 71 kindergartners through the end of third grade. Thirty-four students who were recommended-and-enrolled in a transition room program (RE) were compared to 37 recommended-but- not enrolled (RNE) students using reading and math achievement results from the CTBS and the state's Mandated Achievement Test (SMAT). Reported effect sizes revealed that at the end of first grade, RE students out-performed RNE students on academic achievement test scores. The effect sizes for CTBS reading, CTBS math, and CTBS total were very large ($ES = 0.85, 0.94, 0.98$, respectively). However, at the end of second grade, effect sizes declined in magnitude ($ES = 0.44, 0.89, \text{ and } 0.68$ for SMAT reading, SMAT math, and SMAT

total, respectively). By the end of third grade, the effect sizes for SMAT reading, SMAT math, and SMAT total had declined even more ($ES = .00, 0.72, 0.49$, respectively).

In Mantzicopoulos' (2003) study, it was unclear if the findings supported the transition room program. Although effect sizes dropped dramatically in reading by the end of third grade, the effect sizes for math remained moderate at the end of third grade. An average ES of 0.67 was calculated for the study.

Beebe-Frankenberger et al. (2004) compared a group of 224 second-grade students on academic competence based on school retention/promotion decisions. The cohort consisted of 64 retained students, 41 at-risk students, 46 special education students, and 73 promoted students. Using F values reported in the study, achievement scores from the SAT and the WJ were transformed into effect sizes. SAT reading was found to have the greatest effect size in the comparison, with SAT spelling having the least ($ES = 0.79$ and 0.56 , respectively). A comparison of a smaller sub-sample of 137 low achievers revealed lesser effects than the original sample. In this sample, the reading scores from the WJ achievement test revealed effect sizes of 0.60, 0.46, and 0.46 for letter and word recognition, passage comprehension, and word attack, respectively. None of the 10 effects calculated in the study had an effect size below 0.40, and the average effect size was 0.62. The results of the study found evidence that retained students may not have the necessary ability to catch up.

Since higher academic standards and increased accountability have become more important in the field of education, interventions such as grade retention are being used more by educators in order to help low-achieving students meet minimum academic standards. Jacob and Lefgren (2004) examined retention treatment effects for students in the third and sixth grade experiencing an accountability policy in the Chicago Public School system. The math and

reading sections of the Iowa Test of Basic Skills (ITBS) were used to assess the effect of grade retention on student achievement. Data provided in the study were reported in t values which were converted into effect sizes. The eight effects calculated in the study produced a very small average effects size of 0.11. The authors suggested that grade retention had a fairly small but positive impact on student achievement scores for third-grade students and may increase academic achievement for low-achieving third graders. However, there were little if any effects on academic achievement for sixth-grade students.

Roderick and Nagaoka (2005) also examined the experience of third and sixth graders who were retained under Chicago's testing policy. The authors observed the progress of the students in the study for two years after retention and estimated the short-term effects of retention on reading and math achievement. Promotional requirements for students under Chicago's policy were based on test scores on the ITBS. The results of the achievement scores in the study were reported in t values and were converted into effect sizes for this study. A total of 12 effects were used to calculate an average effect size of 0.16. According to Roderick and Nagaoka's (2005) findings, there was no evidence that retention led to improved achievement growth two years after the promotional gate for third graders, and grade retention was associated with lower achievement growth for sixth graders. Therefore, the retention of low-achieving students did not help.

Hong and Raudenbush (2005) analyzed data from the Early Childhood Longitudinal Study-Kindergarten cohort (ECLS-K) with the technique of multi-level propensity score stratification. The authors reported means and standard deviations for both retained and promoted students for reading and math achievement within each individual-level propensity stratum in schools that permitted grade retention. These data were used to calculate effect sizes

for this study. The results indicated that promoted students scored higher than retained students in both reading and math achievement ($ES = 0.61, 0.47$, respectively). An average effect size of 0.54 was calculated for the study.

The findings of Hong and Raudenbush's (2005) study revealed that students who were retained would have learned more had they been promoted. The average loss in academic growth for retainees, after one year, was almost two-thirds of a standard deviation (half a year's expected growth) for each subject. In addition, the authors concluded that kindergarten retention left most retainees further behind and challenged the practice of regularly using kindergarten retention as an answer for the academic struggles experienced by adolescent children.

In a longitudinal study conducted by Silbergitt, Appleton, Burns, and Jimerson (2006), the reading growth of 147 students from first- through eighth-grade was examined. Participants were divided into three groups consisting of 49 retained students, 49 promoted students, and 49 randomly selected students (control group). Each student completed a series of reading curriculum-based measures (R-CBM) as they progressed through school, and the compiled data were used to compare growth rates for the retention, matched, and random groups. After transforming t values into effect sizes, comparisons between the retained vs. matched group and the retained vs. random group revealed effect sizes of 0.05 and 0.63, respectively. An average effect size of 0.34 was calculated for the study. Compared to similarly performing promoted students, retained students did not show any gains or losses in their growth rates as a result of retention. However, the growth curve of the randomly selected group was significantly larger than the progress of the retained students.

In another study, Silbergitt, Jimerson, Burns, and Appleton (2006) examined the effects of early vs. later retention on R-CBM reading growth for the 49 retained students from the

previously mentioned study. The students were divided into one of two groups based on grade retention, with 27 early retained students (K-2) and 22 later retained students (3-5). The authors' findings revealed that the reading growth of the students who were retained early were comparable to the students retained later. Thus, retention in elementary school was not supported whether it was early or later. The average effect size for the study was 0.34. The confidence interval for the study ranged from - 0.21 to .89 and was found to be statistically insignificant.

Using the 1988-1992 National Education Longitudinal Study (NELS), Moller, Stearns, Blau, and Land (2006) examined math and reading achievement trajectories of high school students over three time periods who had been retained prior to eighth grade. All students who completed a cognitive test that was administered in the eighth, tenth, and twelfth grades were included in the study's analytic sample ($N = 9,045$). Growth modeling, similar to hierarchical modeling, was used to predict both initial scores in the eighth grade and growth in scores between the eighth and tenth, and tenth and twelfth grades. This technique allowed the authors to see how retention impacted achievement gains in high school, controlling for students' initial eighth grade scores. To calculate the effect sizes for initial score and growth on the math and reading scores of retained students, the data reported in the study were converted into t values and then transformed into Cohen's d . This revealed an effect size of 0.49 on initial score and an effect size of 0.20 on growth. An average effect size of 0.34 was calculated for the study.

Moller et al. (2006) found that retention is one of the strongest predictors of both initial scores and rates of growth. Utilizing growth modeling to predict math and reading scores, one model revealed that students who were retained prior to eighth grade had initial achievement scores 5.09 points lower than promoted students. Thus, being aware of past experiences with retention throughout a student's school career can be beneficial when predicting achievement

trajectories. Furthermore, Moller et al. found that students who were retained experienced fewer gains in achievement than promoted students.

Lorence and Dworkin (2006) followed third-grade students who failed the Texas Assessment of Academic Skills (TAAS) reading test through their sophomore year of high school. The study examined 38,445 White, Hispanic, and Black students, of which, 1,244 low-performing students were required to repeat third grade. The remaining students were placed in fourth grade. The authors used same-grade comparisons over same-age comparisons in order to obtain academic achievement differences between the retained and promoted students who had covered the same material. Results for eight grades were examined.

Instead of relying on traditional tests of statistical significance to assess the extent of the reading performance differences between retained and promoted students, Lorence and Dworkin (2006) calculated effect sizes for each grade examined within each of the three comparison groups. Effect sizes were obtained by taking the adjusted mean difference between the retained and promoted groups and dividing by the standard deviation of the control group, or the socially promoted pupils. Results revealed very large effects for third-grade repeaters in the White, Hispanic, and Black comparison groups with effect sizes of 1.96, 1.57, and 1.57 respectively. In fourth grade, the effect sizes for White, Hispanic, and Black students dropped considerably ($ES = 0.64, 0.60, 0.61$, respectively) but large effects still existed. Eighth-grade reading scores showed the smallest effects between the retained and promoted students in each comparison group with effect sizes of 0.13 (White), 0.12 (Hispanic), and 0.08 (Black). In tenth-grade, the effect size for the White students was 0.13, and the effect sizes for Hispanic and Black children were 0.41 and 0.40. Using the 21 reported effects, an average effect size of 0.57 was calculated for the study. According to the study's authors, low-performing students, regardless of race, who were required

to repeat third-grade exhibited increased reading performance despite the fact that some effects were very small.

Gleason et al. (2007) conducted a one-year longitudinal study to examine the short-term effects of grade retention on peer relations and academic performance of 350 at-risk first graders. Data were collected in first grade and one year later, at which time 63 students were repeating first grade and 287 were in second grade. Woodcock-Johnson reading and math scores were used to measure the students' academic achievement, employing both age and grade standard scores. Results were reported in F values and transformed into effect size d , revealing effect sizes of 1.02 and 1.10 for same-age reading and same-grade reading and 0.54 and 1.11 for same-age math and same-grade math, respectively. An average effect size of 0.94 was calculated for the study.

Gleason et al. (2007) found that promoted students scored higher in year two on the reading and math age-standard scores, and the retained students scored higher in year two on the reading and math grade-standard scores. Thus, same-age comparisons favored the promoted students while same-grade comparisons favored the retained students. Positive short-term effects of early grade retention were found in the study suggesting that repeaters benefited academically from the extra year in first grade.

Using the Early Childhood Longitudinal Study-Kindergarten cohort (ECLS-K), Burkam et al. (2007) studied nationwide patterns of kindergarten retention and its consequential academic effects on literacy and mathematics achievement of 12,780 students. The authors examined the students' achievement test scores using multivariate models at the end of kindergarten and at the end of first grade. Test scores were standardized into ES units and reported in the study.

Burkam et al. (2007) found that students who repeated kindergarten ended the school year slightly behind first time kindergarteners in literacy achievement ($ES = 0.11$). However, kindergarten repeaters were considerably behind other first-time kindergarteners the following year ($ES = 0.88$). Similarly, kindergarten repeaters and first-time kindergarteners who were retained did not perform as well on mathematics achievement as the other first-time kindergartners who were promoted ($ES = 0.10$ and 0.85 , respectively). In addition, kindergarten repeaters displayed lower achievement levels in literacy and mathematics at the end of first grade than children who spent only one year in kindergarten ($ES = 0.47$ and 0.31 , respectively). The authors concluded that repeating kindergarten seldom leads to improved literacy and mathematics achievement in kindergarten and first grade. The average effect size for the 18 effects in the study was 0.24 .

Hong and Yu (2007), also using ECLS-K data, examined the effects of kindergarten and first-grade retention relative to promotion on 21,409 students' reading and math achievement during the elementary years. Propensity score stratification was used to eliminate selection bias due to a substantial number of pre-treatment covariates. Reading and math achievement data were reported in t values, which were transformed into effect sizes. This transformation revealed harmful effects in both reading ($ES = 0.26$) and mathematics ($ES = 0.21$) at the end of the treatment year for the retained kindergarten students, with effect sizes in reading ($ES = 0.05$) and math ($ES = 0.05$) diminishing considerably four years after retention. Similar results were found on the first grade retainees' reading and math achievement results as well. In fact, Hong and Yu stated that kindergarten and first-grade retainees never achieved more on average than they would have if they had been promoted in the first place. An average effect size of 0.11 was calculated for this study.

In a 12-year longitudinal study, Jimerson and Ferguson (2007) examined achievement outcomes by comparing low-achieving promoted students with retained students through adolescence. Participants in the study were 137 students who were identified in kindergarten through second grade, of which 72 were followed through eleventh grade. The Science Research Associates' Survey of Basic Skills (SRA) test, the Stanford Achievement Test (SAT), the Armed Services Vocational Aptitude Battery (ASVAB), and grade point averages (GPAs) were used to measure the academic achievement of participants over the years. Jimerson and Ferguson reported the results for the analyses of achievement in effect sizes (i.e., reported η^2). However, for this study, effect sizes were calculated by transforming F values reported in the study to Cohen's d .

Jimerson and Ferguson (2007) reported effect sizes (η^2) for each academic outcome that ranged from 0.11 to 0.22 (small effects). Conversely, F values transformed into Cohen's d revealed effect sizes that ranged from 0.40 to 0.61 (medium to large effects). Nevertheless, according to the authors, the results of the study failed to validate the helpfulness of grade retention on academic achievement relative to comparison groups of promoted students. An average effect size (Cohen's d) of 0.50 was calculated for the study.

In order to determine how effective grade retention was with respect to academic progress in mathematics and language, Bonvin et al. (2008) studied 83 pairs of matched (retained vs. promoted) second grade students from German- and French-speaking Swiss schools. Achievement tests in mathematics and language were created on the basis of the educational programs in both Swiss linguistic regions and administered three times between May 2001 and May 2003. Consequently, the first measurement point was taken before the decision was made to retain any students. Both same-age and same-grade comparisons were examined in the study

with the results of the variance analysis reported in an effect size measure known as eta squared (η^2). Since Eta Squared is comparable to R^2 from multiple linear regression, the square root of η^2 was taken, substituted as r , and converted to Cohen's d .

As with most same-age and same-grade comparisons, Bonvin et al. (2008) found advantages for promoted students when compared to same-age retained students and advantages for retained students when compared to same-grade classmates. The largest effect for same-age comparisons of retainees and promoted poor achievers was found in mathematics ($ES = 0.81$) with the largest effect for same-grade comparisons of retainees and poor performing promoted students found in language ($ES = 0.77$). In addition, the retained students' performance was compared to all new classmates of the retainees (norm group) since the beginning of the repeated grade. This comparison revealed that the academic gains of the retainees decreased, and eventually fell below their classmates at the end of third grade. Mathematics was found to have the highest effect size in the comparison ($ES = 0.67$). Finally, performance differences between the comparison groups (retained and poor performing promoted students) were compared to the norm group. This same-grade comparison showed retained students benefited from the additional year, getting significantly closer to the mean of the norm group in math and language than the control group of promoted students. Language had the largest effect size in the comparison ($ES = 0.77$). The 16 effects in the study ranged from 0.20 to 0.81 with an average effect size of 0.55 calculated for the study. It is suggested by the authors that grade retention be avoided at the primary school level.

Wu et al. (2008b) examined the effects of first-grade retention on the growth of the Woodcock-Johnson mathematics and reading scores over three years using linear growth curve modeling on an academically challenged sample. Initially, 784 never-before retained first grade

students were identified based on low literacy scores at the beginning of their first grade year. Using optimal matching procedures, 97 pairs of retained and promoted students were matched based on their propensity scores (predicted probability of being retained) and were included in the study's growth curve analysis. The results from the effects of retention on the WJ math and reading scores were reported using Wald- z scores. Since a Wald- z score is a Chi-square distribution, it was transformed into an r value (Equation 4) and converted to Cohen's d (Equation 3) in S-PLUS®. This conversion revealed small effects in both math and reading ($ES = 0.26$ and 0.16 , respectively). The average effect size calculated for the study was 0.21 . In addition, the study was found not to be statistically significant when the confidence interval was calculated and revealed that it included zero ($- 0.06$ to 0.48).

The authors concluded that grade retention decreased the rate of growth for mathematics but did not have a significant effect on reading in the initial two years following retention. No evidence of overall beneficial short-term effects of grade retention was provided. In addition, retention seemed to be less beneficial for older students than for younger students (Wu et al., 2008b).

Extending their previous study, Wu et al. (2008a) examined the short- and longer-term effects of grade retention on the changes of WJ math and reading W (Rasch-modeled W scores) and grade scores over four years. The previous study analyzed only W scores. Drawing from the same sample of participants as the earlier study, 98 pairs of students were matched using propensity scores. Descriptive statistics revealed an effect size of 0.07 between the retained and promoted students' reading and math scores at the beginning of the study which indicated that the matching procedure was successful in eliminating initial differences in academic achievement. A two-piece linear growth curve model was employed in the study and results were

reported using Wald- z scores. The Wald- z scores were transformed into effect sizes in the same manner as described in the previous study. The transformation revealed effect sizes of 0.28, 0.31, 0.38, and 0.40 from grade retention (Slope 2 – Slope 1) for math W scores, reading W scores, math grade scores, and reading grade scores, respectively. An average effect size of 0.21 was calculated for the study. Like the previously mentioned Wu et al. (2008b) study, the confidence interval concluded that the study was not statistically significant (- 0.06 to 0.48).

Wu et al. (2008a) found that when using W scores, retained students experienced a gradual increase in both math and reading achievement in the short-term but a more rapid increase in reading achievement in the longer-term than did the promoted students. Grade standard scores revealed that retained students had a rapid increase in the short-term but a quicker decrease in the longer-term in both math and reading achievement than did promoted students. Benefits of grade retention seemed to fade as retained students encountered an unfamiliar and more challenging curriculum.

Willson and Hughes (2009) examined reading and math W scores from the Woodcock-Johnson Tests of Achievement from an analytical sample of 769 first-grade students. Of the participants, 165 students were retained in first grade, and 604 students were promoted. The authors reported an effect size of 0.52 for reading and an effect size of 0.28 for math and concluded that the retained students performed lower than the promoted children in academic competence. Willson and Hughes stated that students' academic competence was the greatest variable connected with, and primarily responsible for, retention in first grade. An average effect size of 0.40 was calculated for the study.

Dong (2010), using data from the 1998-1999 ECLS-K cohort, investigated whether or not retained kindergarten students performed better academically than they would have done had

they been promoted. The sample of participants included 8,391 promoted and 281 retained students who were either first-time kindergarteners or kindergarten retainees in the 1998-1999 school year. Using a linear regression model, the average retention effects on the retained kindergarteners' academic performance in reading and math were reported by Control Function (CF) and Ordinary Least Squares (OLS) estimates. The CF estimates controlled for both random and non-random selection into retention schools, and the OLS estimates were controlled by either using or not using pre-treatment scores. Overall, 16 effects were reported with the CF estimates revealing larger effects than the OLS estimates. Effect sizes by CF ranged between 0.18 and 0.61. Effect sizes by OLS ranged between - 0.26 and 0.40. OLS estimates not controlling for pre-treatment test scores produced a negative sign on the treatment effect. However, adding pre-treatment test scores to the OLS switched the sign to positive. An average effect size of 0.27 was calculated for the study.

The results of the study revealed that retained students on average performed better than they would have performed had they been socially promoted. Math test scores were still significant up to third grade while the effect on the reading test scores was not. Dong's (2010) study found that holding students back in kindergarten had positive but diminishing effects on their academic performance up to third grade. Therefore, kindergarten retention may give low performing students a chance to close the achievement gap.

Griffith et al. (2010), using National Education Longitudinal Study (NELS) data, examined the reading growth rates of 878 secondary students who had been retained between kindergarten and eighth grade and 878 students in a matched comparison group who had never been retained. The result of the one-sample *t*-test was transformed into an effect size using

Equation 2. Reading achievement was the only dependent variable in the study, and the average effect size of the study was 0.31.

The findings revealed a significant discrepancy between the scores of the retained and promoted groups on reading achievement in the eighth grade, with the difference benefiting the promoted group. In addition, students in the comparable promoted group had greater growth over the high school years than their retained peers.

Interestingly, a comparison of the average effects across the two decades revealed an average effect size of 0.66 for the 10 significant 1990-1999 studies while an average effect size of 0.42 was revealed for the 16 significant 2000-2010 studies. While retention was found to have very negative effects on the academic performance of students in both decades, the lower average effect size for the 2000-2010 studies may be the result of better designed and higher quality studies.

Summary

This chapter examined the findings from the comprehensive review of the studies included in the meta-analysis. Studies were reported in chronological order to better understand the development of grade retention research from 1990-2010. In addition, average effect sizes and confidence intervals were calculated and reported for each study. Chapter V describes the implications of the research findings and discusses implications for practitioners and recommendations for future research.

CHAPTER V

DISCUSSION AND RECOMMENDATIONS

Introduction

This chapter reviews the major methods used in the study and describes implications of the research findings. The chapter begins with a statement of the problem and a review of the methodology and provides an analysis of the findings identified in Chapter IV. The chapter then provides implications for practitioners and recommendations for future research.

Statement of the Problem

This study examined the relationship between grade retention of elementary students and its effect on academic performance by evaluating the available scientific literature on the subject. The primary hypothesis in undertaking this study was that there is a positive relationship between grade retention and students' academic performance.

Retention has been a source of dispute in American public education for years, and its use as an educational intervention for low-performing students has varied since the early 1900s. Elementary school retention is embedded in the hierarchical nature of early schooling. The primary reason for retention is the belief that doing so allows students an opportunity to master the curriculum they have failed and become proficient enough for future academic success (Dauber et al., 1993). In the education literature, retention is viewed as negative in many published studies and meta-analyses. However, Lorence (2006) argued that education researchers have been too quick to reach the conclusion that retention is harmful. In fact, Allen et al. (2009) challenged the widely held view that retention has a negative impact on academic achievement.

The literature on retention effects shows that data can be found that appears to support arguments for or against retention. The truth be known, researcher/publication bias, study quality, age and/or grade comparisons, methodological limitations, absence of a randomized experimental group, failure to control for pre-existing differences in students, and length of study, only to name a few, all contribute to the reporting of conflicting results in the grade retention literature. It is these conflicting results that cause administrators, teachers, and parents to wonder about what to do with students who have fallen behind their peers academically. Thus, it is important to examine the results of the most current and best-designed grade retention studies available to answer this question.

Review of the Methodology

This study is a meta-analysis of the literature regarding the relationship between grade retention and the academic achievement of elementary students. The study provided a quantitative summary of findings from available studies published between 1990 and 2010.

An extensive systematic review of the literature was conducted using bibliographic databases and other sources, resulting in the review of hundreds of references from articles, books, book chapters, conference papers, and dissertations. From these references, 68 articles were identified as potential studies for inclusion in the meta-analysis. Each study was reviewed to determine if it met the criteria (e.g., a minimum of one reported academic outcome measure of grade retention, involved K-6 students at time of retention, reported effect size or had sufficient data to calculate effect size, relevant to this study's research questions) for inclusion in the meta-analysis, narrowing the number of potential studies to 43. After the abstraction of data and an in-depth evaluation of these 43 studies, 12 more studies were eliminated. The eliminated studies did not include data to calculate effect size, did not relate to this study's research questions, or

reported only significant findings. Therefore, 31 studies were included in the meta-analytic calculations, and the data were provided in Chapter IV.

A meta-analysis focuses on the magnitude of the effect observed in a study. Meta-analysis uses a metric, called an effect size, to determine the strength of the relationship between two variables that are being evaluated. Lipsey and Wilson (2001) described the effect size statistic as the standardization of study findings that permit meaningful numerical comparison and analysis across studies. As opposed to narrative reviews that usually rely on *p*-values to either accept or reject the null hypothesis, a meta-analysis relies on an effect size which has much more practical significance than a *p*-value. Many times, researchers interpret non-significant results to mean that there is no effect. However, just because something is not significant does not mean it does not have an effect. A *p*-value can only report that an effect is not zero, and to report that an effect is not zero is to miss the purpose (Borenstein, Hedges, Higgins, & Rothstein, 2009). Average effect sizes were computed for each study, and a mean effect size and confidence interval were calculated for each study to test the null hypotheses to determine if the effect sizes were likely to be a true difference or whether they could have occurred by chance.

Analysis of the Findings

The general expression of the null hypothesis in a meta-analysis is to test the significance of an average effect size. To achieve the objectives of identifying the magnitude of the effect of grade retention of elementary students on academic achievement, four research questions and two generally expressed null hypotheses were examined.

Research question 1: What is the research consensus on the effects of grade retention on academic performance of students?

Research question 2: If initial gains are evident in academic achievement from repeating a grade in elementary school, does the research show that they persist in high school?

Research question 3: Does the research show that retaining a student in kindergarten improves the academic performance of the student?

Research question 4: Does the research show that retaining students early (Grades 1-2) benefits students' academic performance more than retaining them later (Grades 3-6)?

The first null hypothesis is: there is no significant ($\alpha = .05$) relationship between grade retention and academic performance of students.

The second null hypothesis is there is no significant ($\alpha = .05$) difference in academic performance, over time, for students who receive grade retention as an intervention compared to those students who did not. Both null hypotheses were rejected.

The study hypothesis is that there is a positive relationship between grade retention and students' academic performance. That is, grade retention is an effective intervention and produces its intended result.

The findings related to each of the primary research questions examined in this study are reported below.

Question 1: What is the research consensus on the effects of grade retention on academic performance of students?

This study provided a meta-analysis of studies related to the effects of grade retention on academic achievement for primarily two dependent variables: reading and mathematics. Data for these two variables were obtained from various state-mandated standardized achievement tests and/or basic tests of academic achievement administered to each of the study's participants.

Thirty-one studies provided a total of 274 academic outcomes related to the effects of grade retention and students' academic performance. However, five studies (Jimerson et al., 1997; Kundert et al., 1995; Silberglitt, Jimerson, et al., 2006; Wu et al., 2008a; Wu et al., 2008b) insignificant at the 95% confidence interval were not included in the average effect size calculation. Thus, the remaining 26 studies (245 academic outcomes) revealed an average effect size of 0.50 which indicated that the retained students scored 0.50 standard deviations lower than the promoted students on the academic outcome measures. Effect sizes for the 26 studies ranged from 0.11 to 1.17 with only four studies reporting improved academic performance as a result of grade retention (Dong, 2010 at 0.27; Gleason et al., 2007 at 0.94; Jacob & Lefgren, 2004 at 0.11; Lorence & Dworkin, 2006 at 0.57). One study (Jacob & Lefgren, 2004) reported benefits for retained students up to third grade, but these benefits diminished by sixth grade. In addition, Gleason et al. (2007) examined retention effects on students' academic achievement only one year post-retention. The short-term effects suggested the retained students benefited from the extra year of schooling; however, the long-term effects were not examined. Therefore, caution should be used when considering the implications of the study.

Using Hattie's (2009) guidelines for effect size magnitude when judging educational outcomes, effect sizes of 0.20, 0.40, and 0.60 are suggested as small, medium, and large effects, respectively. Therefore, an effect size of 0.50 is considered a medium effect and exceeds the $d = 0.40$ benchmark referred to as "the hinge-point." According to Hattie, effect sizes above the hinge-point fall into the "zone of desired effects" where real-world differences can be observed.

Question 2: If initial gains are evident in academic achievement from repeating a grade in elementary school, does the research show that they persist in high school?

Originally, 12 studies were related to this research question, but after eliminating three statistically insignificant studies (Jimerson et al., 1997; Kundert et al., 1995; Silberglitt, Jimerson, et al., 2006), an average effect size of 0.60 (large effect) was calculated from nine studies consisting of 59 academic outcome measures (see Table 5). Effect sizes ranged between 0.27 and 1.97. Eight of the nine studies concluded that retained students performed significantly lower than promoted students. The study contributing to the largest effect size of 1.97 (McCoy & Reynolds, 1999) investigated the effects of retention on student achievement up to age 14 (over 5 years post-retention).

Lorence and Dworkin (2006) was the only study to report positive effects for retained elementary students now in high school ($ES = 0.57$). The authors used same-grade comparisons in order to obtain achievement differences between retained and promoted students who had covered the same material. However, same-grade comparisons normally favor retained students while same-age comparisons normally favor promoted students (Gleason et al., 2007). The type of comparison used to evaluate the effectiveness of grade retention must be considered when interpreting results.

Question 3: Does the research show that retaining a student in kindergarten improves the academic performance of the student?

To answer this question, nine studies containing 95 academic outcome measures were examined. Jimerson et al. (1997) was originally included but, as mentioned previously, was removed due to insignificance (see Table 6). The effect sizes for the kindergarten retention studies ranged from 0.11 to 0.77 with an average effect size of 0.46 (medium effect). Eight of the nine studies either found that kindergarten retainees never achieved more academically than they

would have achieved had they been promoted, or they found that the retainees performed lower academically than their promoted classmates.

On the other hand, Dong (2010) found that kindergarten students performed better academically than they would have performed had they been socially promoted ($ES = 0.27$). However, holding students back in kindergarten had diminishing effects on academic performance up to third grade. Although kindergarten retention appeared to give low performing students a chance to catch up, the effect size (0.27) does not exceed the hinge-point of $d = 0.40$. Therefore, one may question if kindergarten retention makes a visible difference in academic achievement after all things are considered.

Question 4: Does the research show that retaining students early (Grades 1-2) benefits students' academic performance more than retaining them later (Grades 3-6)?

Initially, six studies related to the timing of grade retention were examined for this question. After eliminating two insignificant studies (Kundert et al., 1995; Silbergitt, Jimerson, et al., 2006), four studies with 35 academic outcome measures were used to calculate a medium average effect size of 0.50 (see Table 7). Effect sizes ranged from 0.27 to 0.77. Johnson et al. (1990) had the largest effect size at 0.77. Johnson et al. found that grade retention was not effective regardless of the timing. Pierson and Connell (1992), with an effect size of 0.60, suggested that retention in the early years (Grades 1-2) was beneficial to students' academic performance in comparison with social promotion. McCoy and Reynolds (1999) found that early retained students had lower reading achievement scores than later retained students ($ES = 0.37$). In addition, McCoy and Reynolds found that both early and later retention was associated with lower achievement scores in math. Finally, Meisels and Liaw (1993), with an effect size of 0.27, found that the timing of retention was not consistently associated with improved academic

performance and not recommended at any time. Therefore, due to the fact that the majority of the studies in this meta-analysis do not support grade retention as an effective intervention, it is probably irrelevant to debate early versus later retention. Retention is retention regardless of when it occurs.

Table 5

Studies Related to Research Question 2

Study/Year	# Students Retained	# Students Not Retained	Average Effect Size
Griffith et al. (2010)	878	878	0.31
Jimerson (1999)	29	150	0.48
Jimerson & Ferguson (2007)	74	63	0.50
Lorence & Dworkin (2006)	1,244	37,201	0.57
McCoy & Reynolds (1999)	315	843	1.97
Meisels & Liaw (1993)	3,203	13,420	0.27
Moller et al. (2006)	1,805	7,240	0.34
Phelps et al. (1992)	46	24	0.59
Silberglitt, Appleton et al. (2006)	49	98	0.34
Average Effect Size			0.60

Table 6

Studies Related to Research Question 3

Study/Year	# Students Retained	# Students Not Retained	Average Effect Size
Burkam et al. (2007)	933	10,328	0.24
Cosden et al. (1995)	36	194	0.42
Dennebaum & Kulberg (1994)	53	42	0.55
Dong (2010)	281	8,391	0.27
Hong & Raudenbush (2005)	471	7,168	0.54
Hong & Yu (2007)	471	10,255	0.11
Johnson et al. (1990)	20	37	0.77
Mantzicopoulos (2003)	34	37	0.67
Phelps et al. (1992)	46	24	0.59
Average Effect Size			0.46

Table 7

Studies Related to Research Question 4

Study/Year	# Students Retained	# Students Not Retained	Average Effect Size
McCoy & Reynolds (1999)	315	843	0.37
Meisels & Liaw (1993)	3,203	13,420	0.27
Johnson et al. (1990)	20	37	0.77
Pierson & Connell (1992)	74	164	0.60
		Average Effect Size	0.50

Conclusions

As presented in this meta-analysis, and other major studies on grade retention, the majority of evidence does not support retention to improve academic achievement. Therefore, this study's hypothesis that there is a positive relationship between grade retention and students' academic performance was rejected. However, despite the overwhelming data in the literature against the use of grade retention, it continues to be used as an intervention for low-performing students. In fact, many educators and parents encourage the use of grade retention when students lag behind their peers academically and/or socially. In addition, the push for higher academic standards and more accountability has put a tremendous amount of pressure on school systems, administrators, and teachers to ensure their students perform well on state mandated achievement tests in order to make Adequate Yearly Progress (AYP). As a result of this pressure, many educators turn to grade retention instead of socially promoting students that are not academically prepared for the next grade. Whether it is because they are not familiar with the current research on grade retention, whether they believe retention works based on past experiences, or simply because they think they have no other choice, educators should avoid the use of retention and look for other methods of intervention to improve students' academic performance.

Implications for Practitioners

In many grade retention situations, the teacher is the most powerful variable in determining whether a child will be retained (Johnson et al., 1990). Therefore, it is extremely advantageous to the child if his or her teacher is knowledgeable of the most current grade retention literature provided in this meta-analysis. Teachers and administrators commonly believe that retaining a student is beneficial, especially in the early grades of elementary school, because they only see the short-term gains and do not see the later difficulties that the large majority of retainees experience. One particular long-term difficulty associated with grade retention is dropping out of high school. Rumberger (1995) found that grade retention was the single most powerful predictor influencing a student's decision to leave school. Thus, elementary teachers and administrators need to focus more on the long-term success of their students and find ways other than retention to increase achievement. Retaining a student to do the same thing twice is a bad idea. Educators should provide research-based interventions with a focus and the resources to correct well documented, individual problems. According to Martin (2009), students are better served by staying with their age-appropriate peers and receiving the necessary interventions.

In addition, the research literature on grade retention should be used to develop school systems' promotional policies. Due to increased accountability and higher academic standards, some school systems require students to meet minimum academic standards on state-mandated tests in order to be promoted to the next grade. Since retention is ineffective in improving academic achievement, and is very costly, it is not wise to adopt a policy that makes retention mandatory. Also, Hattie (2009) found that the threat of grade retention is not a motivating force

for students. Therefore, retention policies should be written so school administrators have flexibility when retention issues occur.

Educators, researchers, and policy-makers have a duty to ensure that all students receive instruction that is appropriate for their specific educational needs. If children are required to perform at a one-size-fits-all level to show that they have not been left behind, then educators are destined to failure. On the other hand, if students are being taught using best practices that specifically focus on each student's learning needs, educators can be successful (Beebe-Frankenberger et al., 2004). Therefore, research and practice should focus on how to best meet the instructional needs of low-achieving students without using grade retention as an intervention

Implications for Researchers

Although the majority of published literature on grade retention does not support the use of retention as an effective intervention for improving students' academic achievement, there may be many unpublished studies that support grade retention. The problem lies in the fact that many studies are not published due to either very small or insignificant effects. As discussed earlier, just because an effect is small, or statistically insignificant, does not mean that it has zero effect. Therefore, it would be very beneficial to the field of education to publish all well-designed empirical studies of high quality, regardless of their findings. Then, practitioners would be better informed and could make better decisions concerning the long-time controversy of whether to retain or not retain low-performing students.

Tying promotion/retention decisions to standardized tests is relatively new. Therefore, very few educational systems have sufficient data to evaluate the impact of grade retention on students' academic performance under such a policy. Since almost all of the current research on grade retention has centered on teacher-initiated retention and academic achievement (Lorence &

Dworkin, 2006), future research is needed to evaluate the relationship of linking retention and promotional standards to high-stakes testing. Such research would be beneficial to educators and policy-makers.

The use of retention has overwhelming negative effects in the literature on academic achievement. However, retention may benefit some students. Future research could identify what type of students benefit from retention. In addition, more long-term studies examining the effects of retention through high school are needed, as well as, the effects of retaining older students and retaining students more than once. Also, because more and more immigrants are coming to the United States to live and work, many English Learner (EL) students are being educated in public schools. Thus, more research on the retention of EL students needs to be conducted.

Summary

This chapter reviewed the major methods used in the study and described the implications of the research findings. The chapter provided an analysis of the findings identified in Chapter IV. The chapter also provided specific data regarding the findings for each of the research questions evaluated in this study. Implications for practitioners and researchers were discussed, along with suggestions for further research.

As a practitioner, this study has been very meaningful to me because grade retention is something that I deal with every year. Most of the time, it is hard to convince teachers that grade retention is not effective for improving students' academic achievement. This is due to the fact that teachers only see short-term results, which are usually positive, and they believe retention works. What they do not know is that most initial gains from retention are lost within two to three years post-retention. Therefore, my personal purpose has been to provide educators,

parents, and policy-makers the evidence from grade retention studies published within the last 20 years so they can make informed decisions about retaining low-performing students.

In conclusion, a common problem to all meta-analyses is the possible existence of studies impacting the question under investigation that were never published nor reported. Published research studies tend to be biased toward positive findings, and studies are often discarded if it is apparent that statistically significant findings will not be forthcoming (Wolf, 1986). Therefore, estimating the number of unreported studies necessary to overturn a significant meta-analytical conclusion (Fail-Safe N) is advantageous. In this study, a “Fail-Safe N ” of 10,890 was calculated. Simply stated, more than 10,000 insignificant studies are needed to disprove the results of this meta-analysis. Thus, it is pretty clear that the results of this study are significant.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analysis.

- Abbott, M., Wills, H., Greenwood, C. R., Kamps, D., Heitzman-Powell, L. & Selig, J. (2010). The combined effects of grade retention and targeted small-group intervention on students' literacy outcomes. *Reading and Writing Quarterly*, 26(1), 4-25.
- Allen, C. S., Chen, Q., Willson, V. L., & Hughes, J. N. (2009). Quality of research design moderates effects of grade retention on achievement: A meta-analytic, multilevel analysis. *Education Evaluation and Policy Analysis*, 31(4), 480-499.
- Bali, V. A., Anagnostopoulos, D., & Roberts, R. (2005). Toward a political explanation of grade retention. *Educational Evaluation and Policy Analysis*, 27(2), 133-155.
- Barnett, K. P., Clarizio, H. F., & Payette, K. A. (1996). Grade retention among students with learning disabilities. *Psychology in the Schools*, 33(4), 285-293.
- *Beebe-Frankenberger, M., Bocian, K. M., MacMillan, D. L., & Gresham, F. M. (2004). Sorting second-grade students: Differentiating those retained from those promoted. *Journal of Educational Psychology*, 96(2), 204-215.
- *Bonvin, P., Bless, G., & Schuepbach, M. (2008). Grade retention: Decision-making and effects on learning as well as social and emotional development. *School Effectiveness and School Improvement*, 19(1), 1-19.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons.
- *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.
- Byrd, R. S., & Weitzman, M. L. (1994). Predictors of early grade retention among children in the United States. *Pediatrics*, 93(3), 481-487.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. Ed.). New York: Academic Press.

- *Cosden, M., Zimmer, J., Reyes, C., & Gutierrez, M. (1995). Kindergarten practices and first-grade achievement for Latino Spanish-speaking, Latino English-speaking, and Anglo students. *Journal of School Psychology, 33*(2), 123-141.
- Cosden, M., Zimmer, J., & Tuss, P. (1993). The impact of age, sex, and ethnicity on kindergarten entry and retention decisions. *Educational Evaluation and Policy Analysis, 15*(2), 209-222.
- Dauber, S. L., Alexander, K. L., & Entwisle, D. R. (1993). Characteristics of retainees and early precursors of retention in grade: Who is held back? *Merrill-Palmer Quarterly, 39*(3), 326-343.
- *Dennebaum, J. M., & Kulberg, J. M. (1994). Kindergarten retention and transition classrooms: Their relationship to achievement. *Psychology in the Schools, 31*, 5-12.
- *Dong, Y. (2010). Kept back to get ahead? Kindergarten retention and academic performance. *European Economic Review, 54*(2), 219-236.
- Ferguson, P., Jimerson, S. R., & Dalton, M. J. (2001). Sorting out successful failures: Exploratory analyses of factors associated with academic and behavioral outcomes of retained students. *Psychology in the Schools, 38*(4), 327-341.
- Glass, G. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*(10), 3-8.
- Glass, G., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*, Beverly Hills, CA: Sage.
- *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. *Elementary School Journal, 107*(4), 327-340.
- *Griffith, C. A., Lloyd, J. W., Lane, K. L., & Tankersley, M. (2010). Grade retention of students during grades k-8 predicts reading achievement and progress during secondary schooling. *Reading and Writing Quarterly, 26*(1), 51-66.
- Guèvremont, A., Roos, N. P., & Brownell, M. (2007). Predictors and consequences of grade retention: Examining data from Manitoba, Canada. *Canadian Journal of School Psychology, 22*(1), 50-67.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.
- Hedges, L. & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.

- 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 on nonpromotion on elementary and junior high school pupils: A mea-analysis. *Reviews 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.
- Hong, G., & Raudenbush, S. W. (2006). Evaluating kindergarten retention policy: A case study of causal inference for multilevel observational data. *Journal of the American Statistical Association*, 101(475), 901-910.
- *Hong, G., & Yu, B. (2007). Early grade retention and children's reading and math learning in elementary years. *Educational Evaluation and Policy Analysis*, 29(4), 239-261.
- 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), 226-244.
- Jimerson, S. R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st Century. *School Psychology Review*, 30(3), 420-437.
- *Jimerson, S. R. (1999). On the failure of failure: Examining the association between early grade retention and education and employment outcomes during late adolescence. *Journal of School Psychology*, 37(3), 243-272.
- *Jimerson, S., Carlson, E., Rotert, M., Egeland, B., & Sroufe, L. A. (1997). A prospective, longitudinal study of the correlates and consequences of early grade retention. *Journal of School Psychology*, 35(1), 3-25.
- *Jimerson, S. R., & Ferguson, P. (2007). A longitudinal study of grade retention: Academic and behavioral outcomes of retained students through adolescence. *School Psychology Quarterly*, 22(3), 314-339.
- *Johnson, E. R., Merrell, K. W., & Stover, L. (1990). The effects of early grade retention on the academic achievement of fourth-grade students. *Psychology in the Schools*, 27, 333-338.

- *Kundert, D. K., May, D. C., & Brent, D. (1995). A comparison of students who delay kindergarten entry and those who are retained in grades k-5. *Psychology in the Schools, 32*, 202-209.
- Lazarus, P. J., & Ortega, P. (2007). Universal pre-kindergarten in conjunction with universal screenings: An antidote to grade retention. *Journal of Educational Research & Policy Studies, 7*(1), 54-75.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.
- Lorence, J. (2006). Retention and academic achievement research revisited from a United States perspective. *International Education Journal, 7*(5), 731-777.
- *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.
- Mantzicopoulos, P. Y. (1997). Do certain groups of children profit from early retention? A follow-up study of kindergarteners with attention problems. *Psychology in the Schools, 34*(2), 115-127.
- *Mantzicopoulos, P. (2003). Academic and school adjustment outcomes following placement in a developmental first-grade program. *Journal of Educational Research, 97*(2), 90-105.
- Mantzicopoulos, P., & Morrison, D. (1992). Kindergarten retention: Academic and behavioral outcomes through the end of second grade. *American Educational Research Journal, 29*(1), 182-198.
- Martin, A. J. (2009). Age appropriateness and motivation, engagement, and performance in high school: Effects of age within cohort, grade retention, and delayed school entry. *Journal of Educational Psychology, 101*(1), 101-114.
- McCombs-Thomas, A., Armistead, L., Kempton, T., Lynch, S., Forehand, R., Nousiainen, S., Neighbors, B., & Tannenbaum, L. (1992). Early retention: Are there long-term beneficial effects? *Psychology in the Schools, 29*, 342-347.
- *McCoy, A. R., & Reynolds, A. J. (1999). Grade retention and school performance: An extended investigation. *Journal of School Psychology, 37*(3), 273-298.
- *McLeskey, J., & Grizzle, K. (1992). Grade retention rates among students with learning disabilities. *Exceptional Children, 58*(6), 548-554.
- *Meisels, S. J., & Liaw, F. R. (1993). Failure in grade: Do retained students catch up? *Journal of Educational Research, 87*(2), 69-77.

- *Moller, S., Stearns, E., Blau, J.R., & Land, K.C. (2006) Smooth and rough roads to academic achievement: Retention and race/class disparities in high school. *Social Science Research*, 35, 157-180.
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: Government Printing Office.
- No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).
- Okpala, C. O. (2007). The perceptions of kindergarten teachers on retention. *Journal of Research in Childhood Education*, 21(4), 400-405.
- Owings, W. A., & Kaplan, L. S. (2001). Standards, retention, and social promotion. *NASSP Bulletin*, 85, 57-66.
- Owings, W. A., Magliaro, S. (1998). Grade retention: A history of failure. *Educational Leadership*, 56, 86-88.
- Pagani, L., Tremblay, R. E., Vitaro, F., Boulerice, B., & McDuff, P. (2001). Effects of grade retention on academic performance and behavioral development. *Development and Psychopathology*, 13, 297-315.
- *Phelps, L., Dowdell, N., Rizzo, F. G., Ehrlich, P., & Wilczenski, F. (1992). Five to ten years after placement: The long-term efficiency of retention and pre-grade transition. *Journal of Psychoeducational Assessment*, 10, 116-123.
- Picklo, D. M., & Christenson, S. L. (2005). Alternatives to retention and social promotion: The availability of instructional options. *Remedial and Special Education*, 26(5), 258-268.
- *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.
- *Reynolds, A. J. (1992). Grade retention and school adjustment: An explanatory analysis. *Educational Evaluation and Policy Analysis*, 14(2), 101-121.
- Roderick, M. (1994). Grade retention and school dropout: Investigating the association. *American Educational Research Journal*, 31(4), 729-759.
- Roderick, M. (1995). Grade retention and school dropout: Policy debate and research questions. *Research Bulletin of Phi Delta Kappa Center for Evaluation, Development, and Research*, 15, 1-5.

- *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.
- Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal*, 32(3), 583-625.
- Shepard, L. S., & Smith, M. L. (1990). Synthesis of research on grade retention. *Educational Leadership*, 47, 84-88.
- *Silberglitt, B., Appleton, J. J., Burns, M. K., & Jimerson, S. R. (2006). Examining the effects of grade retention on student reading performance: A longitudinal study. *Journal of School Psychology*, 44, 255-270.
- *Silberglitt, B., Jimerson, S. R., Burns, M. R., & Appleton, J. J. (2006). Does the timing of grade retention make a difference? Examining the effects of early versus late retention. *School Psychology Review*, 35(1), 134-141.
- Smink, J. (2001). Alternatives to retention. *NASSP Bulletin*, 85, 3-17.
- Tanner, C. K., & Galis, S. A. (1997). Student retention: Why is there a gap between the majority of research findings and school practice? *Psychology in the Schools*, 34(2), 107-114.
- Tomchin, E. M., & Impara, J. C. (1992). Unraveling teachers' beliefs about grade retention. *American Educational Research Journal*, 29(1), 199-223.
- Willson, V. L., & Hughes, J. N. (2006). Retention of Hispanic/Latino students in first grade: Child, parent, teacher, school, and peer predictors. *Journal of School Psychology*, 44, 31-49.
- *Willson, V. L., & Hughes, J. N. (2009). Who is retained in first grade? A psychosocial perspective. *The Elementary School Journal*, 109(3), 251-266.
- Witmer, S. M., Hoffman, L. M., & Nottis, K. E. (2004). Elementary teachers' beliefs and knowledge about grade retention: How do we know what they know? *Education*, 125(2), 173-193.
- Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Newbury Park, CA: Sage.
- *Wu, W., West, S. G., & Hughes, J. N. (2008a). Effect of retention in first grade on children's achievement trajectories over 4 years: A piecewise growth analysis using propensity score matching. *Journal of Educational Psychology*, 100(4), 727-740.

*Wu, W., West, S. G., & Hughes, J. N. (2008b). Short-term effects of grade retention on the growth rate of Woodcock-Johnson III broad math and reading scores. *Journal of School Psychology, 46*, 85-105.

APPENDIX A

IRB Exclusion Letter

Office for Research
Institutional Review Board for the
Protection of Human Subjects

June 6, 2011

THE UNIVERSITY OF
ALABAMA
R E S E A R C H

Anthony Bright
ELPTS
College of Education
The University of Alabama

Re: IRB Requirement for Doctoral Dissertation Research

Mr. Bright:

This letter comes as a response to your communication received June 3, 2011. Following review it has been determined that activities outlined within the project description do not meet the criteria for human subjects research as set forth within UA IRB Form #31 titled "Human Research Determination Checklist".

Because the activity is not considered to be research involving the use of human subjects, the activity does not require IRB approval and is therefore excluded from review by the IRB. If you have any questions or if I can be of further assistance please do not hesitate to contact me.

Sincerely,

Carpantato T. Myles/MBM, CM
Director of Research Compliance & Research Compliance Officer
Office of Research Compliance
The University of Alabama



152 Rose Administration Building
Box 8701 I7
Tuscaloosa, Alabama 35487-01 17
(205) 348-5152
FAX (205) 348-8882

APPENDIX B

Initial Screening Instrument

ST-01

Johnson, E. R., Merrell, K. W., & Stover, L. (1990). The effects of early grade retention on the academic achievement of fourth-grade students. *Psychology in the Schools*, 27, 333-338.

Participants	Methods	Findings
<p>This study investigated the academic effects of early (kindergarten or first-grade) retention on a group of fourth-grade students who had been retained at the K-1 level ($n = 20$, 14 males, 6 females).</p> <p>Two comparison groups were used:</p> <ol style="list-style-type: none"> 1. Fourth-grade students recommended for retention at K-1 level but not retained ($n = 17$, 8 males, 9 females). 2. Fourth-grade students who had made normal progress through the grades ($n = 20$, 11 males, 9 females). <p>Total of 57 fourth-grade students from 4 public school districts in a western state.</p> <p><i>Gender:</i> 33 males 24 females</p>	<p><i>Statistical Tests:</i></p> <p>MANOVA, One-way ANOVA, Post-hoc pairwise comparisons (Scheffe')</p> <p>Effects reported in standard deviations.</p> <p><i>Data Collection:</i></p> <p>Cumulative academic file of each subject was examined to obtain standardized academic achievement data.</p> <p>Metropolitan Achievement Tests, Sixth Edition (MAT)</p> <p>MAT content/subject areas used included Total Reading, Total Math, Reading Comprehension, Math Calculation, Spelling, Vocabulary, and Language Arts.</p>	<p>No significant differences in academic achievement level were found between the retained and recommended-for-retention but not retained groups. However, both groups were significantly lower on several academic achievement measures than their normal comparison peers.</p> <p>The results of this study indicate that grade retention as an academic intervention was ineffective.</p>