

NAVIGATING THE COMPLEXITIES OF LEGISLATION: HOW ELEMENTARY
SCHOOL PRINCIPALS INTERPRET AND IMPLEMENT FLORIDA'S
THIRD-GRADE RETENTION POLICY

By

COURTNEY CALDWELL ZMACH

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2006

Copyright 2006

by

Courtney Caldwell Zmach

ACKNOWLEDGMENTS

Many people contributed to my educational background, to shape the educator and researcher that I have become. It is difficult to know where to begin, because each is important. I thank my steadfast supervisory committee chair (Dr. Richard Allington), cochair (Dr. Anne McGill-Franzen), and members (Dr. Gregory Camilli, Dr. Thomas Dana, and Dr. Sandra Russo). Each of them supported my scholarship and development in becoming a reading researcher. Dissertations can be a lonely journey and I thank those who spent time listening to me. At different stages and in different ways, Dick Allington, Anne McGill-Franzen, Greg Camilli, Tom Dana, Sandra Russo, Jenn Graff, Evan Lefsky, Jeff Miller, Katie Solic, and Lunetta Williams provided listening ears. I value their time, suggestions, and unwavering support.

My thanks go to the many members of the Just Read, Florida! Office and the Florida Department of Education for taking time to answer questions and for gathering data files for my study. I am indebted to Hajnalka Peto and Jenny Williams for the time they spent transforming state data files, so I could create the databases used in my study.

I thank my parents and family, who supported my decision to uproot my life as a classroom teacher, to return to school. From as early as I can remember, my parents shared with me the value of education. Their support and encouragement led me this far. My former colleagues and students deserve thanks for inspiring me to reach for higher educational goals. I also express gratitude to my new colleagues at the American Institutes for Research (AIR), for supporting my continued scholarship.

Of course, this research would not have been possible without the cooperation of the twelve Florida school districts. Hurricane Season 2004 was an especially trying time for Floridians. Special thanks go to the elementary school principals who gave their time to participate in this study.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS	iii
LIST OF TABLES	vii
LIST OF FIGURES	viii
ABSTRACT	x
CHAPTER	
1 INTRODUCTION TO STUDY	1
Introduction.....	1
Policy Problem.....	2
Study Approach	3
Perspectives	4
Purposes of this Study	5
2 REVIEW OF THE LITERATURE	6
Introduction.....	6
Research on Retention	7
High-Stakes Framework	18
Florida Policy Context.....	31
Summary	46
3 METHODS	48
Introduction.....	48
Participants	49
Data Collection	51
Procedures.....	54
Design	63
4 RESULTS	76
Introduction.....	76
Multilevel Models.....	76

Summary.....	85
5 DISCUSSION.....	87
Introduction.....	87
Research Question 1	87
Research Question 2	89
Research Question 3	91
Policy Considerations	97
Policy Recommendation.....	104
Future Research	107
Conclusion.....	109
APPENDIX	
A FLORIDA PRINCIPAL SURVEY	111
B REVIEW OF SURVEY PILOT STUDY	120
C DATA SOURCES	122
D SAMPLE FROM STATISTICAL SYNTAX.....	123
E RESULTS TABLE FOR ALL MODELS	124
LIST OF REFERENCES.....	126
BIOGRAPHICAL SKETCH	137

LIST OF TABLES

<u>Table</u>	<u>page</u>
2-1 Florida retention policy good cause statewide exemptions.....	36
2-2 Supports identified in Florida’s legislative intent	41
3-1 Time in years: Pre-policy versus post-policy	49
3-2 Degree of support category, <i>Florida Principal Survey</i>	54
3-3 Survey response rates by district size.....	57
3-4 Usable sample versus returned responses by district size	60
3-5 Third-grade percent retained usable sample principals versus the total population.....	61
4-1 Results of fitting a taxonomy of multilevel models for change to the log percent retained data, Models A through E ($n = 102$).....	86
5-1 Third-grade FCAT reading achievement Level 1	99
C-1 Sources of data	122
E-1 Results of fitting a taxonomy of multilevel models for change to the log percent retained data, Models A through F2 ($n = 102$).....	125

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
2-1 High-stakes framework	18
3-1 Item 9, <i>Florida Principal Survey</i>	52
3-2 Item 14, <i>Florida Principal Survey</i>	53
3-3 Item 19, <i>Florida Principal Survey</i>	53
3-4 Employment history, <i>Florida Principal Survey</i>	54
3-5 Population versus usable sample mean percent retained by year	61
3-6 Mean SES values over 5-year period for usable sample schools	62
3-7 Mean retention rates by year	65
3-8 Examples of principals' third-grade retention rates over time	66
3-9 Third-grade retention over time	67
3-10 Histogram of percent retained before the transformation	70
3-11 Histogram of percent retained after the transformation	70
3-12 Examining the homoscedasticity assumption	71
3-13 Multilevel model for change taxonomy	73
3-14 Number of zero percent retained values removed by year	74
4-1 Scatterplot of log percent retained plus one	78
5-1 Examples of mean predicted retention trends over time	88
5-2 Mean predicted third-grade retention rates by poverty level over time	90
5-3 Example plotting retention rates using principals' mean SES values before and after policy implementation	90

5-4	Mean predicted third-grade retention rates by principals' current retention belief over time.....	92
5-5	Principals' retention belief by the mean predicted third-grade retention rates over time disaggregated by poverty group.....	95
5-6	Mean predicted third-grade retention rates by degree of support over time	96
5-7	Degree of support by mean predicted third-grade retention rates over time disaggregated by poverty group	98

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

NAVIGATING THE COMPLEXITIES OF LEGISLATION: HOW ELEMENTARY
SCHOOL PRINCIPALS INTERPRET AND IMPLEMENT FLORIDA'S
THIRD-GRADE RETENTION POLICY

By

Courtney Caldwell Zmach

May 2006

Chair: Richard L. Allington
Cochair: Anne McGill-Franzen
Major Department: Teaching and Learning

Retention policies affecting third graders have become increasingly common as schools, districts, and states across the country work to comply with the federal mandate for all third-grade students to read at or above grade level by 2014. Florida is one such state to enact a high-stakes retention policy as a way to meet accountability challenges. Third graders in Florida who fail the state assessment, the Florida Comprehensive Assessment Test (FCAT), are held back in grade.

This exploratory study gained understanding of Florida's third-grade retention policy by focusing on the elementary school principal ($n=102$) as the unit of analysis. To examine the Florida policy, it seemed useful to better understand patterns of third-grade student retention over a 5-year period, retention beliefs of principals, and degree of support given to students affected by the policy. The overarching research question of this study was "What is the trend of third-grade retention practices before and after

implementation of Florida's third-grade retention policy and how is this trend impacted by other variables?" Using a combination of items from a researcher-designed survey of elementary-school principals across twelve districts, and data provided by the Florida Department of Education, I used a multilevel statistical model for change examining within-person change and between-person differences in change, beginning with the 1999-2000 school year third-grade retention rates. Study variables also included measures for pre- and post-policy, poverty, trend within time, retention belief, and degree of support.

Results show that the trend within time and poverty at the school level dramatically affected retention rates across this sample. Students in higher-poverty schools were impacted most greatly by the policy. In other words, lower achieving students in higher-poverty schools were more likely to be promoted pre-policy and because of this, the mandate had more impact on students in higher-poverty schools than in lower-poverty schools. Regardless of the actual impact on reading achievement, Florida's retention policy has eliminated social promotion in third grade.

CHAPTER 1 INTRODUCTION TO STUDY

Introduction

Even with the national agenda to “leave no child behind,” substantive segments of the student population are held back in grade because of mandatory retention policies. Today, thousands of students in Chicago, New York City, and across Texas and Florida (and elsewhere) have been affected by the use of high-stakes assessments to make promotion and retention decisions.

Retention in grade (sometimes called nonpromotion, being held back, flunking, or failing) is a policy that retains a student in the same grade until the student meets requirements to move to the next grade (Jackson, 1975; Jimerson, 2001; Shepard & Smith, 1989). Retention is not a new idea: it is an old and highly contested practice that has fluctuated across history. Earlier research indicated reasons for retaining students, including low performance on standardized assessments, youngness for the grade, physical size, immaturity, and even poor attendance. Today, some believe the conventional wisdom that if a child does not fare well on a standardized assessment, then that child is ill prepared for the next grade level. Once again, retention has resurged because of mandates tied to meeting accountability criteria.

The antithesis of retention is social promotion, where students are promoted from grade to grade, with same-age peers. Social promotion is something of an enigma, as many state departments of education have not collected information determining how widely it has infused the system. Equally troubling, limited information is available

regarding how retention decisions are made, or whether standardized assessments are a determinant for retaining students (U.S. Department of Education, 1999). In the Council of Chief State School Officers (CCSSO) 2001-2002 Survey of Student Assessment Programs (SSAP), 10 states reported that standardized assessments played a role in retention decisions (personal communication, September 14, 2004). However, we know that this number has grown, as Florida (in 2003) joined the ranks using reading achievement scores from the Florida Comprehensive Assessment Test (FCAT) for promotion and retention decisions.

Policy Problem

What truly is the crux of the problem at hand? Not performing on grade level is seen by some as a social problem (Oakes, 1999). To fix this perceived problem, lawmakers instituted what they believed to be a solution—retention in grade. If the children did not learn to read at a proficient level the first time in third grade, will another year in third-grade be the solution? It would appear that the policy is working when more students achieve a passing score. Retention in grade has been researched for decades and proven ineffective (Allington & Walmsley, 1995; Shepard & Smith, 1989); yet Florida persisted in enacting a retention policy to remedy the perceived social problem.

The challenge we face is more than whether failing a student is justified. It comes down to how we, as educators and researchers, can most effectively help students who struggle with reading, especially those who fail to make a passing score on high-stakes assessments. Nevertheless, retention policies tied to reading achievement are becoming increasingly common as schools, districts, and states across the country contend with the No Child Left Behind of 2001 mandate for all third-grade students to read at or above

grade level by 2014 (US Department of Education, 2001). Florida is one such state to enact a high-stakes retention policy to meet the challenges of accountability.

One wonders how this retention policy benefits Florida's students. As reported in a dataset provided by the Florida Department of Education (FLDOE), over 27,000 of Florida's third-grade students were impacted by the retention policy after the 2002-03 school year. Of this number, after 1 year of retention as part of the State's chosen intervention method to help increase student achievement, only 59% passed the FCAT after their 2nd year in third-grade (Florida Department of Education, 2004c). There are a great number of areas awry in education today, but retention used as a reading intervention is questionable. State-level policymakers believed that Florida's retention policy (F.S. 1008.25) would help mend the perceived problem of third graders not reading on grade level, but the research base does not support that belief.

Study Approach

In the early stages of [policy] implementation, summative measures usually are inappropriate. More appropriate questions for analysis involve the extent to which necessary resources are available to support implementation, whether there is evidence of good-faith efforts to learn new routines, or indication of commitment and support within the implementing system for policy strategies and goals.

—Milbrey McLaughlin (1987, p. 176)

On considering how best to capture the spirit of my study, I decided to collect data using a survey questionnaire, and to use historical retention data and poverty indicators provided by the FLDOE. My exploratory study aimed to gain understanding of the interpretation and implementation of Florida's retention policy by focusing on the elementary school principal as the unit of analysis. To illuminate the current situation in Florida, it seemed useful to better understand

- Patterns of third-grade student retention over a 5-year period

- Impact of school-level poverty
- Other variables such as a principals' retention belief and the degree of support provided for students affected by the new state policy

The overarching research question of my study was “What is the trend of third-grade retention practices before and after implementation of Florida’s third-grade retention policy and how is this trend impacted by other variables?” To answer this question, I designed a multilevel model considering factors such as the school’s level of poverty, while determining whether past retention practices appear to influence the level of current support offered to students at risk for retention or currently retained.

Perspectives

The worldview of the researcher bases socioeconomic status (SES) as a predictor for students’ educational successes. “[Poverty] is a condition, like gravity, that affects virtually everything” (Bracey, 2003, p. 46). Foci throughout my study presuppose that economic disadvantage places many schools and their students at risk for failure. The correlation between student achievement and poverty level is widely accepted (Linn, 2004).

Florida elementary school principals were selected to gather information first-hand, from a range of schools across the state. Principals are vital stakeholders in making decisions at the school level, and being informed about the happenings in their schools. Principals’ historical retention practices were built into my study design to examine retention over time in relation to retention belief. Selection of principals in no way suggests that principals deserve sole responsibility for matters concerning retention; however, as the figurehead of leadership in a school, they serve an important role.

Purposes of this Study

Results from my study serve multiple purposes. As a result of the third-grade retention policy, thousands of third-grade students have been retained or re-retained (Florida Department of Education, 2004c); however, this policy has brought forward several mandates and initiatives geared to support students who struggle with reading. Because of the mandate to use particular supports, examining the relationship between degree of support available and retention rates may help current principals assess (or reassess) their practices to provide the interventions proven to benefit students at risk for failure.

Stakeholders from school, district, and state levels may have a clearer vision of how this retention policy has been enacted. We know that beliefs held by those who implement policy play an important role (McGill-Franzen, 2000; Spillane, 1998); therefore, we must understand how these beliefs affect the retention policy. Regardless of whether one concurs with or opposes this policy, it is important to understand whether retention beliefs held by administrators across the State impact the proportion of students retained, or even how supports are distributed to help students. This research may help administrative leaders (and district and state leaders) consider patterns of retention in a different way. Results are unique to Florida and may not be generalizable to other states; however, other states practicing retention or considering implementing a retention policy might gain insights that may benefit their locale.

CHAPTER 2 REVIEW OF THE LITERATURE

Introduction

Today in Florida, a retention policy prevents third-grade students from progressing to the fourth grade when they fail to master grade-level standards (or expected learning) on the state assessment, the FCAT. Given this policy impact on students, why is Florida's retention policy (Section 1008.25, Florida Statutes) worth study from the vantage point of an elementary school principal? As the primary instructional leader of a school, the principal has the extraordinary task of remaining abreast of policy changes, while communicating vital information to staff members and parents. The school principal serves as a gatekeeper of information for the instructional staff, and for the students they serve and protect. The principal is largely accountable for all the happenings in his or her school. In the hierarchical system instituted in districts across Florida, it is important to understand how administrators (in this case, principals) support those students at risk for retention or already retained in grade.

Over the past decades, other investigations studied the effects of retention policies, including student achievement, dropout rates of students affected by retention decisions, and beliefs about retention (Shepard & Smith, 1989). Since Florida implemented the retention policy in 2003, ending social promotion for third graders who fail to read on grade level, many facets are worth studying that can add to and extend the existing body of knowledge. While one can choose to focus on retention at the classroom level (examining issues related to students, teachers, or classroom instructional practice), I

believe the principal holds direct influence on the aforementioned. Few studies have examined the role of principals and how they enact retention policies. Since the principal acts as an intermediary or messenger disseminating information from the district (or state) to the classroom level, how a principal interprets and implements this retention policy is vital to our understanding and evaluation of what makes a difference for those at risk for retention.

Research on Retention

First, we must consider why retention in grade has resurfaced as a commonly accepted practice, or intervention of choice, in states and districts across our country. To understand where we are today, let us go back and review a brief history of how educators, in the United States, came to practice retention.

In the 19th century, as America developed and expanded its public school system, formalized instruction typically occurred in ungraded one-room schoolhouses with students ranging in age (Ruhl, 1984), where students progressed through the education system learning content. By the mid-nineteenth century, the European education system began to influence multiple aspects of the United States' education structure (Balow & Schwager, 1992; Ruhl, 1984). Soon, many facets of the United States' system became graded—everything from schools, students, curricula, teachers (Balow & Schwager, 1992). A graded, or sequential system, turned learning into discrete parts, as students were held accountable to learn specific information that built upon one year to the next.

As student learning became linked to yearly increments, the introduction of textbooks helped standardize the curriculum (Jackson, 1975). With the use of textbooks, it became clear that students did not learn at the same rate. Hence, originated one consequence when textbook content was not mastered—retention in grade. For students

who did not master the material, this solution required them to repeat the same grade level to continue learning that material (Holmes & Matthews, 1984; Jackson, 1975).

Decades later, Leonard Ayres described in his book *Laggards in our Schools* (1909) a trend that Dr. William Maxwell, Superintendent of the New York City schools, reported in 1904. Maxwell observed discrepancies in the normal ages of students within the New York City school system. He calculated a staggering 39% of children in the elementary grades over the standard age for a given grade. This inconsistency between age and grade when students fall backward in grade was termed “retardation” (Ayers, 1909). This fluctuation was the result of retention in grade and research ensued to learn more about this phenomenon.

One of the first retention studies, The Backward Children Investigation, led by Ayers in 1907 studied children who were not progressing through school at typical rates—those who were over-age for the grade. Ironically, Ayers’ areas of inquiry are quite similar to present-day wonderings—learning more about differences in student characteristics, the age of students when starting school, teacher effectiveness, and curriculum differences. His research questions included, “How many of the children in our schools fail to make normal progress from grade to grade and why do they fail? How many of the children drop out of school before finishing the elementary course and why do they drop out of school before finishing the elementary course and why do they drop out? What are the facts and what are the remedies?” (Ayers, 1909, p. 2).

Ayers calculated the “survival” and “mortality” rates of students by establishing the proportion of students who remained in the school system without dropping out. Based on his calculations, most students remained in the system through the end of grade school

with a severe drop off in attendance during the junior high years and few retainees remained to attend high school. The retained were more likely to leave school before their promoted same age peers. His findings established that students, who do not succeed, simply drop out (Ayers, 1909).

Evidence has mounted for nearly a century that retention does not benefit students. In the 1930's and 1940's, researchers conducted studies dispelling the notion that retention helps children. In fact, Scott and Ames (1969) reported that the "negative findings on the effects of non-promotion were so uniform in the 1930's and 1940's that many investigators considered the question closed..." (p. 433). However, Goodlad (1952) revisited retention when it surfaced as a hot topic of the day with educators asking the question "What is best for the development of this child?" (p. 150).

Over fifty years ago, Goodlad recognized that retention was not a one-size-fits-all solution asserting that "promotion on the basis of fixed minimum standards is not adequate" (Goodlad, 1952, p. 150). He tested hypotheses dealing with differences in social and personal adjustment between children who repeated a grade and those who did not repeat. Based upon his findings, he found promotion to be a more justifiable educational practice, taking into account the best interests of the child. Goodlad recommended that 1) children be treated as individuals, not subjected to system-wide policy, 2) teachers use facts related to achievement, intelligence as well as human growth and development in decision making, and 3) "[the] instructional needs of the pupil should take precedence over matters of administrative expediency in dealing with questions involving promotion and nonpromotion" (p. 154).

Using Goodlad's (1952) recommendations, Lobdell (1954) conducted a study where one school district adhered to criteria guiding both teachers and principals in making retention decisions for low achieving students. Criteria for these decisions were both general and specific. Briefly, general criteria included what would be best long-term for the child, including 1) holding a child back only one time before sixth grade, 2) avoiding repetition of sixth grade and finally, 3) children should not spend the repeated year with the same teacher as the previous year. Specific criteria suggested that many aspects be taken into account beyond the score on a standardized assessment, including current grades, intelligence scores, age, size, social characteristics and the attitude of the parents toward the child's progress (Lobdell, 1954).

Essentially, all children were to be treated individually when making retention decisions. Lobdell (1954) asserted that "holding back a child for two or three or more years beyond his normal graduation age, as practiced in past generations, can in no way be defended" (pp. 335-336). Both the teacher and the principal were involved in making these decisions; however, the role of the principal was more of a judge. Principals insured that the facts were present and that the best interests of the child were taken into account (Lobdell, 1954). Retained students who were subjected to these criteria showed short-term success, however the long-term effects were not discernable from Lobdell's available data. Like Goodlad, Lobdell (1954) concluded that broad policies for retention needed to be replaced with criteria to guide the decision making process, as this was believed to be in the best interests of the child.

Changes in promotional policies was the subject of the Hall and Demarest (1958) piece in *The Elementary School Journal*. They identified three levels of promotional

policy: 1) **Grade-standard**: students learned a predetermined amount to move to the next grade, 2) **Continuous-promotion**: as in social promotion, where students move through the grades with same-age peers, and 3) **Continuous-progress**: students are retained based upon individual retention decisions (Hall & Demarest, 1958). This study examined the policy shift in Phoenix, Arizona treating children as individuals to make retention decisions. The move to “continuous-progress” was seen as a positive shift for children (Hall & Demarest, 1958).

Research on retention was sparse in the 1960s. However, as the decade neared an end, Chase (1968) and Scott and Ames (1969) published their work on the subject. Both studies omitted students retained for reasons such as low intelligence, brain damage, or emotional disturbance as these were not viewed as necessarily correctable via retention and included only those who were immature because of youngness or behavior. Findings from these studies supported retention as a means to increase maturity.

Nonetheless, over the years, research has consistently shown that retention causes negative long-term consequences. Efforts to increase achievement by means of retention simply present more hurdles for the student to overcome. Grissom and Shepard (1989) created a structural equation model, or causal model, in hopes of explaining the causes for dropping out of school. Although they found that substantially more students who were retained drop out of school early compared to their promoted peers, other factors, in addition to flunking, contributed to the dropping out, such as the need for the student to work or attend to other family matters. Research on dropping out of school has been studied in more recent times with similar results to Ayers (Roderick, 1994). Roderick (1994) also found that the retained were more likely to dropout leaving school early.

There are other long-term educational outcomes worth consideration. Retained students are more likely to be placed in special education programs (Barnett, Clarizio, & Payette, 1996; Guthrie, 2002; McGill-Franzen & Allington, 1993), and students placed in special education or other remedial programs rarely escape their special education label (McGill-Franzen, 1987; McGill-Franzen & Allington, 1993).

Retention Research in Chicago

The Consortium on Chicago Schools Research at the University of Chicago studying the implementation of Chicago Public Schools' 1996 retention policy provides a wealth of insight into the effectiveness of that policy (e.g., Roderick, Bryk, Jacob, Easton, & Allensworth, 1999; Roderick & Engel, 2001; Roderick, Nagaoka, Bacon, & Easton, 2000). When students in Chicago did not meet promotional gate cut-off scores on the standardized assessment, Iowa Tests of Basic Skills (ITBS), they were held back. Promotional gates in Chicago occur at third, sixth and eighth grades.

After analyzing 2 years of retention records, these researchers determined that many children were being placed in special education programs after the retention decision was reached. In fact, they calculated that between 17 and 20% of retainees were recommended for special education placements (Nagaoka & Roderick, 2004). One caveat Nagaoka and Roderick (2004) acknowledged was the difficulty discerning whether these students had unidentified difficulties prior to the policy or whether these students were being pushed into special education programs to improve accountability reports. Both explanations are plausible (Allington & McGill-Franzen, 1992; McGill-Franzen & Allington, 1993). In addition, one must strongly consider that lack of teacher expertise is another possible explanation for these referrals. When teachers use

all they have in their bag of tricks, they commonly turn elsewhere for guidance—special education is an intuitive choice.

Some retained students did not meet the criteria to move through the promotional gates after their second attempt. These students remained in the same grade for a third year in the same grade; these students are called the double retainees. The Chicago team worked to determine whether the double retainees were better off after these retentions (Roderick et al., 2000). After equating the scores of the third-grade ITBS to the fourth- and fifth-grade tests, the researchers found little conclusive evidence to suggest that double retentions benefited Chicago's third graders. In fact, after three years in third grade, about 80% were promoted to fourth grade, while about 10% of students were placed in special education programs.

Thus far, results from the study in Chicago confirm past research efforts that retention does not work. In fact, they have found “little evidence that students who were retained did better than their low-achieving counterparts who were promoted” (Nagaoka & Roderick, 2004, p. 45). Retained students who were offered additional chances to meet the promotional criteria struggled to reach the cut point. Consortium researchers support early intervention; however, they caution that this does not mean high-stakes assessments need to be moved to earlier grades (Roderick et al., 1999).

Reviews of Retention Research

Jackson (1975) presented the first review of research on retention by dividing thirty studies published between 1911 and 1973 into three groups according to the research design. Based on his analyses of naturalistic, pre-post and experimental design, he found

that retention may benefit some students, although more students seem to benefit from promotion. Jackson also suggested that the retention is unfounded as an intervention.

Next, we examined evidence from using meta-analysis, a research method that takes into account the quantitative outcomes of multiple research studies (Cooper & Harris, 1994; Light & Pillemer, 1984). The steps of meta-analysis are important to evaluate highly contested issues, as it helps compensate for biases in the sample and findings (Camilli, Wolfe, & Smith, in press; Holmes & Matthews, 1984). One benefit of meta-analysis, and the reason why these studies are presented here, is that this technique is much more “comprehensible to the reader than lengthy recounting of each individual study’s methods and results” (Shepard & Smith, 1989, p. 16).

Two noteworthy meta-analyses conducted over a decade ago aimed to learn more about the effects of retention on students (Holmes, 1989; Holmes & Matthews, 1984). The 1984 meta-analysis systematically reviewed 44 studies that fit the selection criteria. Essentially, research studies needed to have specific features for the meta-analytic calculations. Holmes and Matthews (1984) required original studies to contain these criteria: 1) report the effects of retaining students in grade school or junior high, 2) provide enough data to calculate effect sizes, and 3) compare retained students to promoted students. After statistically combining the results of these studies, overwhelmingly the evidence confirmed that retention does not work ($ES = -.37$). The sign of the effect size indicates whether combined outcomes of the studies had a positive or negative effect on students. An effect size of $-.37$ suggests that the negative effects of retention far outweigh the positive.

Only 5 years after Holmes and Matthews published these results, Holmes (1989) performed another meta-analysis using the original forty-four studies along with an additional nineteen studies conducted since that review. In the 1989 analysis, 9 studies showed positive effects, benefiting the retained. Although the overall effect size was slightly smaller in the 1989 analysis ($ES = -.15$ vs. $ES = -.37$), the overall direction of the results confirm that retention does not support the development of students academically or personally (Holmes, 1989). Additionally, Holmes (1989) conducted a secondary analysis to ascertain why there was a difference between these effect sizes merely 5 years later. He found that studies fell into two types, positive and negative. Most of the positive studies were conducted in suburban schools with few minority students, contained remediation plus intervention, and compared grade level peers as opposed to same age peers. When studies were matched on IQ, achievement tests, socioeconomic status, gender, and grade, the effect size was $-.30$, after controlling for possible differences in the samples.

More recently, Jimerson (2001) conducted a meta-analysis examining the results of studies dating from 1990 to 1999. Compared to the meta-analyses cited above, selection criteria for this analysis included accepting only studies with matched comparison groups. With an effect size equaling $-.31$, meta-analytic methods repeatedly substantiate that retention is not an effective method to increase student achievement (Jimerson, 2001). In his conclusion, Jimerson (2001) recommended moving beyond questioning whether to retain students and to move forward considering which interventions are effective to remediate students who struggle academically.

Retention Trends

Studies exist documenting trends of retention practices over time. Of particular interest here is Allington and McGill-Franzen's study (1992) investigating school-extending practices in the New York State. These researchers wondered whether there was a shift in retention to Grades K to 2 in an effort to delay the predicted third-grade retention. This clever move allowed a delay in public accountability reports (Allington & McGill-Franzen, 1992). Early identification of students in Grades K to 2 who would likely be retained in third grade were held back in earlier grades in efforts to remediate them prior to the mandated flunking in third-grade, *if* they failed the state assessment.

Retention Beliefs

“The contradiction between research and actual classroom practice is deep-seated in a belief system that has been delivered over the past 90 years” (Reitz, 1992), or over the past 100 years now. Two studies included here, focus on educators' beliefs about retention, were conducted during an era when “social promotion” was a more acceptable option. Byrnes (1989) conducted a study on the attitudes of students, parents and educators on the practice of retention through two means, survey and interviews, in a large city in the southwestern United States. We concentrate here on the questionnaire sent to forty-five principals and assistant principals with a 78% return rate. Findings from the principal questionnaire determined that most principals (74%) favored retention for students who were not meeting the grade level standard. Principals also favored immaturity (54%) as a cause for holding students back in grade. Contrary to popular belief, in that study retention was not seen in a negative light. Significant differences were detected when respondents were asked who should have the final judgment in retention decisions. Teachers felt this was their responsibility; whereas principals felt

this fell in their domain. Alternatives to retention were explored with teachers and principals choosing options such as smaller class sizes and more individualized instruction compared to other ideas that they perceived as changing the current school practices. For instance, options not selected included “flexible entry age, transitional maturity classes, and multi- and non-graded school structures” (Byrnes, 1989, p. 114). Conclusions from this study showed that teachers, principals and parents believe retention benefits those students not performing on standard or who were thought to lack maturity. Byrnes (1989) described retention as an intuitive choice, even though research still did not conclusively suggest a benefit for students.

Another study using a multi-method approach focused on the attitudes of teachers of grades K to 7 (Tomchin & Impara, 1992). Using a questionnaire and interviews, Tomchin and Impara found that primary grade teachers (grades K to 3) are more likely to retain students, compared to teachers from middle grades (grades 4 to 7). Primary grade teachers do not believe harm will follow a retention decision, whereas teachers in older grades were more skeptical. On the questionnaire, almost 98% of teachers disagreed with the statement that “Children should never be retained” (Tomchin & Impara, 1992). As principals are typically former classroom teachers, it is of interest in the present study to determine the retention beliefs held by principals, especially now that social promotion is constrained by policy. “Parents, teachers, and principals seem to play a crucial role in the decision-making process, and generally a veto from any of these consultants can result in promotion instead of retention, regardless of performance on competency tests” (Nikalson, 1984 and Rose et al., 1983 as cited in Jimerson, Carlson, Rotert, Egeland, & Sroufe, 1997, p. 4). In Florida, a veto is not possible.

High-Stakes Framework

To understand my policy study, one must be familiar with the high-stakes framework (Figure 2-1) within which retention is situated. Also, as students' reading success is pivotal to the national goal for students to read on grade level by 2014, we see how retention interconnects with the literature of educational reform, accountability, and high-stakes standardized assessments. We begin with a review of the essentials to understand this high-stakes framework beginning with an overview of national educational reform exploring past and present movements focusing on standards-based reform, seeing how reform entwines with accountability and standardized assessments. Next, we examine the uses and possible misuses of standardized assessments and the relationship to retention. Finally, by focusing on how policy is interpreted and implemented, we work toward a common understanding. Adding to the background of my study, efforts were made to focus on elementary school principals and the State of Florida within each area of the framework.

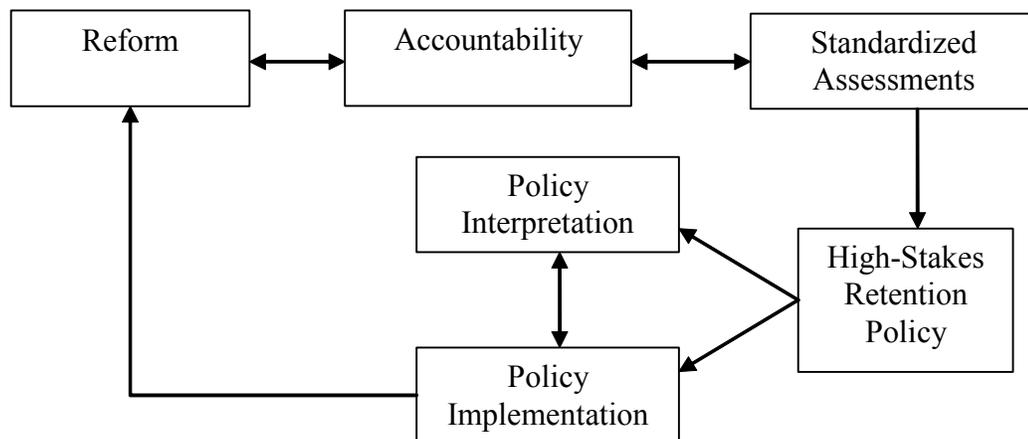


Figure 2-1. High-stakes framework

Reform Movements

In considering where we are today, let us remember past policy, or “first generation” policy (McLaughlin, 1992 as cited in McGill-Franzen, 2000, p. 800). The current talk about educational equity, and raising achievement are not new ideas. These have been policy goals for at least 40 years, when the federal government passed the Elementary and Secondary Education Act (ESEA) of 1965, in particular Title I, aimed to improve the education for students from impoverished backgrounds (McGill-Franzen, 2000). For 40 years, high-need schools enrolling at least 40% of families from economically disadvantaged backgrounds have received funds from the largest federal funding program, Title I. Economic disadvantage is determined by the percentage of students who are eligible for a free or reduced price lunch. One goal of Title I funding has been to close the achievement gap by providing all students with the opportunity to attain a high-quality education, regardless of their economic disadvantage or advantage. Educational attainment for all students was on the national educational agenda. With reform movements continuing to permeate the United States’ educational system and gain strength, there is impetus for broad, sweeping transformations and having them happen instantaneously. Reforms are one method of instituting change.

A Nation at Risk: The Imperative for Educational Reform is widely regarded as the origin of modern American educational reform (National Commission of Excellence in Education, 1983). However, such reports moved us into the next generation of education policy (McGill-Franzen, 2000). It called for the American public to restructure the educational system and raise standards (National Commission of Excellence in Education, 1983). Specifically, one recommendation applied to retention in grade. Recommendation C, Number 8 suggested redesigning the grade leveling system

lessening importance on age restrictions for placing and promoting students, instead it emphasized that academic progress and students' needs should serve more to guide promotion and graduation decisions (National Commission of Excellence in Education, 1983).

Politically, support for retention appears to be bipartisan. By the late 1990s, many lawmakers took note of the commission's advice and scrutinized students' academic progress. In his 1998 State of the Union Address, then-President Clinton "joined a host of other political leaders, from the Democratic mayor of Chicago to the Republican governor of Texas, all calling for an end to the promotion of students whose achievement does not meet the expectations for that grade" (National Research Council, 1999, p. 41). We note here that when President Clinton publicly denounced "social promotion" again in his 1999 State of the Union Address, he also praised the efforts of Chicago Public Schools for raising student achievement when it ended social promotion (Clinton, 1999). Although five years later, it remains to be seen whether the Chicago policy has helped the students held back in grade (Nagaoka & Roderick, 2004).

The United States Department of Education (hereafter, USDOE) later released the document, *Taking Responsibility for Ending Social Promotion: A Guide for Educators and State and Local Leaders* (1999), where President Clinton directed states and localities to practice retention as a way to increase student achievement. Here, he outlined the steps and rationale for ending social promotion, or the practice of simply moving children from grade-to-grade with their age cohort. To accomplish this goal, President Clinton described a plan that included rigorous standards aligned to the

curriculum, reduced class sizes, well-prepared teachers, and extra support for students through after-school or summer programs (U.S. Department of Education, 1999).

National policy spotlights how children learn to read, the best pedagogy for teachers, and measuring reading growth as children learn how to read. Policymakers have been led to believe that there is one way, or a best way, for teaching students how to read as evidenced by a review of research that validates only experimental or quasi-experimental research (Allington, 2002b). As an example, the Congressionally requested report of the National Reading Panel (National Institute of Child Health and Human Development, 2000) focused their review of existing research using a specific set of methodological guidelines. In conducting their analyses, they concentrated on five components of literacy, including 1) phonemic awareness, 2) phonics, 3) fluency, 4) vocabulary, and 5) comprehension, using studies that met stringent guidelines. The emphasis on systematic and explicit scientifically-based reading research has some policymakers believing that only particular instructional methods work best for teaching children to read. While the National Reading Panel (NRP) provided many interesting and worthwhile recommendations, the report restricted review of other components known to be essential to teaching children to read such as writing instruction and independent reading. This review of research was central to recommendations set forth in the reauthorization of the Elementary and Secondary Education Act, the No Child Left Behind Act of 2001 (Bush, 2001).

The resounding theme of “every child” is clearly seen in the No Child Left Behind (NCLB) framework (Bush, 2001), the most recent wide-sweeping educational reform. Four pillars, as described on the NCLB website, are the cornerstones of this reform,

including 1) stronger accountability, 2) more local freedom, 3) proven methods, and 4) choices for parents (U.S. Department of Education, 2002). Each of these, when joined together, stimulate change as efforts to reform the American educational system take place.

Central to implementing NCLB is understanding the funding structure of this mandate. The two main sources of federal funding are Title I and Reading First grants. Title I funds were always designated for use in schools with higher levels of impoverished students; however, NCLB also provided a host of new stringent guidelines regulating how these funds may be spent. For our purposes here, it is fundamental to know that these funds must be spent upon scientifically-based methods and strategies, as defined by NCLB (and the NRP). In fiscal year 2004, the federal government appropriated \$18.5 billion dollars for Title I.

With the advent of NCLB came another federal grant called Reading First. As the name implies, this grant is geared specifically at boosting literacy achievement for students in the primary grades. Beginning with fiscal year 2002, \$900 million dollars were appropriated for this grant with similar funds available for the next five fiscal years (U.S. Department of Education, n.d.). Florida was one of the first states awarded a Reading First grant and as of May 2005, Florida's Department of Education has received an additional \$100 million of federal support for districts (Florida Department of Education, 2005b). However, to receive monies districts (or consortiums of districts) must win a sub-grant from the state's Reading First funding. To obtain these highly competitive grants, applicants must document how the monies will be spent while adhering to the State grant approved by the federal government.

In Florida, Reading First grants are reviewed by the FLDOE, as well as faculty affiliated the Florida Center for Reading Research (FCRR) at Florida State University (Florida Department of Education, 2005b). Eligibility for Reading First funds differ from Title I. Title I focuses on improving the educational achievement for students from high poverty backgrounds; Reading First focuses efforts on educational achievement for students who struggle with learning how to read in a way that prohibits reaching accountability. Reading First schools must come from districts with 15% or more of enrolled students from high-poverty backgrounds, and schools must have 10% or more of economically disadvantaged students, as opposed to the 40% level of economic disadvantage to receive Title I. As of the 2004-05 school year, Florida has over 400 Reading First schools.

Link to Accountability

Accountability has been part of policy discussion for years, however the signing of the NCLB Act of 2001 (Bush, 2001) pushed it into prominence. One key element of this reform is a rigid accountability requirement—schools and districts across the nation must meet what is called Adequate Yearly Progress (AYP) by 2014. Federal policy provides “a single definition of adequate yearly progress, the amount by which schools must increase their test scores to avoid some sort of sanction—an issue that in the past has been decided jointly by states and the federal government. And the federal government has set a single target date by which all students must exceed a state-defined proficiency level—an issue that in the past has been left almost entirely to states and localities” (Elmore, 2002, p. 31). State departments of education have the enormous task of developing accountability plans to comply with the federal mandate (Erpenbach, Forte-Fast, & Potts, 2003). Hence, the era of accountability is upon us and there is much

information to disseminate from federal, state and district policy initiatives—this concerns every elementary school principal across our nation.

Accountability legislation can include both rewards and sanctions (Massell, 2001). NCLB provides serious sanctions for schools and districts not meeting AYP that increase in severity with each year a school or district does not meet the standard. For some, the final penalty will result in a full restructuring, meaning that school personnel may be fired, schools could be converted into a charter school, or even turned over to a private enterprise. NCLB holds the system responsible for educating children. However, in the case of retention in grade as a measure of success, the onus of accountability falls on the student (Massell, 2001). Similarly, McGill-Franzen (1987) asserted in her case about students' placement in compensatory and special education programs, that we "place the burden of the problem on the student" (p. 488). I suggest the same is true of retention. Many problems that students face are out of their control and as a result, they suffer. Although the system pays, in a sense when accountability is publicized, it is the student who may carry the life-long consequence of failing a grade.

Standardized Assessments

In this section, we see how standardized assessments are inextricably tied to both reform and accountability. To add to the context of my study, the reading portion of Florida's standardized test, the FCAT, will be the focus of this section. The FCAT was first administered in spring 1999 as part of the state's response to national education reform. In 1999, Governor Jeb Bush brought high-stakes to Florida using the criterion-referenced portion of the FCAT as part of the state reform plan, the *A+ Plan for Education*, to boost student achievement. This plan assigned letter grades to successful schools providing monetary rewards, while sanctioning low performing schools (Dorn,

2004). In 2003, Florida attached another high-stake consequence—retention in grade—for third graders failing the FCAT-SSS, or FCAT.

The FCAT is given in grades 3 to 10, with two parts making up the reading portion of the FCAT (Florida Department of Education, 2003c). One part is criterion-referenced, assessing the state mandated standards-based curriculum, the Sunshine State Standards (SSS). Officially called the FCAT-SSS, most often it is simply called the FCAT. Scores are reported as Achievement Levels (Level 1 (low) to Level 5 (high)), and Developmental Scale Scores (DSS), which accumulate over time beginning in third grade. According to Florida state law, students who score Level 1 fail the assessment. Interestingly, the FCAT-SSS is said to be one of the most challenging state assessments currently in use, especially in terms of passage length and text difficulty of the passages (Hiebert, 2002).

The other portion of the FCAT is norm-referenced (FCAT-NRT), comparing Florida students to a national sample. Florida's norm-referenced test is a modified version of the Standard Achievement Test (SAT). Classroom teachers typically administer both parts of the FCAT with scoring conducted by contractors hired by the state. Florida's assessments are secure, meaning the tests are highly protected before, during and after administration (Florida Department of Education, 2000b).

It is widely endorsed that one assessment should not be the sole determinant in a high-stakes decision, such as retention in grade (American Educational Research Association, 2000; International Reading Association, 1999). Although Florida, too, claims to support this, children receive mandatory retention upon achieving Level 1 on the FCAT (Warford & Openshaw, 2004b). Technically, the FCAT is not the sole

determinant for retention, however other options would not be considered if it were not for a student failing the FCAT. It is the reverse of the adage—innocent until proven guilty. Instead, Florida third graders are guilty of failure, unless they qualify for a good cause exemption, described in more detail later.

Researchers have been studying retention for decades using standardized assessment data. With many retention decisions tied to high-stakes assessments, it is important to consider the legitimacy of these decisions and the relationship to education placements. It has been argued that some placement practices, such as retention and inclusion in special education programs, are the result of standardized assessments and the stakes attached to them (McGill-Franzen & Allington, 1993). After studying low-achieving children in multiple contexts, concern surfaced that some schools retained students before third grade to delay accountability penalties by classifying more students in special education programs. These efforts appeared to increase the reported achievements on a high-stakes assessment in these schools (McGill-Franzen & Allington, 1993).

Lastly, the State of Florida has relied on core reading programs as the panacea to help students read more proficiently, with a goal of increased performance on FCAT to meet accountability measures. In 2002, the State Textbook Adoption Committee found six core reading programs meeting the adoption criteria, meaning these programs are research-based and aligned to the states' reading curricula (Florida Department of Education, 2003b). In a study of third grade core reading programs used in Florida, two programs were selected for a content analysis focusing on the instruction provided to students (McGill-Franzen, Zmach, Solic, & Love Zeig, in press). Using school-level

data, these researchers worked to understand the relation between the percentage of third-grade students scoring Level 1 on the FCAT to the core program used, determining if schools using a particular program had an academic advantage. Using each school's percentage of students eligible to receive free or reduced price lunch as a proxy for poverty, McGill-Franzen et al. (in press) found that school level of poverty was a better predictor for academic achievement than the core program used. High-stakes standardized assessments have many unintended consequences, particularly for those most at risk for retention—students attending schools with higher rates of poverty.

High-Stakes Retention Policy

Built on the idea of setting high expectations for all students, standards-based reform resulted in standardization of the curriculum. It seems inherently logical that we want to test this knowledge to determine the effectiveness of instruction on students' learning as measured by standardized evaluations. However, assessments became high-stakes after attaching consequences and rewards. When then-President Clinton publicly called for the end of social promotion in his 1999 State of the Union address, it was clear that federal policy makers felt a decisive measure was needed to establish accountability (National Research Council, 1999). Thus, the emphasis on standardized assessments given in schools across the nation changed to hold students and schools accountable to achieve certain academic standards at specific points in a school career. Assessments moved from low- to high-stakes as they became used for retention and promotion decisions, or in grading schools. Retention policies based on students' performance on high-stakes assessments are becoming more common.

Policy Interpretation and Implementation

The last stage of the high-stakes framework described here involves how policy is interpreted and implemented. Policy logic, as described by Allington (2001), “proceeds on the assumption that implementing particular policies will have some intended effect” (p. 275)—there are two essential facets to this logic. The first presumes that intended policy will shape (or reshape) instruction. The subsequent presumption is that this reshaping will generate the desired results via change—sometimes in instructional practices (Allington, 2001). Those who study educational change know the complexities of this process (Allen, Cary, & Delgado, 1995; Flinders & Thornton, 1997; Fullan, 2001; Tye, 2000). Principals who handle change the best are more successful in implementing new programs and policies (Fullan, 2002).

Vast amounts of research have occurred on the interpretation and implementation of educational policy with teachers. Thinking about the school-level context, research has shown that the policies introduced by state lawmakers often have few similarities to actual practice (McGill-Franzen, 2000). Researchers have encountered teachers who were attempting to implement policy, but in actuality were creating their own policies based upon the intended legislation (Spillane & Jennings, 1997, cited in McGill-Franzen, 2000, p. 901). One would suspect that the same is true of school administrators. As policies are transformed from legislative intent to practice, there can be some loss in the translation. “What actually is delivered or provided under the aegis of a policy depends finally on the individual at the end of the line” (McLaughlin, 1987, p. 174). If we characterize the typical elementary school teacher as the “end of the line,” then the school principal acts as the intermediary between the district and the teacher. This junction

serves an important purpose in understanding what ultimately happens at the end of the line.

Resistance to policy is one reason for a policy's demise, when in fact, for many, it is a lack of understanding on the part of those charged with the implementation (Darling-Hammond, 1990). Policy analysts call for awareness on how policy is enacted locally. This insight can help guide the successful implementation of the said policy. The institutionalization of a policy depends on this understanding. Other problems in policy implementation exist. Lack of a core understanding is illustrated in the case study example of a teacher named Carol Turner, who in her mind, perceived that she was successfully implementing the new *California Mathematics Curriculum Framework* (Ball, 1990). With a lack of understanding of the new mathematics framework and a new textbook in hand, Carol was under the misguided impression that following the new textbook constituted enacting the new curriculum. Although it is unclear how a textbook can be the impetus for change, nonetheless in this case, it was viewed as the "messenger of change" (Ball, 1990, p. 257).

The notion of a messenger is vital. Much of the literature overlooks the role of the principal, who can be seen more in the role of policy messenger, than anything else. Nevertheless, even as a messenger, who may be deemed powerless by an instructional staff, the principal is the gatekeeper of knowledge through which the details of the policy permeate to the classroom teacher (Ball, 1990). As the instructional leader of a school, how a teacher learns the essentials of a policy is the responsibility of the principal or their designee. Valencia and Wixson (2004) explain a model illustrating a multi-level system through which policies are shaped as they move toward the core of education—the

classroom. As policy trickles through the system, there are many possible outcomes. Depending upon factors such as opinion or perception of a policy, there can be positive or negative conclusions to a particular policy (Valencia & Wixson, 2004).

We know implementation of policy is not simple—it is “a complexly interactive process without beginning or end” (Lindblom & Woodhouse, 1993, p. 11). It takes direction to achieve the desired results. With that said, past research has found that when it comes to policy implementation there are many school principals who feel less than adequate when enacting new policy (Musella, 1989). Senior administrators in a large school district requested Musella observe an in-service session as principals learned about a new governmental policy. As Musella witnessed this training, he noticed concerns regarding ambiguity of the policy goals, nonspecific policy guidelines, frustration about the lack of input in the policymaking process, inadequate funding resources, and insufficient time to enact changes. In an effort to share his observations, without pointing out the obvious faults in their process, Musella shared research on typical reasons for opposition to change (Zander, 1962, cited in Musella, 1989). A strikingly obvious, but often overlooked reason for policy failure is that “those responsible for changing are not involved in the planning” (Musella, 1989, p. 95). The principals observed felt powerless; they were not part of the change process; however, they were expected to enact a major governmental policy.

How administrators implement and understand new policy is important (Allington, McGill-Franzen, & Schick, 1997). Simply because a policy exists, does not ensure that it is acted upon, or adequately understood. Allington et al. (1997) described many important characteristics in their qualitative study of administrators’ understanding of

learning disabilities. Although many of the administrators interviewed supported state and federal programs, one essential finding of this study was that administrators perceived “creating and funding almost any intervention was someone else’s responsibility” (Allington et al., 1997, p. 231). These findings are fundamental to the third-grade retention policy in Florida. How principals implement the retention policy may rest more upon their beliefs, as policy implementers do “not always do as told” (McLaughlin, 1987, p. 172).

Florida Policy Context

Although the federal government endorses the use of retention in grade, this is a policy matter for local and state policymakers, such as a state department of education or a school district. The Florida policy mentioned thus far in my study was not the first attempt to enact a statewide retention policy. In 1999, the Florida Legislature did not pass a law for retention decisions based solely on a student’s FCAT score for fourth-grade students (Florida Department of Education, 2003a). Instead, student performance throughout the school year was deemed appropriate to make promotion and retention decisions, not simply a score from the FCAT. At that time, the Legislature decided that retention decisions should be left to the discretion of individual school districts. In August 1999, in lieu of a formal policy, the FLDOE recommended three options to school districts for students who were not meeting their district’s progression plan. Options included to, “(1) remediate students before the beginning of the next school year and promote, (2) promote and remediate in the following year with intensive remediation, and (3) retain and remediate” (Florida Department of Education, 2003a, p. 4). Even after making these recommendations to districts, the State tried again to enact a retention policy and this time met success.

Policy Enactment

The 2002 Legislature passed a rewrite of the Florida School Code (F.S. 1008.25) instituting a mandatory retention policy into effect on January 7, 2003, immediately ending social promotion in Florida schools (Florida Statutes, 2002). Of interest to this study is the mandatory third-grade retention policy enacted during the 2002-03 school year.

“The new law set the Grade 3 reading FCAT as the critical gateway to identify students who, after remediation, are still unable to demonstrate reading proficiency and clearly need more time to learn the basic skill of reading” (Florida Department of Education, 2003a, p. 2). Here is what third-grade students encounter: The law asserts that third-grade students, who do not achieve Level 1 on the reading portion of the FCAT, fail. Students performing at Level 1, as defined by the State, are said to have experienced little success with the standards-based curriculum, the Sunshine State Standards (Florida Department of Education, 2000a). This policy focuses on identifying students who need a stronger literacy foundation “...regardless of the reason that is causing it—even a learning disability, limited English proficiency, or a disadvantaged background—needs to be addressed and corrected before the student can be expected to move successfully on to the more difficult work of the higher grades” (Tremor & Butler, 2004, p.1). The Legislation has mandated school districts to “allocate remedial and supplemental instruction resources to students in the following priority: (1) first—students who are deficient in reading at the end of grade 3, and (2) next—students who fail to meet performance levels required for promotion” (Florida Department of Education, 2000a, p. 2).

Retention loopholes

As mentioned earlier, loopholes (or narrow exceptions) written into this law, called the “good cause exemptions,” allow some children to forego retention (Florida Department of Education, 2004d). Currently, six good cause exemptions allow promotion to fourth grade, if documented. Third-grade students meeting one of the following good cause options are eligible for promotion to fourth grade, even after scoring Level 1 on the FCAT

1. Limited English proficient (LEP) students having two years or less of English instruction
2. Students with disabilities (SWD) not participating in the state assessment because their individualized education plan (IEP) or Section 504 plan indicates that FCAT is not an appropriate assessment
3. Students scoring at the appropriate level on an alternative assessment
4. Students meeting district determined criteria to score the equivalent to Level 2 on FCAT via a good cause portfolio
5. Students with disabilities (SWD) previously retained in grade receiving two or more years of remediation
6. Students who have been retained two years in any grade, kindergarten to grade three, and have received two or more years of remediation

(Florida Department of Education, 2004d)

Here is further explanation for Options 3 through 6. Option 3, the FLDOE recommends districts use the Stanford Achievement Test (SAT) as the alternative assessment for students who fail the FCAT-SSS. Because of the challenging nature of the FCAT, “this is one of the reasons why, in third-grade, a student must achieve at the 51st percentile on the nationally normed SAT9 test of reading in order to insure that they have sufficient reading ability to achieved above Level 1 on the FCAT” (Torgesen, 2004, p. 2). On May 17, 2005, the FLDOE provided a new cut score for the FCAT-NRT, now

version 10 of the SAT, to at or above the 45th percentile as announced in a memorandum for District School Superintendents (Warford & Openshaw, 2005). Districts electing to use the SAT-10 as their alternative assessment will adhere to the 45th percentile or above, while districts still using the SAT-9, described by Torgesen (2004), must continue using the 51st percentile or above as the criteria. Depending on the version, SAT-9 or SAT-10, students must meet the designated cut-point to be eligible for a promotion with good cause.¹

Option 4 allows a good cause portfolio designed by school districts. Each district had the opportunity to create a portfolio for classroom teachers to compile on behalf of students identified at risk to fail the FCAT. The FLDOE notes that only teachers may initiate and compile these portfolios. The state recommended four districts, Citrus, Clay, Orange and Pasco, as examples for other districts to use as models to establish their portfolio criteria (Florida Department of Education, 2005). Good cause portfolios are designed to be an on-going, collection documenting students' classroom performance equivalent to Level 2 or higher on the FCAT. If a student fails the FCAT, the teacher submits a portfolio to the school principal to determine if the portfolio meets the criteria to promote the student for good cause.² "If the school principal determines that the students should be promoted, the principal must recommend it in writing to the district superintendent. The district superintendent must accept or reject the school principal's recommendation in writing" (Florida Department of Education, 2004a, p. 8).

¹ It is unclear how the FLDOE determined the new cut-point for the SAT-10. Based on information provided by the test publisher, Harcourt Assessment, the SAT-10 cut-point mandated by the FLDOE is lower than the equivalent calculated by the publisher. Harcourt equates a 51st percentile rank on the SAT-9 to a 46th percentile rank for the total reading score, not the 45th percentile.

² For more about good cause portfolios, please visit www.firn.edu/doe/commhome/progress/proghome.htm.

The latter two good cause promotion exemptions are complex with numerous scenarios by which a student could be promoted. Complexities arise when understanding the options available to those students previously retained in K to 3. Option 5 relates to students with disabilities (SWD) retained in any grade, kindergarten to grade three. For these students, they may be promoted to fourth-grade after being retained one time; however, they must have received at least two years of reading remediation. Hypothetically, a SWD who was retained in first grade and received extra remediation during the year of retention and in second grade could be promoted to fourth grade after failing the third-grade FCAT. Another hypothetical student may have been identified in second grade as a struggling reader. However, this SWD student received remediation in second grade and was promoted to third grade. After a second year of remediation, this student would not be eligible for promotion because the child was never retained. After a second year in third grade, this child would be eligible for a promotion with good cause, if they still do not pass the third-grade FCAT.

The final good cause exemption, option 6, allows promotion for good cause to any “students who have received the intensive remediation in reading for two or more years, but still demonstrate a deficiency in reading and who were previously retained in K-3 for a total of two years” (Florida Department of Education, 2004a, p. 7). This good cause promotion set a limit on the number of times a student may be retained in K to 3, provided they received the appropriate remediation (Warford & Openshaw, 2004b), although Florida does not currently limit the number of retentions a student may experience in elementary school. As can be seen, good cause exemptions are stringent

and most complex. In a sense, upon failing the FCAT, students' promotional decisions are treated on a case-by-case basis.

Lastly, state law requires districts to report the specific type of Good Cause Exemption used to promote a student each year, beginning with the 2002-03 retentions (Warford, 2004b). Using data provided by the FLDOE, Table 2-1 disaggregates the number of Florida students promoted with good cause after the first and second years of the policy.

Table 2-1. Florida retention policy good cause statewide exemptions

	2002-03		2003-04	
	#	%	#	%
1) Limited English proficient students with fewer than 2 years in English as a second language program	2,974	22.80	2,511	11.79
2) Students with disabilities (SWD) not participating in statewide assessment as per individualized education plan (IEP)	1,016	7.79	1,647	7.73
3) Students who demonstrate proficiency on an alternate assessment	3,307	25.35	3,845	18.05
4) Students who demonstrate proficiency through a portfolio	1,514	11.61	3,468	16.28
5) Students with disabilities (SWD) retained <i>once</i> with 2 or more years of remediation	3,637	27.88	7,906	37.12
6) Students retained <i>twice</i> with 2 or more years of remediation	598	4.58	1,924	9.03
TOTAL PROMOTED WITH GOOD CAUSE	13,046		21,301	

Note. Data provided courtesy of Florida Department of Education (L. Fleming, personal communication, June 6, 2005).

Let us compare the numbers of students scoring Level 1 on the FCAT to the numbers of students promoted for good cause to understand the impact of this policy.

Over 43,000 students scored Level 1 on the FCAT in 2002-03, and the FLDOE reported,

as shown in Table 2-1, that after the 2002-03 school year over 13,000 students were promoted with good cause to fourth grade. Based on this information, it would appear that nearly 30,000 students were retained after the first year of policy implementation. However, according to data also collected by the FLDOE, state files document that over 27,000 third-grade students were retained after policy implementation and after application of good cause exemptions and we will rely upon this number.³

Policy Guidelines

Much of the retention policy echoes the mandates set forth in NCLB. Identifying students who struggle with reading and evidencing their reading achievement are central to both national and statewide reform efforts. However, like retention, determining the best technique to help students who struggle with reading has a long and contentious past (Allington & McGill-Franzen, 2000). Rather than debate here which techniques are best, I present the mandates and suggestions endorsed by the FLDOE in the 2004 Legislative Intent and other technical reports disseminated to Florida schools via the state paperless communication system (Florida Department of Education, 2004a; Tremor & Butler, 2004; Warford & Openshaw, 2004a, 2004b, 2004c, 2005).

The Legislative Intent (hereafter, Intent) for the state third grade policy (F.S. 1008.25, Section 6) has three main areas of focus, including: 1) determining reading proficiency, 2) parental notification, and 3) developing comprehensive plans to address the Intent goal for all students to read at or above grade level. Two new initiatives were also announced by FLDOE officials in the 2004 Intent (Florida Department of Education,

³ School districts reported data to the State during the same collection period in August 2003. The reason for these discrepant numbers is unclear; however, for the purposes of my study we will rely on the numbers disaggregated by school (i.e., approximately 27,000 third grade students retained in 2002-03), not the total extrapolated by the researcher, even though these should have been nearly the same. One FLDOE official suggested the amount may be discrepant due to students moving. Although this suggestion would only apply to state numbers if students were moving out-of-state, not within the state.

2004a; Warford & Openshaw, 2004a). One, called *Reading Enhancement and Acceleration Development (READ) Initiative*, is a retention prevention program aimed at students in grades kindergarten to three, including newly retained third graders. Multiple lines of legislation outline several steps and sub-steps that comprise this initiative (s. 1008.25(7) (b) 7, F.S.; Bill page 17, line 1 – page 19, line 23) (Florida Statutes, 2002). Another initiative calls for establishing classrooms for “retained third grade students who subsequently score at Level 1 on the reading portion of FCAT” (F.S. 1008.25(7) (b) 8; Bill page 18, lines 1 – 23) (Florida Statutes, 2002). These classrooms, called an *Intensive Acceleration Class*, have specifications outlined in the legislation. A brief overview of how at risk students are identified, how parents are notified and how students are supported follows.⁴

Identification

First, all students in Kindergarten through Grade Three (hereafter, K to 3) have their reading proficiency determined and monitored. The FLDOE is focused on identifying students in K to 3 who lack the required proficiency to pass the third grade FCAT, defining these students as “exhibiting a substantial deficiency in reading” and “must be given intensive reading instruction following the identification.” To make diagnoses, elementary schools are required to assess their students “regularly” to identify the “exact nature of the student’s difficulty in learning to read.” Reading First school students must be assessed using the statewide assessment system, the Dynamic Indicators of Basic Early Literacy (DIBELS; Good & Kaminski, 2002). Non-Reading First schools must also identify their students, however in addition to the DIBELS, these schools may

⁴ To provide the reader the opportunity to hear the tone of the 2004 Legislative Intent, specific terms or wording used in the Legislative Intent were included here using quotation marks.

select from other assessments. Newly identified students are then subject to a remediation plan that must be provided “during regular school hours.” All of this can be likened to a critically-ill person being treated as an out-patient.

Notification

Upon labeling a student as reading “deficient,” parents must be notified as soon as possible with a “description and explanation, in terms understandable to the parent.” It is expected that parents be consulted in developing an academic improvement plan (AIP) and that parents are provided a list of the supplemental instructional services and supports to be provided “until the deficiency is corrected.” Schools are to keep parents apprized of their child’s progress. It is also required that parents be provided strategies to help their child with reading. At the end of third grade, parents are also alerted that if the services and supports do not correct the deficiency, “as demonstrated by scoring at Level 2 or higher on the statewide assessment test in reading for grade 3,” then the student must be retained, unless they qualify for promotion with good cause.

Retention plus remediation

In addition to diagnosing students and sharing this news with parents, the Intent specifies how to remediate students. Schools and districts must follow the regulations of the READ Initiative for those students who qualify. This Initiative aims to prevent retention for students in K to 3, and for the newly retained third graders, it offers an accelerated, or intensive approach, to help prepare students for promotion to fourth grade. Each student must receive extra support during the regular school day, which is in addition to the mandatory 90-minutes of daily, uninterrupted reading instruction utilizing a scientifically-based program. These are the state approved and Florida Center for Reading Research (FCRR) reviewed core reading programs focusing on the five areas of

reading identified by the NRP: 1) phonemic awareness, 2) phonics, 3) fluency, 4) vocabulary, and 5) reading comprehension (National Reading Panel, 1999). Ongoing progress monitoring must also be provided. All core subject areas, such as science, mathematics, social studies, are to be incorporated into the school day, in addition to the above.

Double retention plus remediation

In 2004-05 a new initiative began for students who were already retained in third grade, but did not pass the FCAT again. Each district must establish an Intensive Acceleration Class and qualified students must participate. The goal of this program is to accelerate learning for students to gain two years worth of material in one year. Theoretically, the FLDOE suggests that successfully remediated students could advance to fifth grade and by-pass fourth grade. This child could potentially rejoin his or her same age cohort. The Intent does not suggest that students would advance to the fourth grade via this program. To accomplish this extraordinary feat, the Intent explicitly describes several rules for districts (Florida Department of Education, 2004a, pp. 11–12). It recommends reduced teacher-student ratios, however ratios are not defined. The FLDOE notes that class sizes are expected to be smaller than other third-grade classes. Teacher-student contact time is expected for most of the day. In other content area subjects, such as Mathematics or Science, students are expected to be taught from the fourth-grade Sunshine State Standards. As well, the fourth-grade Language Arts strand is used. This provision helps ensure students will have the necessary background knowledge to cope if promoted directly to fifth grade. Students in an Intensive Acceleration class also have services of a speech-language pathologist, if needed.

Students are to be assessed weekly. Districts must monitor all students taking part in this Initiative.

Central supports

Several supports are embedded within the Intent for all identified students with reading difficulties. State law requires districts to offer certain supports, while others are suggested, but not mandated. Six supports (Table 2-2), central to my research, and their purpose within the Florida retention policy context, are described here (Florida Department of Education, 2004a).

Table 2-2. Supports identified in Florida's legislative intent

Required	Recommended
Focus on early intervention	Extended learning opportunities
Current academic improvement plan (AIP)	Transitional class available for retained students
On-going portfolios meeting state requirements	Mentor or tutor with specialized training

Focus on early intervention. Students who are identified as early as kindergarten are provided extra supports. As part of the establishment of the READ initiative described earlier, third grade students are provided an accelerated curriculum and students in grades K to 2 are provided intervention to foster their reading development. The central tenet here is to help prevent third-grade retention. No laws exist to retain children before third grade. Prevention can benefit students experiencing difficulty (Allington, 2002a; Allington & Walmsley, 1995).

Current academic improvement plan (AIP). An Academic Improvement Plan (AIP) formally documents the remedial strategies provided to the student over a period. The FLDOE requires schools to collaborate with parents when making or revising an

AIP. The AIP must name which of the five areas of reading identified by the NRP (1999) is lacking.

On-going portfolios meeting state requirements. All retained third graders must have an “active, ongoing portfolio” that may be used as part of the good cause promotion options. As mentioned earlier, these portfolios are district-created, so there is expected variation across districts. Essentially, these teacher initiated and compiled good cause portfolios must contain evidence that accurately reflects whether the student can achieve comparable to Level 2 on the third-grade FCAT. Both the school principal and district superintendent will review and review these portfolios.

Extended learning opportunities. All students who have trouble with reading must be provided extra intensive support beyond the 90-minute reading block. The FLDOE advocates the use of extended learning outside the regular school day, although by definition, this would be in addition to the extra support that must be provided during the regular day beyond the 90-minutes of instruction. Extended learning may occur before- or after-school, or on the weekend, such as a Saturday School. Another option includes the use of an extended school year beyond the minimum 180-days. Extended learning, as suggested by the FLDOE, is different from summer school. All students who fail the third grade FCAT must be provided the opportunity to attend summer school, or as they are called in Florida, Summer Reading Camp.

Transitional class available for retained students. The Intent mandates that retained students must be offered the choice of a transitional type setting with the purpose of producing learning gains. The goal of such classes, as noted by the FLDOE, is “*what* is being provided to help the student catch up, not *where* it is being provided” (Florida

Department of Education, 2004a, p. 13) [emphasis in original]. Configurations may vary, and schools may choose transitional classes that contain third- and fourth-grade students or re-retained third graders only. Districts may elect to offer the transitional classroom at a central location. This type of class is akin to transitional classes preparing students for first grade (e.g., pre-first grade, junior first grade, readiness room) (Shepard, 1989).

Mentor or tutor with specialized training. All parents of retained students must be provided with either “supplemental tutoring in scientifically research-based reading services in addition to the reading block” (Florida Department of Education, 2004a, p. 10). This tutoring may happen before and/or after school. Alternatively, students may be provided with a mentor or tutor with specialized reading training, as opposed to using a scientifically research-based program.⁵

The FLDOE provided many alternatives for districts to select with their allocated funds. Other options recommended in the Intent include: reduction in teacher-pupil ratio, although the FLDOE does not endorse a capacity or class size. The Intent also recommends more “frequent” progress monitoring, which may include the assessments found within the scientifically research-based programs. A frequency is not prescribed.

Implementing the policy

Policymakers acknowledge the complexities of the process (Hart, 1996). To cope with the policy fluctuations, local school districts are provided with communications and meetings as a means to guide changes and aid understanding in a timely manner. They serve as a way to distribute the recipe with state policymakers telling districts *what* needs

⁵ In September 2004, the FLDOE surveyed districts to learn which of the three options they selected to provide parents of retained students. Districts are required to provide at least one of the following: 1) tutoring with a research-based program, 2) a tutor or mentor trained in reading, or 3) a “Read at Home” contract.

attention and *how* to do it. To assist struggling readers, Florida mandates schools and districts use special classes, use certain types of curriculum materials, as well as provide summer school and make certain services available. Timely explanations are needed to assist policy implementers to understand mandates as this law continues to evolve.

The FLDOE is committed to providing updates and technical assistance papers concerning this policy via an open-access paperless communication system. Although many documents are intended for school district superintendents or other district leaders, interested members of the public are free to register for an e-mail service that delivers documents upon release. As of June 2005, twenty-four notices dating back to August 26, 2002 were posted on the FLDOE website regarding this policy, plus related attachments (Florida Department of Education, 2005c), although many other documents have been distributed since the last entry on January 11, 2005. Recent documents are presumably available elsewhere on the FLDOE website.

Following the first year of policy implementation (2003-04), the FLDOE sent an electronic memorandum to Florida's district superintendents outlining amendments of two Florida state educational laws (F.S. 1002.20 and 1008.25) (Warford & Openshaw, 2004a). As part of this paperless communication system, a carbon-copy (cc:) of this memorandum was also sent electronically to the Assistant Superintendents for Curriculum and Instruction, Directors of Student Services, Directors of Elementary Education, Directors of Exceptional Student Education, Elementary School Principals, and Elementary Guidance Supervisors of each district. Revisions to 2004-05 student progression plans require changes, many of which relate or interrelate to the third-grade retention policy. Essentially, the memorandum provides a lengthy list of requirements

for each elementary school with a copy of the bill attached. Specifically, the FLDOE requires each school to

assess the reading ability of each K-3 student, provide parents with notification of any reading deficiency, implement a detailed academic improvement plan (AIP), and provide intensive reading instruction; revise the required notice to parent of third grade students with substantial reading deficiencies to include information about additional evaluations, portfolio review and assessment to determine whether the student is ready for promotion, and information on the district's specific criteria and policies for mid-year promotion; define mid-year promotion; make a technical correction related to students who were previously retained in grade three; for third graders who are retained, require appropriate intensive interventions, including the provision of summer reading camps; specify the activities and supports to be provided to retained third graders, including the use of a state-identified reading curriculum that meets certain specifications; provide for intensive acceleration for students currently retained who score Level 1 in reading; require a report to the State Board of Education on the interventions provided; and require the option of placement in a transitional setting for retained third graders (Warford & Openshaw, 2004a, p. 1).

After school districts had opportunity to review the wide-ranging edicts from the memorandum outlining the legislative changes, the FLDOE issued another memorandum with a chart clarifying the old and new legislative intents (Warford & Openshaw, 2004b). This chart, covering 14 pages, provides comparison of the original intent to the current modifications (Florida Department of Education, 2004a; Warford & Openshaw, 2004b). This serves as evidence that the FLDOE does provide support to districts in their efforts to decipher and enact changes in a timely manner.

Even with these technical assistance reports, questions and confusion exist as implementing this policy continues. In October 2004, members of the FLDOE Just Read! Florida⁶ office traveled to multiple sites across the state holding meetings, offering assistance to principals and district-level administrators. I had the opportunity to attend

⁶ Members of the Florida Department of Education Just Read! Florida office participated in writing the Legislative Intent for the retention policy. This office oversees the technical assistance of this policy. Also, these October meetings were originally slated for September 2004, but were rescheduled due to hurricanes.

the October 7th meeting held in Tampa, Florida. The purpose of the meeting was for attendees to receive clarifications to their implementation questions, and for attendees to provide input prior to the impending passage of the mid-year promotion rule. Several matters were discussed at length; however, I will provide just two examples. First, the FLDOE, as described earlier, introduced two new initiatives, READ and Intensive Acceleration, for the 2004-05 school year. When these initiatives were introduced and as they were written in the Legislative Intent, they were referred to by these titles.

However, the FLDOE released another document that described these initiatives as a “tiered” system. At the meeting, members of the FLDOE used the titles interchangeably, causing confusion. After explaining that “Tier 1” (identified as at risk for retention) and “Tier 2” (retained once) students participate in the READ Initiative, while “Tier 3” (retained two times in the same grade) students participate in Intensive Acceleration Classes, confusion appeared to ease. Another area of questioning related to the number of times students may be retained. Participants received advice regarding how to handle specific cases for students with multiple retentions explaining how students who received two years of intensive intervention may not be retained in a grade for more than two years. After two years of intensive intervention, students would be eligible for a good cause exemption. The observed scenario was reminiscent of what Musella (1989) described in his work observing an in-service training. Similarly, I heard concerns regarding ambiguity of the interventions and unclear policy guidelines.

Summary

Themes resonating from the past research on retention indicate that 1) students need to be treated as individuals when making a life-altering retention decision and 2)

students who are at risk for retention or retained need adequate support to help them achieve their highest potential. As we can see, retention is deeply rooted within the high-stakes framework presented and examined within this chapter. In Florida, simply because FCAT achievement data appears to point toward success, how has the state managed to raise achievement while drastically increasing the number of students retained in grade? In order gain more insight into this policy, we work to understand how Florida elementary school principals provide support within the given framework.

CHAPTER 3 METHODS

Introduction

Historically, retention has been a dichotomous debate with little middle ground; however, in the previous decade, it has gained acceptance with more states using it as a component of a state accountability scheme. Florida is one such state to enact a policy preventing promotion in grade when students do not meet set criteria. The impetus of this policy is to help the state meet the national challenge to have all third-grade students reading on grade-level by 2014. Florida is determined to beat the national goal, calling for districts to have all third-graders reading on level by 2012. This retention policy is relatively new with little known about school-level implementation.

The foundation of my study was rooted in a multilevel model considering factors, such as level of poverty in their schools, while determining whether the past retention practices of a principal influenced the level of current support offered to students at risk for retention or currently retained. This multilevel model allows the reader to look broadly—discerning whether retention practices have changed since the start of this policy. Using 5 years of percent retained data enabled examination of differences in retention rates before and after policy implementation. As recommended by Allington (2001) for the study of policy, data from pre-policy and post-policy provide opportunity for comparison. To better understand the school years discussed in the balance of this study, Table 3-1 displays the school years and their relation to the start of the retention policy.

Table 3-1. Time in years: Pre-policy versus post-policy

<i>Year</i>	<i>School Year</i>	<i>Retention Policy</i>
1	1999-2000	
2	2000-2001	Pre-Policy
3	2001-2002	
4	2002-2003	
5	2003-2004	Post-Policy

Participants

Population

To study the Florida third grade retention policy, Florida principals were selected. This State has 67 traditional public school districts encompassing over 1,800 elementary schools. Each school district covers an entire county.

Sample

For this study, districts were selected both purposefully and randomly for a total of 920 possible public elementary schools.¹ The six districts selected with purpose are among the largest of Florida's school districts, while the other six districts were selected at random from the remaining available districts. This sample included districts from each geographic region of the State – the panhandle, the north, the south, landlocked areas, the Gulf coast, and Atlantic coastal regions. Within and across all districts, combinations of urban, suburban and rural schools exist (National Center for Education Statistics, 2002).

Prior to submitting proposals to conduct research in these districts, ten of the twelve districts expressed interest in participation, responding positively with a letter of

¹ Charter schools were not included in this study.

interest to the researcher. The other two districts welcomed submission of the research request. Of these, all districts willing agreed to partake in the study after independent review by district representatives. Across all districts, I sought formal school district approval to request principals' voluntary participation in the *Florida Principal Survey*. One large district restricted access to a cluster of thirty elementary schools participating in a district research evaluation, reducing the available sample from 920 elementary school principals to a maximum of 890 possible elementary school principals.

Confidentiality Protection

In the district approvals, it was agreed that I may refer to districts by enrollment group, as opposed to specific district names. Based on their PK to 12 district enrollments gathered using the Common Core of Data Build a Table tool (National Center for Education Statistics, 2004), I classified these 12 districts into three groups: 1) small, 2) mid-size, and 3) large. Districts with fewer than 30,000 students enrolled were classified as small; districts with greater than 30,001 students enrolled were deemed mid-sized. The other six districts, with over 100,000 students enrolled, were classified as large. This classification system may be unique to Florida as its districts are countywide. To protect the identities of the participating districts, specific enrollment figures are not displayed. Also, specific school and principal names will not be revealed, nor released publicly. Each principal was assigned a unique identification code as an additional safeguard to protect the identities of the respondents. Appendix A shows the *Florida Principal Survey* that participants received.

Data Collection

Survey Instrument

As this policy stretches across Florida, survey design was selected as a data collection method because it is representative, objective, quantifiable and systematic (Isaac & Michaels, 1995). The survey questionnaire, researcher created, referred to as the *Florida Principal Survey* contains three distinct sections. Part 1 has four items requesting respondents to identify their name, school and district, as well as the unique identification number provided in their consent letter. Part 2 contains items for principals to share their attitudes and practices related to the retention policy. Most topics for the survey were derived after reviewing the Intent memo the FLDOE sent to schools and districts outlining the legislative changes (Warford & Openshaw, 2004b).

For my study, the focus of item construction was the support available for at risk students, including those at risk for retention, those currently retained, or those retained again in third grade.² The closed-ended items in this study were fixed-alternatives. One item offered two response options (Figure 3-2) and six items were categorical by design with response choices, “yes, no, or not sure” (Figures 3-3 and 3-4). However, the “not sure” option was not selected by this sample. For the purposes of this study, these items were treated as dichotomous, “yes, no.” Part 3 of the survey requested the principal to enter the years of career experience as an educator, the total years as a principal, and the length of time spent as principal at their current school. A summary of the pilot survey review follows in Appendix B.

² Other items included in the survey are for a separate report.

Retention belief

Item 9 served to identify the retention belief of principals as related to the use of one standardized assessment result to determine a retention decision. As current policy bases the decision on FCAT performance, it was salient to this study to understand principals' responses. By design, principals were asked to select between the two options. Principals were provided the response choices shown in figure 3-2. They could click or check the option best matching their current belief.

Directions: Knowing that individual circumstances can exist, what is your response to the following statement?	
"On the whole, retention in grade can benefit 3rd grade students who score Level 1 on the FCAT"	
Item Number	Response Choices
Q9	Yes, retention is a beneficial option for students who fail the FCAT.
	No, retention is not a beneficial option for students who fail the FCAT.

Figure 3-1. Item 9, *Florida Principal Survey*

Resources for at risk students

The next two groupings of items provided an opportunity for principals to report on the presence of different types of supports or monitoring systems currently used in their schools. As legislation has strongly suggested that schools and districts use particular supports to remediate their students, I wanted to learn which supports were being used by schools. Survey items are presented in figures 3-3 and 3-4.

Degree of support

Based on the data collected, it would not be possible to make determinations about

Directions: Think about programs or initiatives currently in place at your school designed to help at-risk students.

Item Number	Survey Item
Q14b	Does your school emphasize early intervention programs in K - 2 to prevent retention in 3 rd grade more now than in the past?
Q14c	Do ALL retained 3rd grade students in your school have a current AIP?
Q14e	Is a transitional class (a class designed for the retained and/or re-retained) available for retained students?

Figure 3-2. Item 14, *Florida Principal Survey*

Directions: Which, if any, of the following are used as strategies with retained third grade students?

Item Number	Survey Item
Q19a	A mentor or tutor with specialized reading training
Q19b	Extended school day (such as After School Program, Saturday School, or Extended School Year) as defined in the NCLB legislation
Q19c	Ongoing portfolios that meet state portfolio requirements

Figure 3-3. Item 19, *Florida Principal Survey*

the quality of these supports as implemented in each school; however, I decided to examine the number of supports reported as implemented to determine if any patterns exist. A new variable created for the planned analyses lends itself to help explain retention patterns for schools using fewer or more supports, that is the degree of support provided in each school. A categorical variable called *suppsys3grp* was created based on principals' responses to the item numbers in Table 3-2. Schools reporting use of any 1 or 2 supports provided a lower degree of support options; schools using any combination of 3 or 4 supports provided a medium degree of support options; and schools using 5 or 6 supports provided a higher degree of support options.

Table 3-2. Degree of support category, *Florida Principal Survey*

Item number	Support type
Q14b	Focus on Early Intervention
Q14c	Current Academic Improvement Plan (AIP)
Q14e	Transitional Class available for retained students
Q19a	Mentor or Tutor with specialized training
Q19b	Extended Learning Opportunities
Q19c	On-going portfolios meeting state requirements

Employment history

Between these items (Figure 3-4), there was enough information, for the purposes of this study, to determine how long the respondent held a principalship at their current school. The purpose of the length of the principalship at the same school will be described in more detail later.

Item Number 28. How many years have you been a principal? (Including 2004-2005 school year)	_____
Item Number 29. How many years have you been principal at your current school? (Including 2004-2005 school year)	_____

Figure 3-4. Employment history, *Florida Principal Survey*

Procedures

The first phase of the data collection began after the start of the 2004-05 school year with an on-line survey sent electronically to the elementary school principals of the twelve participating districts. Approximately 50% of the districts volunteered to have a district official contact their principals to acknowledge their approval of the study. All participating principals had Internet access.

Principals' e-mail addresses were determined using multiple sources. Most districts do not maintain databases to share e-mail addresses outside their school district Intranet. Small districts were willing to provide or check e-mail addresses gathered. Some e-mail addresses were gathered from district and school websites, while hundreds were created using a "formulaic" method (e.g., a seven character name followed by the district URL or FirstName.LastName@district.k12.fl.us) suggested by district officials. Nearly two hundred schools were telephoned to request the e-mail address of their principal. Surprisingly, the least effective method of obtaining accurate an e-mail address was calling, as many school employees did not readily know this information. Every reasonable effort was made to ensure the survey reached the intended sample. A three-stage follow-up sequence to collect data was planned (Dillman, 2000).

Research decisions during the collection period

During the data collection period, four hurricanes inundated Florida. Every school district in the state was affected by at least one hurricane resulting in school closures. In the 12 districts from this study, the cumulative number of district closures due to the four hurricanes was evenly distributed by district size with an average of 24 days lost.³

Two areas of data collection were potentially problematic. First, because of the impact of the hurricanes, a decision was needed whether to continue data collection in these districts. On September 29, 2004, the FLDOE announced that the FCAT would proceed as planned with some provisions for those districts most greatly affected to have more time to administer the assessment (Warford, 2004a). The State Education Commissioner John Winn encouraged districts "to restore a sense of normalcy by

³ To protect the identities of the districts, specific numbers are not given.

continuing to focus on student achievement” (Florida Department of Education, 2004b, ¶ 4). Because of these announcements, data collection proceeded, as there was no hold on the policy under study.

The second problematic area was the low response rate from the on-line survey. After 5 weeks of data collection, a 9% response rate was achieved. A decision was reached to offer an alternate paper-pencil survey after learning that districts would permit a paper version of the survey to be sent to the nearly 800 non-responding principals under current district approvals. All non-responding principals received a copy of the paper survey along with a postage-paid envelope to return the survey directly to the researcher. The on-line format continued to be available for those who chose that format. The paper survey offered in October was nearly identical to the on-line format. Questions used the same wording and were asked in the same order. Minor modifications, were made where the on-line survey instructed respondents to “click” the appropriate response. For example, the paper survey was changed to “check” the response. Non-responding principals were contacted using a two-stage follow-up (Dillman, 2000). I e-mailed principals two times after the paper surveys were mailed requesting their participation using the survey format of their choice.

These revisions resulted in an overall 29% return rate. Of the 255 principals who returned the survey, 137 principals completed the on-line format, while 118 principals completed the paper-pencil format. An additional 2% of the principals declined participation stating lack of time as a reason. Table 3-3 shows the response rates achieved using the district size classification. All returned surveys, including those returned without names, were included in these calculations. Compared to the large

districts, a higher percentage of principals from small and mid-sized districts responded favorably to this survey.

Table 3-3. Survey response rates by district size

<i>Small</i>	<i>N</i>	<i>%</i>
Responded	18	41.9
Declined	2	4.7
Non-response	23	53.4
Total	43	
<i>Mid-Size</i>	<i>N</i>	<i>%</i>
Responded	46	40.0
Declined	4	3.5
Non-response	65	56.5
Total	115	
<i>Large</i>	<i>N</i>	<i>%</i>
Responded	191	26.1
Declined	15	2.0
Non-response	526	71.9
Total	732	

Usable Sample Procedure

While awaiting the close of the survey data collection, several databases of information needed for this study were gathered, including the third grade percent retained calculations and the school-wide percent of students eligible for a free or reduced price lunch.⁴ Of the 255 returned surveys, each district was identified; however, seven surveys were not linked to a principal and were therefore not included, bringing the maximum available to 248 surveys. Every reasonable effort was exhausted to locate missing data. All missing pieces of data were hand-searched to crosscheck for accuracy. Prior to determining the final sample, a 10% random sample was selected to ensure the accuracy of the databases compiled for the study and detected no errors in data entry or data calculation.

⁴ See Appendix C for a list of the data, including the source and year.

For each type of data collected, a systematic procedure determined the usability of a principal in this study. Principals were evaluated to determine how they met the screening criteria. Essentially, each case, or principal, was examined to determine the level of missing data. The criteria included 1) holding a principalship at their present location for 5 or 6 continuous years, 2) assessing the completeness of third grade retention rates, 3) assessing the completeness of school-wide free or reduced price lunch eligibility rates, and 4) assessing the completeness of survey items used in these analyses. Each piece of the screening criteria is described below, as well the how I arrived at the total usable sample. At the end of each procedure a sub-total is provided.

First, as surveys were returned to the researcher, they were entered into a database structured at the person-level where each principal represented a row (or case) (Kreft & de Leeuw, 1998; Singer & Willett, 2003). Then, using the names of the responding principals, I created a matrix of each principal's employment history for school years 1999-2000 through 2004-2005. Three components constituted a principal's employment history which included: 1) the number of years of principalship total as reported on item 28; 2) the number of year of principalship held at their current school, as of 2004-2005, as reported on survey item 29; and 3) the principal needed to still work at the same school, as revealed by the name of the school they provided on the survey questionnaire. Of those who responded to the survey, 121 principals held a principalship for at least 5 or 6 continuous years at the same school and thereby met the time requirement.

The second factor considered third grade retention percentages for the 5-year period examined in this study. Using data files provided by the Florida Department of Education, I created two sets of spreadsheet files. One file contained raw numbers of

students retained in third grade from each of the 5 years used in this study, and the other file contained the numbers of third graders enrolled each year. I merged these files to calculate the percent of third grade students retained in grade for each year examined in this study. There were cases when one or both of the state provided data files did not contain either the raw number of students enrolled or the enrollment needed to determine the percentage. There were 6 principals removed due to missing data from one or more time point. This reduced the usable sample to 115 principals.

The third factor considered the level of missing free or reduced price lunch eligibility percentages for each school. Following the same procedure as determining the percent of students retained in third grade, free and reduced price lunch eligibility databases were created for each of the 5 years from files provided by the FLDOE. Of the 114 principals, there was one principal with many years of free or reduced price lunch eligibility missing; therefore, to avoid erroneous conclusions, this principal was removed. The total usable was reduced to 114 principals.

The survey items used in this study underwent an inspection to assess the completeness of the closed-item responses. The four principals who did not respond to item 9 were removed, as were two others who did not respond to at least half of the items used to determine the degree of support groups. The total usable sample was reduced to 108 principals.

Finally, after one final inspection of the 108 principals, I reexamined all cases and removed two outliers. Since only two schools that primarily serve students with disabilities responded to the survey, these outliers were removed due to extremely small third grade student populations. An additional inspection of the data revealed 4

principals with extreme values. These schools contained percent retained values that appeared to be anomalous. Without revealing too much identifying information, an example is provided. One school contained near zero percent retained for two years, while in the next year nearly a quarter of third grade students were retained. To avoid erroneous conclusions, these principals were removed from the final analysis. In all, there is a usable sample of 102 principals in these analyses. The final distribution of the usable sample by district size is provided in Table 3-4.

Table 3-4. Usable sample versus returned responses by district size

District Size	Total Sent	Returned	Usable	Percent Usable
Small	43	18	11	61.1
Mid-Size	115	46	21	45.7
Large	732	191	70	36.6
<i>Total</i>	<i>890</i>	<i>255</i>	<i>102</i>	<i>40.0</i>

Usable Sample Representativeness

Population and sample data were compared to determine the probability that the sample represents the percent retained for the population of all state elementary schools. Since total population data was readily available in files provided by the FLDOE for each of the 5 years of retention data used in this study, I compared the usable sample principals ($n = 102$ school principals) to the state-level totals⁵ using non-directional z -tests. Table 3-5 displays means, standard deviations, standard errors, z -scores and p -values demonstrating that there was not enough evidence to conclude that the usable sample mean percent retained was unequal to the total population mean percent retained.⁶ For example, as displayed in Table 3-5, for Year 5 (School Year 2003-04) the population

⁵ The number of schools in the Florida system varies by year. See Table 3-9 for the number of schools statewide for years 1999-00 to 2003-04 as reported on the Florida School Indicator's website.

⁶ In other words, the means were statistically equal.

($\mu = 10.722$, $\sigma = 3.514$) was not significantly different from the usable sample ($M = 9.127$, $s = 5.165$), $z = -.454$, $p = .650$. Figure 3-5 visually displays the pattern between mean percent retained of third graders over the 5-year period for each principal comparing the population of principals in the state and the sample of principals usable in this study.

Table 3-5. Third-grade percent retained usable sample principals versus the total population

Year	M	s	S.E. of M	n	μ	σ	σ_{μ} (pop S.E.)	N	z -score	p -value
1	3.077	4.036	.043	72	3.329	1.781	.004	1665	-.141	.888
2	3.118	3.394	.030	100	3.039	1.229	.003	1714	.064	.949
3	3.359	3.974	.035	102	3.338	1.552	.004	1776	.014	.989
4	11.855	7.124	.064	102	14.377	5.250	.012	1829	-.480	.631
5	9.127	5.165	.045	102	10.722	3.514	.008	1848	-.454	.650

Note. Weighted by third grade enrollment at each year

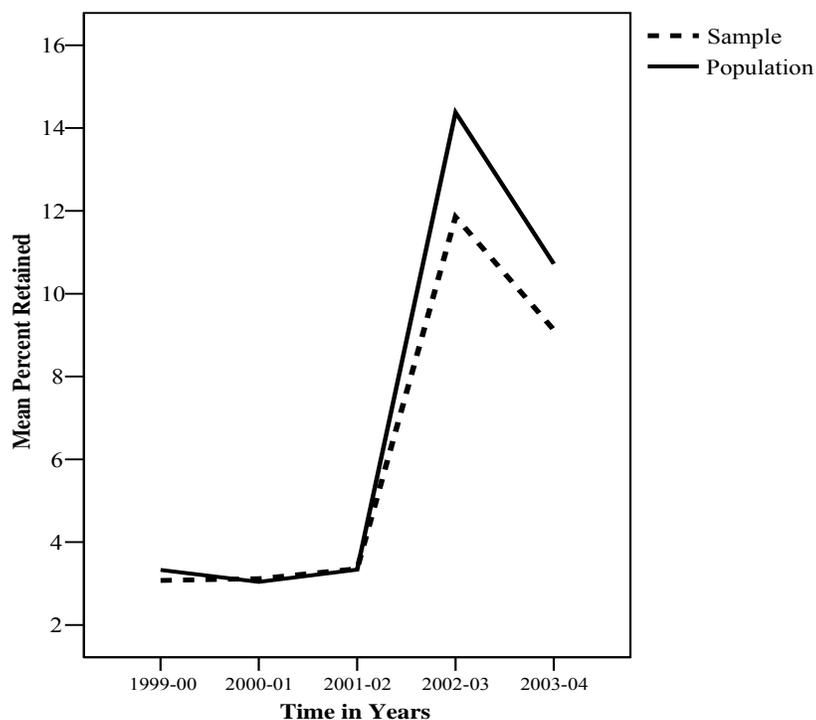


Figure 3-5. Population versus usable sample mean percent retained by year

Additional Variables

Four other variables were designed for the analyses that follow. Two are measures of school-wide level of poverty; while two others consider effects of policy and time.

Socioeconomic status. For this study, the percent of students eligible for free/reduced priced lunch at the school-level was used as a proxy for the socioeconomic status (SES) of the school population. Prior to analysis, I examined the data to determine the variability in school SES for each principal over the 5 years. Since SES lacked variability over years, I created a new variable called SESMEAN (Figure 3-6).

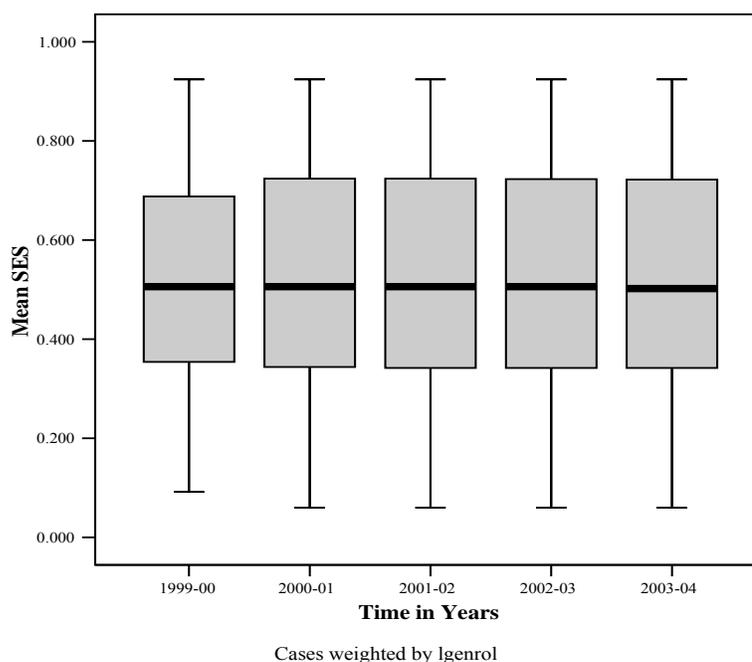


Figure 3-6. Mean SES values over 5-year period for usable sample schools

Mean Centered SES. An additional proxy for poverty used in the analyses that follow is a variable called SESDIFF. This variable subtracts the raw percentage of students eligible for a free or reduced price lunch in a school for each year from the mean percentage of students eligible for free or reduced price lunch for that school over years.

Pre/Post Policy. This is a dummy coded variable used to make a distinction between the pre- and post-policy periods. The years before the start of the retention policy (i.e., years 1, 2, and 3) have been dummy coded as 0; while the final two years (i.e., years 4 and 5) have been coded as 1 (Table 3-1). This Pre/Post Policy variable estimates the effects of the policy.

Trend within Policy Period. A final coding scheme was developed to separately analyze trends over years for pre- and post-policy periods. The variable twt1 was coded to analyze the differences during the pre-policy years; while the variable twt2 was coded to analyze the post-policy differences. Each time point contains a 1-point difference and each trend is centered on zero (Appendix E).

Design

Research Questions

Here, I investigate the possible explanations to the overarching research question “What is the trend of third-grade retention practices before and after implementation of Florida’s third-grade retention policy and how is this trend impacted by other variables?” To answer this question, I use a multilevel statistical model for change (viz., hierarchical linear model, growth curve model, or mixed models) examining within-person change and between-person differences in change beginning with the 1999-2000 school year (Bryk & Raudenbush, 1987; Singer & Willett, 2003). Using longitudinal data, I ask

- How do principals’ retention practices change between policy periods?
- How is this change impacted by poverty level?
- What variables reduce this variation?

The next sections explain how I answered these questions.

Multilevel Framework

A repeated measures analysis of variance (ANOVA) was conducted as a preliminary analysis. Here, the unit of analysis is time where the repeated measure ANOVA treats time as levels of a factor. After inspecting these data structured at the person-level, it was decided that multilevel modeling might lead to a more valid interpretation of the data. In making this decision, I consulted the literature on multilevel modeling (Bryk & Raudenbush, 1987; Hox, 2002; Kreft & de Leeuw, 1998; Singer & Willett, 2003). Kreft and de Leeuw (1998) explain several reasons why one would choose a multilevel model over a typical ANOVA or linear regression. Figure 3-7 shows one limit of examining person-level data. Examining the mean rates of retention may produce misleading results, as it does not consider each individual. Technically, there is variation both between years for a principal and variation between each principal. This can lead to inflated Type I error rates and significant findings that are spurious.

Multilevel modeling easily manages unbalanced cases with missing values (Kreft & de Leeuw, 1998; Singer & Willett, 2003). In fact, unbalanced data is thought to be interpreted more readily using multilevel modeling compared to a technique such as repeated measure analysis of variance (Singer & Willett, 2003). Additionally, in a technique such as an ANOVA, typically cases with missing data are deleted (i.e., listwise or pairwise), whereas they are not in a multilevel model (Hox, 2000; Moskowitz & Hershberger, 2002).

The techniques of *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence* (Singer & Willett, 2003) informed the design of this study. According to Singer and Willett (2003), the first step in building a multilevel model is to determine if it is an appropriate technique. To make this determination, they recommend creating

empirical growth plots to inspect the trajectories using a small sample. To create these growth plots, the data was restructured in a person-period data set where multiple records exist for each individual. If there is sufficient variability between these growth plots, then multilevel modeling for change may be appropriate.

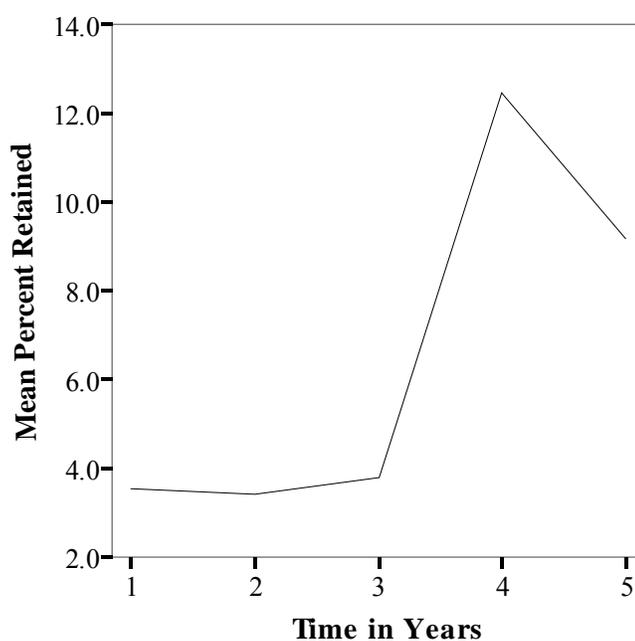


Figure 3-7. Mean retention rates by year. Results of repeated measures analysis of variance of principals' year-by-year retention rates

The figures shown were instrumental in this decision making process. First, the records shown in figure 3-8 display the range of possibilities, illustrating that there is sufficient evidence showing variability amongst principals' retention rates for both the intercepts and the slopes. Figure 3-9 displays the variability amongst all the principals of this study. Most principals' retention rates begin near 0% and rise sharply over the 5-years displayed. Some principals had seemingly higher rates of retention before the state policy, which then declined over time, while a few remained more constant over time. Nevertheless, there is sufficient variation over years to warrant further inspection.

Singer and Willett (2003) further explained three methodological characteristics complementary to multilevel modeling in the context of a longitudinal analysis examining change over time. Here, I state their recommendations and share how I considered their advice. First, they recommended three or more waves of data to detect change. For these analyses, I selected 5 years of data to model change over time. Three years were pre-policy and the latter two were the available data since Florida implemented the retention policy. For my study, I planned to inspect time within each policy period, as well as time before and after policy implementation. Second, they suggested that data should change systematically over time. With the introduction of the retention policy in 2002-03, the systemization of change over time was exemplified by

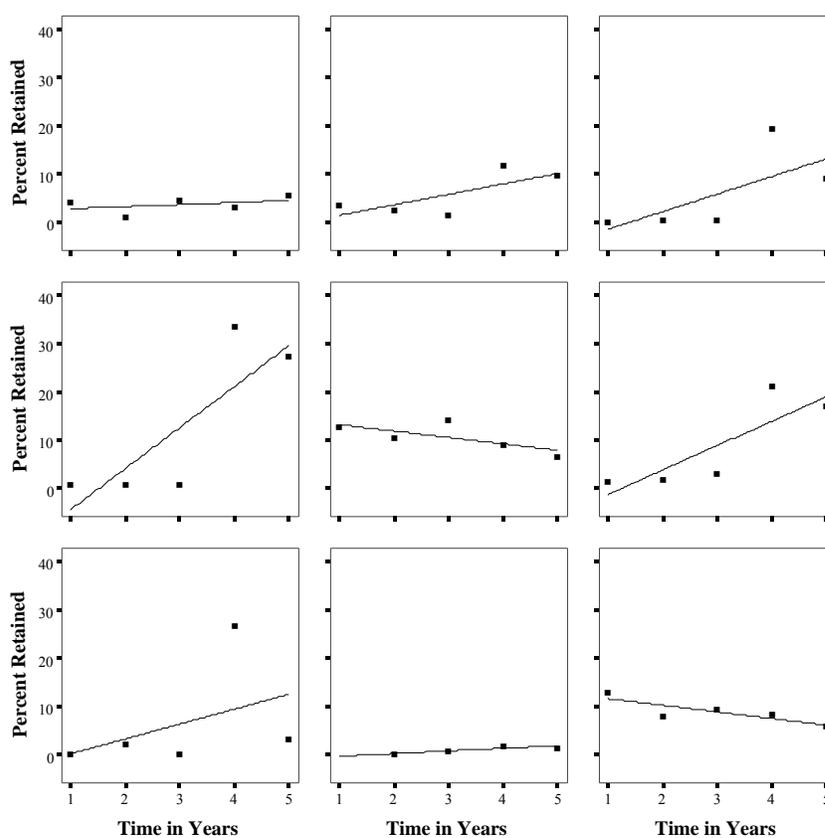


Figure 3-8. Examples of principals' third-grade retention rates over time

the introduction of this policy. Post-policy retention rates changed based upon FCAT results. Finally, they advised researchers to use a noteworthy metric to measure time. For this study, the third grade retention outcome at the end of the school year was used to measure policy years. Since promotional status changes on a year-by-year basis, the school year is the metric for time. Study years were coded as either pre- or post-policy.

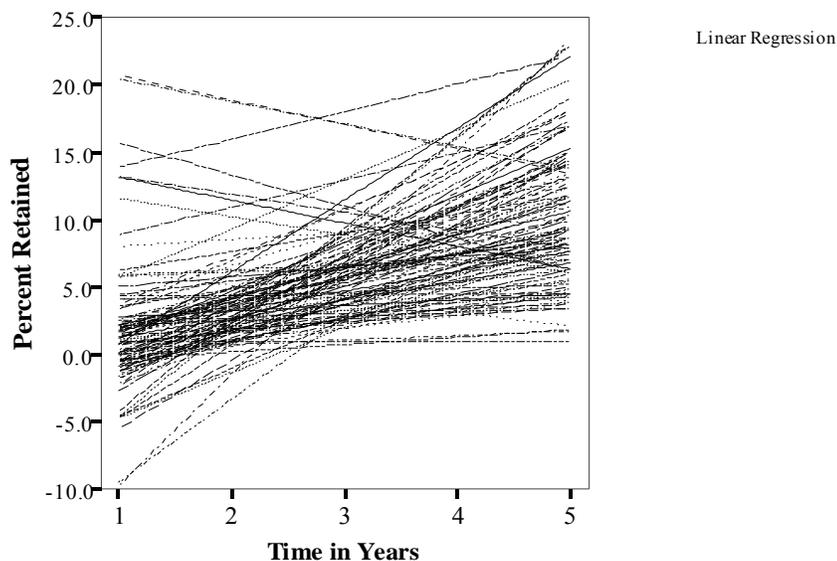


Figure 3-9. Third-grade retention over time. Observed variation in fitted trajectories for all principals (n = 102)

Model Assumptions

All statistical analyses are based on underlying assumptions. Because a multilevel model is more complex, so are the assumptions on which this model was based. Singer and Willet (2003) identified three key assumptions for researchers to consider when determining the tenability for fitting a model to data. These assumptions are the shape, normality, and homoscedasticity.⁷

First, the shape, or linearity, of the dependent variable was inspected to determine whether the means fall on a straight line as a function of the independent variable. That

⁷ All analyses were weighted by the log of school-level third-grade enrollment at each year.

is, does the mean of percent retained for pre- and post-policy connect with a straight line? For variables such as Pre-Post Policy, “there [was] nothing to assess because a linear model is de facto acceptable for dichotomous predictors” (Singer & Willett, 2003, p. 128). Simply put, any two points can be connected by a line.

Second, a key assumption was that these data were distributed normally. That is, the distribution of the dependent variable is normal for both pre- and post-policy. Violations of normality are usually robust; however, for these data, the violation was severe for the dependent variable, percent retained. Figure 3-10 shows the extreme positive skew of the data distribution. The minimum value was 0, and the maximum value was 34.65, suggesting a normal median at approximately 17. However, the clustering of data points near zero resulted in a median of 4.41. The mean of this skewed distribution was 6.29. Using the standard deviation of 6.19 for interpretation would lead to the impossible statement that approximately 95% of retention rates in this sample are between -6.35% and 19.73%. We can refer back to figure 3-9 to see predicted percent retained with negative values.

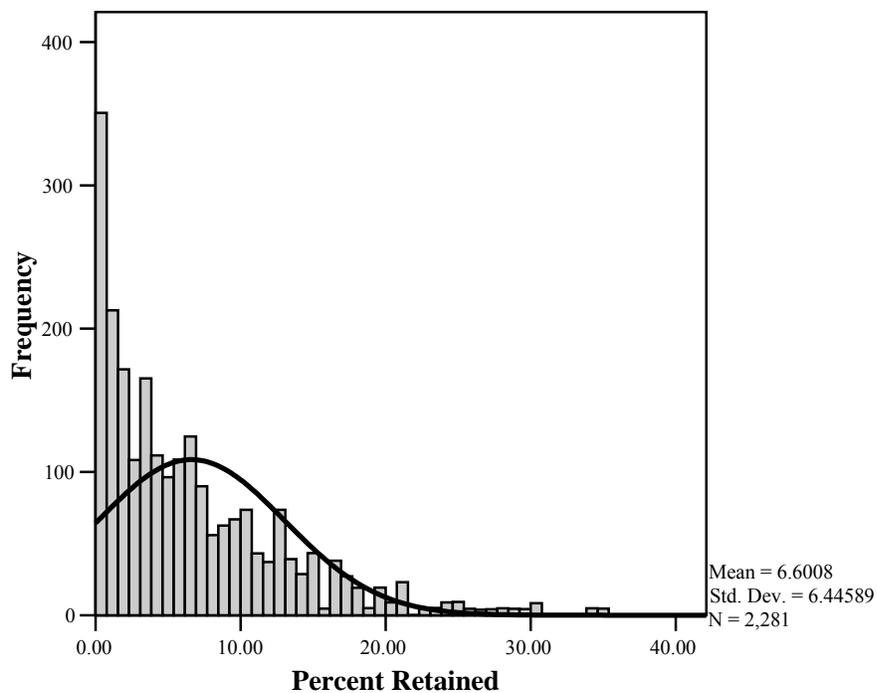
The chosen solution to this problem was a mathematical transformation of the dependent variable into natural log units (McElroy, 2001; Singer & Willett, 2003). Since it is mathematically impossible to take the natural logarithm of a zero value, an alternate metric was needed since this sample contained forty-nine values associated with a zero value. Rather than completely remove these cases using listwise deletion (Wothke, 2000), a decision was reached to add 1 to the percent retained value before taking the logarithmic transformation. The natural logarithm of percent retained plus one normalized the distribution. This transformation of the data permitted zero-values

(Figure 3-11). The resulting distribution had a mean of 1.65, a median of 1.74, and a standard deviation of .917. To assist interpretation, log values can be converted back to the original scale of measurement by taking the antilog and subtracting one (Singer & Willett, 2003). Now, it is more feasible to argue that 95% of the population *log percent retained plus one* are between $(\exp(.00) - 1) = .00$ and $(\exp(3.57) - 1) = 34.52$.

Finally, Singer and Willett (2003) also recommended examining the homoscedasticity, the assumption of equal conditional variances. In other words, we assume that the variance of the dependent variable is the same at all values of predictor, which is to say that the spread of values around the mean dependent variable is the same for all values of the predictor. So, for these data, there are two mean log percent retained values, one pre-policy and the other post-policy. The assumption is that the variation around these means would be the same. Boxplots were examined to determine whether this assumption was met. For the assumption to be met, the interquartile ranges should not be considerably wider or narrower than the other. Figure 3-12 shows the interquartile ranges for two predictors. Visually, we see near equal variances; therefore, homoscedasticity appears to have been met.

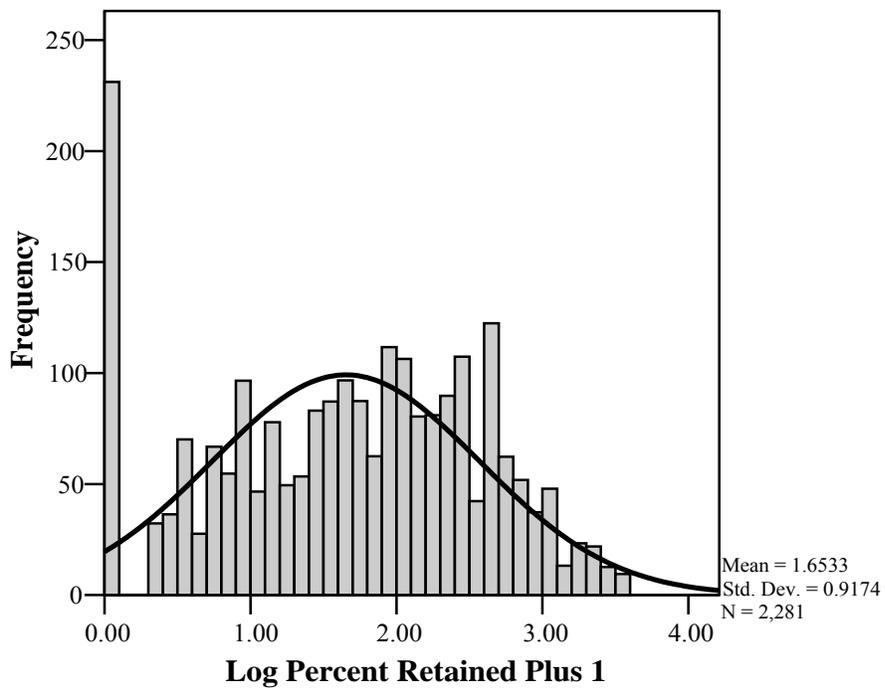
Model Building Process

The rationale for building a multilevel model was twofold: 1) to explain variation in principal retention rates over 5-years while acknowledging the violation of independence since 5-years worth of data are nested within each principal and 2) to add predictors sequentially to test the significance of that predictor, as well as changes in model fit and variance reduction. Many procedures assume statistical independence make some sort of independence assumption. For example, when comparing two groups,



Cases weighted by Igenrol

Figure 3-10. Histogram of percent retained before the transformation



Cases weighted by Igenrol

Figure 3-11. Histogram of percent retained after the transformation

differences between groups should be due to the variable of interest and not some shared feature of group membership (unless those groups are explicitly included in the model). Violations of this assumption can lead to incorrect standard error estimates resulting in a Type I error rate unequal to the conventional nominal $\alpha = .05$ (i.e., the potential for spurious statistical significance) (Singer & Willett, 2003).

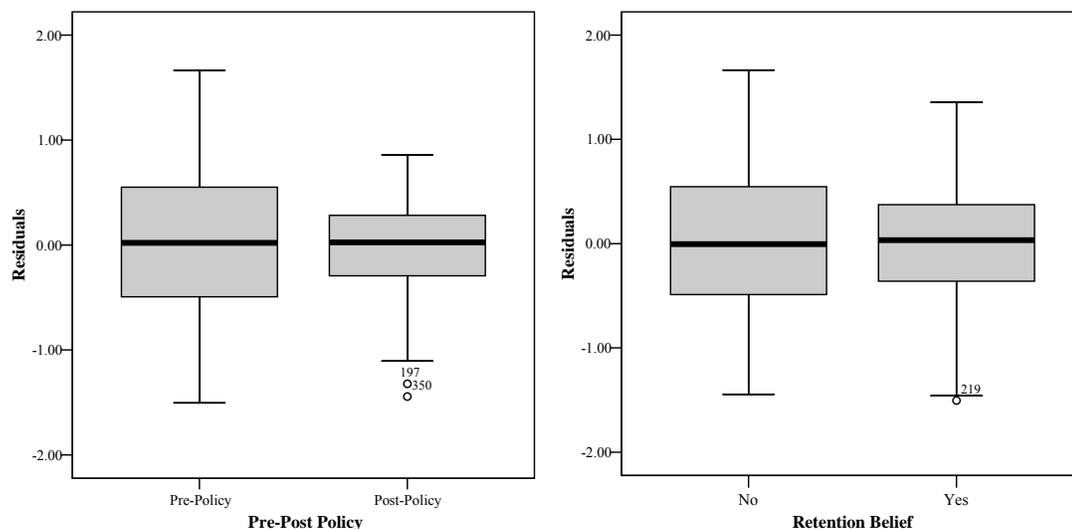


Figure 3-12. Examining the homoscedasticity assumption. Left panel represents the residual by Pre-Post Policy. Right panel represents the residual by Retention Belief.

Within these models, I used the following terms to represent the variables and predictors. In these analyses, $\log(Y_{ij}+1)$ was the natural log of the percent retained for time i in school j with an addition of 1 to permit log transformations of zeros.⁸ Singer and Willett (2003) recommended using a scale to help make the intercept more interpretable. In all models, the intercept parameter β_{0j} indicated the average percent retained for principal where 0 corresponds to Pre-Policy. β_{1j} was the regression coefficient relating time to log percent retained for principal j . γ_{00} was the grand mean log percent retained at

⁸ The addition of 1 to the log is implied unless stated otherwise.

Pre/Post_Policy=0. γ_{10} was the grand mean slope over principals. e_{ij} was the residual for principal j at time i . There are seven main models in this taxonomy labeled Model A through Model F2. Figure 3-13 provides an overview of the model taxonomy included in this study.

To make meaning from this model building process different statistics were examined. Results from these models yield intra-class correlations (ICC) and proportionate reductions in error (PRE). ICC values provided opportunity to compare values associated within a particular model. These values are not compared to other models, whereas PRE values provided context for model comparison. An ICC represents the proportion of total variance that is between-principals for a particular model. A PRE represents the proportionate reduction in variance for a particular variance estimate (e.g., the residual variance) relative to another model. For consistency, all models and sub-models were compared back to Model B, which is the time model called Pre/Post Policy.

Finally, there are many ways to ascertain the fit of the model relative to the fit of other models, including Akaike's Information Criteria (AIC), Bayesian Information Criteria (BIC), and log-likelihood (LL) statistic (Singer & Willett, 2003). This research used Akaike's Information Criteria (AIC) to provide an overall comparison of the model fit. The AIC has the disadvantage of a more subjective interpretation (i.e., no corresponding p -value) and the advantage of being able to compare non-nested models (e.g., one model with a predictor X compared with another model with predictor X and not Y) (Singer & Willett, 2003). Using a smaller-is-better approach, these goodness of fit values allow comparison to Model A, the baseline model.

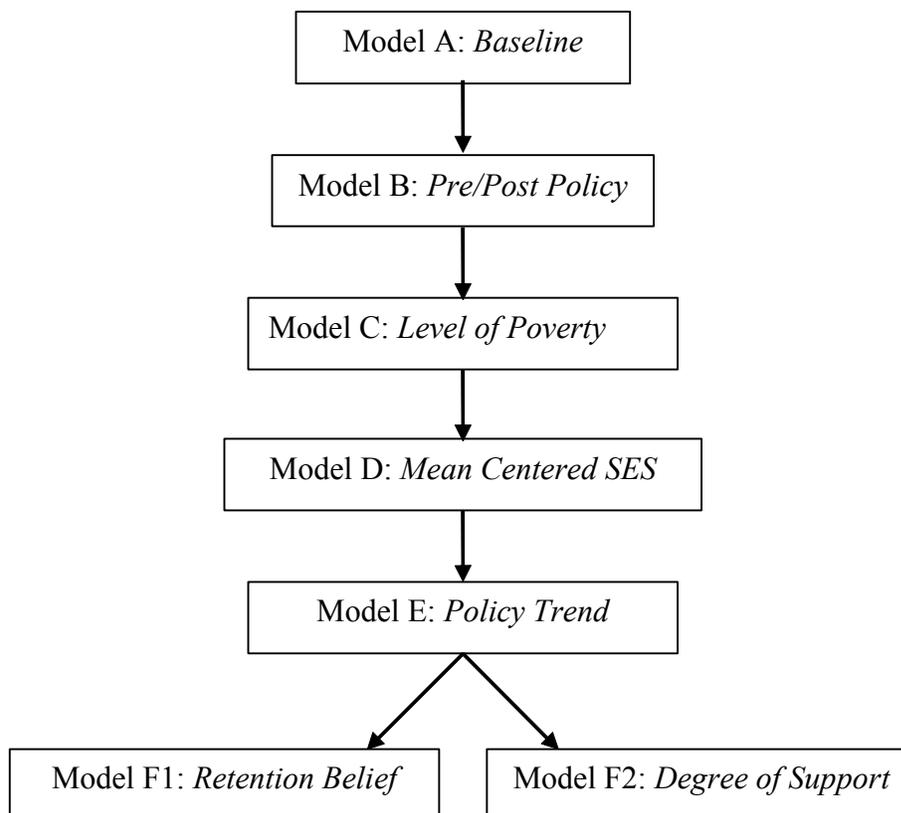


Figure 3-13. Multilevel model for change taxonomy

Study Limitations

Limitations may hinder valid interpretation of these results. First, one study assumption needed to be addressed prior to these analyses. Given the violation of normality, I had to make a decision regarding whether to select a different model to accommodate the data (e.g., specify a binomial distribution with a logit link) or to transform the data to accommodate the model. The latter was selected. Initially, these data were normalized taking the log of the dependent variable, percent retained. This would have resulted in dropping fifty-seven 0% retained values from this study, simply because it is mathematically impossible to take the natural logarithm of zero. Figure 3-14 displays the frequencies of values that would have been removed due to 0% retained by

year. Interestingly, here we see that 0% retained is largely a product of pre-policy retention practices, not post-policy.

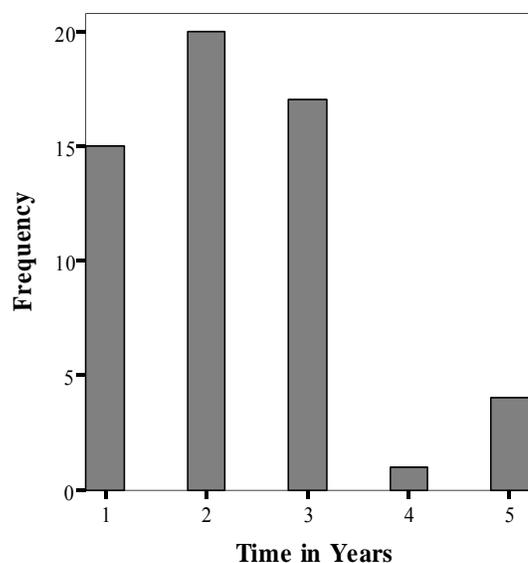


Figure 3-14. Number of zero percent retained values removed by year

This decision could have been justified because of the positive skew of the distribution resulting in the majority of values being close in proximity to the scores being removed. In other words, this was not a case of removing outliers; instead, it was a more sensible technique to handling these data than either permitting the skew or proposing an additional transformation of the data to permit zero-values. However, this technique was not selected because of the richness of the data potentially lost. Because of the pre-post policy focus of my study, it seemed prudent to select a metric that retained these values. Therefore, prior to transforming the percent retained, 1 unit was added to all the values. This adjustment permitted the 0% retained values, and normalized the percent retained distributions.

Finally, after the data collection began, four hurricanes inundated Florida, which may account for the lower than anticipated return on the survey. To minimize the amount

of researcher interaction, influence, or bias, a paper-pencil survey was provided as an alternate option for principals. These self-administered surveys lessened interaction between the researcher and the potential participants and helped minimize social desirability bias (Baumann & Bason, 2004).

CHAPTER 4 RESULTS

Introduction

Results presented for the multilevel taxonomy of models begin systematically with Model A and conclude with Model E. Model building was a successive process with each model building upon the former. A summary table of the results is presented demonstrating how variables, such as time measured as pre- or post-policy, level of poverty, and trend within policy period impact third grade retention in Florida. Other variables, such as a principals' retention belief and the degree of support provided in schools were also tested; however, these variables did not improve model fit and were dropped from the overall analysis (Appendix E). A summary table of the results is available at the end of this chapter (Table 4-1). The resulting models, A through E, help answer the following research questions

- How do principals' retention practices change between policy periods?
- How is this change impacted by poverty level?
- Do retention practices vary more between policy periods or between principals?
- What variables reduce this variation?

Multilevel Models

Model A: Baseline

Model A, the baseline model, is often called a one-way random effects ANOVA model or an unconditional means model (UM) (Raudenbush & Bryk, 2002; Singer & Willett, 2003). It is similar to the typical fixed effects ANOVA; however, the residual

variance is decomposed into two sources (Raudenbush & Bryk, 2002). The sample distribution of the dependent variable, percent retained, was extremely positively skewed. For this reason, the dependent variable was mathematically transformed (McElroy, 2001; Singer & Willett, 2003) to yield a log-normal distribution, as shown in Equation 4-1.

$$\log(Y_{ij}+1) = \beta_{0j} + e_{ij}, \quad (4-1a)$$

$$\beta_{0j} = \gamma_{00} + \mu_{0j}, \quad (4-1b)$$

$$\log(Y_{ij}+1) = \gamma_{00} + \mu_{0j} + e_{ij}. \quad (4-1c)$$

Hence, $\log(Y_{ij}+1)$, the log percent retained for school j at time i , is a function of

- γ_{00} : the grand mean percent retained over all schools before and after policy
- μ_{0j} : the residual for school j representing the distance of that school's percent retained from the grand mean. These residuals are assumed to be normally distributed with a mean of 0 and a variance of σ_{μ}^2 .
- e_{ij} : the residual for school j at time i represent the distance of the school at that time from that school's mean before and after policy. These residuals are also assumed to be distributed normally with a mean of 0 and a variance of σ_e^2 .¹

Model A is a baseline model with no predictors. It serves as a model for comparison to subsequent models. Using Equation 4-1c, the resulting model is shown in Equation 4-2.

$$\log(Y_{ij}+1) = 1.66 + \mu_{0j} + e_{ij}. \quad (4-2)$$

Thus, the grand mean log percent retained was 1.66, $t(96.160) = 34.800$, $p < .001$, corresponding to a grand mean percent retained of $(\exp(1.66)-1) = 4.26\%$. The within-principal variation was significant ($p < .001$) suggesting the need for predictors, while the between-principal variation was not significant ($p = .071$). In fact, 98.3% of the

¹ Variances for future models are assumed to meet the same criteria.

variability in log percent retained was within-principals. The AIC for the baseline model was 1279.43.

Figure 4-1 is a plot that regresses the transformed retention percentages on the principals. It is clear that, over the principals, the predicted regression line suggests little difference while the dispersion of points within principals creates a cloud of variability throughout the graph. It is concluded that efforts should be taken to explain this variability within-principals.

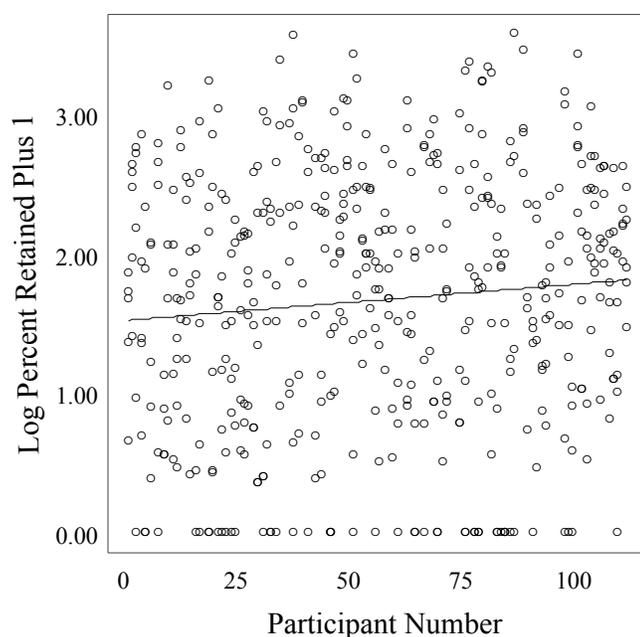


Figure 4-1. Scatterplot of log percent retained plus one

Model B: Conditional on Pre/Post Policy

Model B was the first in this series of growth models aimed to explain the impact policy has on retention rates. For this, a variable representing time before and after the introduction of the retention policy was added to the baseline model. For my study, the time variable was coded such that Pre/Post_Policy=0 corresponds to the school years

1999-00, 2000-01, 2001-02 and Pre/Post_Policy=1 corresponds to the school years 2002-03 and 2003-04.

Distributions were first examined and appear to meet assumptions (Figure 4-2). Pre-policy has a mean of 1.184 which corresponds to $\exp(1.184)-1 = 2.27\%$. The standard deviation of .822 suggests that approximately 95% of retention rates in this sample are between $(\exp(-.46, 2.83)-1)$, or between (0%, 15.95%). Post-policy has a mean of 2.31 which corresponds to $\exp(2.31)-1 = 9.1\%$. With a standard deviation of .595 suggests that approximately 95% of retention rates in this sample are between $(\exp(1.12, 3.50)-1)$, or between (2.06%, 32.12%), as shown in Equation 4-3.

$$\log(Y_{ij} + 1) = \beta_{0j} + \beta_{1j}(\text{Pre/Post_Policy}) + e_{ij}, \quad (4-3a)$$

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \mu_{0j} \\ \beta_{1j} &= \gamma_{10}, \end{aligned} \quad (4-3b)$$

$$\log(Y_{ij} + 1) = \gamma_{00} + \gamma_{10}(\text{Pre/Post_Policy}) + \mu_{0j} + e_{ij}. \quad (4-3c)$$

For the reduced models, β_{0j} was the intercept and is comprised of the grand mean corresponding to Pre/Post_Policy=0 and the residual, μ_{0j} , representing the deviation of school j from the grand mean at Pre/Post_Policy=0. Incorporating Pre/Post_Policy permits the retention rates to vary over time for different schools. Hence, β_{1j} is the slope and is comprised of the grand slope, γ_{10} , which is the change from pre to post policy. The combined model expresses that the log percent retained for school j at time i is a function of the percent retained for pre-policy, the change in percent retained from pre- to post- policy, and the residuals. Using Equation 4-3c, the resulting model is shown in Equation 4-4.

$$\log(Y_{ij} + 1) = 1.17 + 1.14(\text{Pre/Post_Policy}) + \mu_{0j} + e_{ij}. \quad (4-4)$$

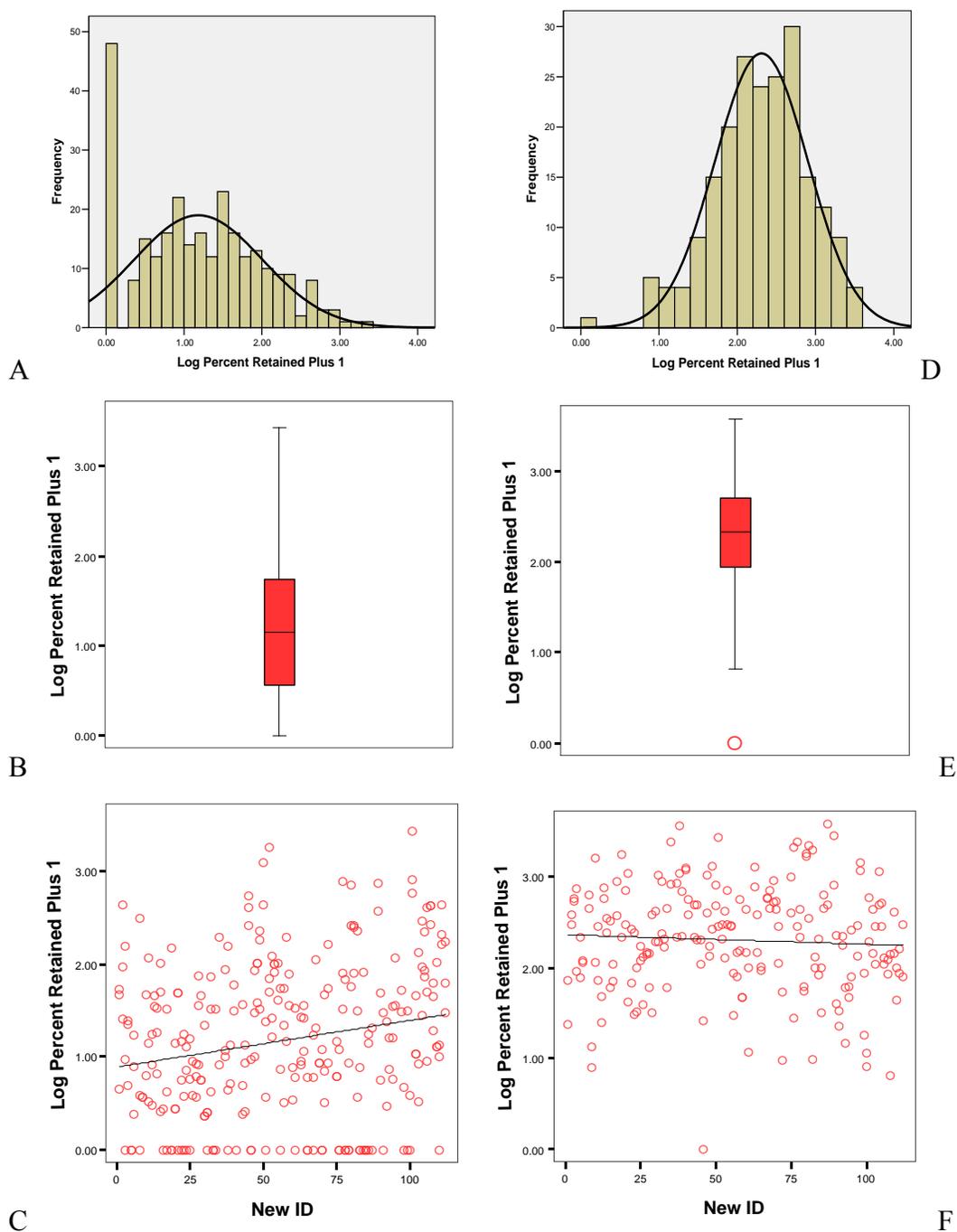


Figure 4-2. Distribution plots of third-grade retention rates. A) Pre-policy histogram. B) Pre-policy boxplot. C) Pre-policy scatterplot. D) Post-policy histogram. E) Post-policy boxplot. F) Post-policy scatterplot.

The estimated mean log percent retained at pre-policy is 1.17, $t(153.09) = 21.692$, $p < .001$, corresponding to $\exp(1.17) - 1 = 2.22\%$. The coefficient for Pre/Post_Policy was

significant at 1.14, $t(375.53) = 19.909$, $p < .001$, suggesting that mean log percent retained increases by $\exp(1.14)-1 = 2.13\%$ from pre- to post-policy. An intra-class correlation coefficient (ICC) of .079 indicates that approximately 8% of the variance in this model is between principals, while approximately 92.1% of the variance is within principals. This means that there is still considerably more within-principal variability (92.1%), than between-principal variability (7.9%). Incorporating policy reduced the within-principal variation by 51.5%, the variation within schools is 1.81, $z = 13.655$, $p < .001$. The proportion of variance within schools was approximately 98%; incorporating pre/post policy as a metric for time reduced it to approximately 92%. Further, model fit improved by 20.8% with the AIC dropping from 1279.43 to 1021.81. The subsequent models will be relative to this model except for the AIC, which is always relative to the baseline model, Model A.

Model C: Conditional on Policy and Level of Poverty

This model introduced a proxy for school-wide level of poverty using the percent of students eligible for a free or reduced priced lunch as the criterion.² Model C is shown in Equation 4-5.

$$\log(Y_{ij} + 1) = \beta_{0j} + \beta_{1j}(\text{Pre/Post_Policy}) + \beta_{2j}(\text{SESMEAN}) + e_{ij}, \quad (4-5a)$$

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \mu_{0j} \\ \beta_{1j} &= \gamma_{10} \\ \beta_{2j} &= \gamma_{20}, \end{aligned} \quad (4-5b)$$

$$\log(Y_{ij} + 1) = \gamma_{00} + \gamma_{10}(\text{Pre/Post_Policy}) + \gamma_{20}(\text{SESMEAN}) + \mu_{0j} + e_{ij}. \quad (4-5c)$$

Using Equation 4-5c, the resulting model is shown in Equation 4-6.

² Using disaggregated SES over years, as opposed to SES for each year, was justified by the lack of variation found in the repeated measures ANOVA described in Chapter 3.

$$\log(Y_{ij} + 1) = .463 + 1.13(\text{Pre/Post_Policy}) + 1.36(\text{SESMEAN}) + \mu_{0j} + e_{ij}. \quad (4-6)$$

The estimated mean log percent retained at initial status is .463, $t(108.427) = 4.859$, $p < .001$, corresponding to $(\exp(.463)-1) = .59\%$. However, this is at $\text{SESMEAN}=0$. The coefficient for Pre/Post_Policy of 1.13, $t(374.619) = 19.758$, $p < .001$, suggests that mean log percent retained increases by $(\exp(1.13)-1) = 2.1\%$ from pre- to post-policy. The coefficient for SESMEAN of .014, $t(97.426) = 8.434$, $p < .001$, suggests that each one-unit change in SESMEAN corresponds to $(\exp(.014)-1) = .014\%$ for both pre- and post-policy.

An intra-class correlation coefficient (ICC) of .030 indicates that 3% of the variance in this model is between principals, while 97% of the variance is between years within principals. This served to balance-out the ICC, so to speak, compared to the variable that examined the effects before and after policy. Model fit improved by 24.5% (relative to the baseline model) with the AIC dropping to 966.34.

Model D: Conditional on Policy, Level of Poverty and Mean Centered SES

The model-building process continued with the inclusion of a variable for mean centered SES called SESDIF . This variable is the percent difference between a school's percent of students eligible for a free or reduced price lunch and the mean of SES over the 5-year period for that school. Model D is shown in Equation 4-7.

$$\log(Y_{ij} + 1) = \beta_{0j} + \beta_{1j}(\text{Pre/Post_Policy}) + \beta_{2j}(\text{SESMEAN}) + \beta_{3j}(\text{SESDIF}) + e_{ij}, \quad (4-7a)$$

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \mu_{0j} \\ \beta_{1j} &= \gamma_{10} \\ \beta_{2j} &= \gamma_{20} \\ \beta_{3j} &= \gamma_{30}, \end{aligned} \quad (4-7b)$$

$$\log(Y_{ij} + 1) = \gamma_{00} + \gamma_{10}(\text{Pre/Post_Policy}) + \gamma_{20}(\text{SESMEAN}) + \gamma_{30}(\text{SESDIF}) + \mu_{0j} + e_{ij}. \quad (4-7c)$$

Using Equation 4-7c, the resulting model is shown in Equation 4-8.

$$\log(Y_{ij} + 1) = .469 + 1.10(\text{Pre/Post_Policy}) + .014(\text{SESMEAN}) + .022(\text{SESDIF}) + \mu_{0j} + \mu_{1j} + e_{ij} \quad (4-8)$$

The estimated mean log percent retained at initial status is .469, $t(109.614) = 5.006$, $p < .001$, corresponding to $(\exp(.469)-1) = .60\%$. However, this is at SESMEAN=0. The coefficient for Pre/Post_Policy of 1.10, $t(380.620) = 19.160$, $p < .001$, suggests that each one-unit change in Pre/Post_Policy corresponds to a $(\exp(1.10)-1) = 2.0\%$ from pre- to post-policy. The coefficient for SESMEAN of .014, $t(98.421) = 8.670$, $p < .001$, suggests that each one-unit change in SESMEAN corresponds to a $(\exp(.014)-1) = .014\%$ change in mean log percent retained when Pre/Post_Policy and SESDIF are controlled at a particular value. The coefficient for SESDIF of .022, $t(471.591) = 3.442$, $p < .001$, suggests that each one-unit change in SESDIF corresponds to a $(\exp(.022)-1) = .022\%$ change in mean log percent retained when Pre/Post_Policy and SESMEAN are controlled at a particular value.

An ICC of .029 indicates that approximately 97% of the variance in this model is between years within principals, while approximately 3% of the variance is between principals. The variation within schools was 1.780, $z = 13.609$, $p < .001$, corresponding to a PRE of 1.5%. The variation between-schools at initial status is .054, $z = 2.664$, $p = .008$, corresponding to a PRE of 65.6%. Model fit improved by 24.7% (relative to the baseline model) dropping to 962.86. There was virtually no change in PRE's, ICC's, or AIC, nevertheless this model was retained because the parameters and variance components were significant.

Model E: Conditional on Policy, Level of Poverty, Mean Centered SES and Trend within Time

The final model formally presented added two variables to account for the policy trend with the pre- and post-policy time periods. These variables, twt1 and twt2, were coded so that twt1 was centered on pre-policy time and twt2 was centered on post-policy time. Model E is shown in Equation 4-9.

$$\log(Y_{ij} + 1) = \beta_{0j} + \beta_{1j}(\text{Pre/Post_Policy}) + \beta_{2j}(\text{SESMEAN}) + \beta_{3j}(\text{SESDIF}) + \beta_{4j}(\text{twt1}) + \beta_{5j}(\text{twt2}) + e_{ij}, \quad (4-9a)$$

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \mu_{0j} \\ \beta_{1j} &= \gamma_{10} \\ \beta_{2j} &= \gamma_{20} \\ \beta_{3j} &= \gamma_{30} \\ \beta_{4j} &= \gamma_{40} \\ \beta_{5j} &= \gamma_{50}, \end{aligned} \quad (4-9b)$$

$$\log(Y_{ij} + 1) = \gamma_{00} + \gamma_{10}(\text{Pre/Post_Policy}) + \gamma_{20}(\text{SESMEAN}) + \gamma_{30}(\text{SESDIF}) + \gamma_{40}(\text{twt1}) + \gamma_{50}(\text{twt2}) + \mu_{0j} + e_{ij}. \quad (4-9c)$$

Using Equation 4-9c, the resulting model is shown in Equation 4-10.

$$\log(Y_{ij} + 1) = .465 + 1.11(\text{Pre/Post_Policy}) + .014(\text{SESMEAN}) + .020(\text{SESDIF}) + .034(\text{twt1}) + -.226(\text{twt2}) + \mu_{0j} + \mu_{1j} + e_{ij}. \quad (4-10)$$

The estimated mean log percent retained at initial status is 0.465, $t(109.806) = 4.952, p < .001$, corresponding to $(\exp(.465)-1) = .59\%$ when Pre/Post_Policy, SESMEAN, SESDIF, twt1 and twt2 are zero. The coefficient for Pre/Post_Policy of 1.11, $t(380.595) = 19.333, p < .001$, suggests that each one-unit increase in Pre/Post_Policy corresponds to a $(\exp(1.109)-1) = 2.03\%$ from pre- to post-policy. The coefficient for SESMEAN of .014, $t(380.595) = 19.333, p < .001$, suggests that each one-unit change in SESMEAN corresponds to a $(\exp(.014)-1) = .014\%$ change in mean log percent retained when Pre/Post_Policy, SESDIF, twt1 and twt2 are controlled at particular values. The coefficient for SESDIF of .020, $t(469.348) = 2.986, p = .003$,

suggests that mean log percent retained changes on average by $(\exp(.020)-1) = .02\%$ when SESMEAN, Pre/Post_Policy, twt1 and twt2 are controlled at particular values. The coefficient for twt1 of .034, $t(385.280) = .720$, $p = .472$, suggests that mean log percent retained changes on average by $(\exp(.034)-1) = .034\%$ when SESMEAN, SESDIF, Pre/Post_Policy and twt2 are controlled at particular values. The coefficient for twt2 of -.226, $t(373.269) = -2.616$, $p = .009$, suggests that mean log percent retained declines on average by $(\exp(-.226)-1) = -.202\%$ when SESMEAN, SESDIF, Pre/Post_Policy and twt1 are controlled at particular values.

An ICC of .030 indicates that approximately 3% of the variance in this model is between years within principals, while nearly 97% of the variance is between principals. Variation within schools is .1.755, $z = 13.570$, $p < .001$, corresponding to a PRE of 2.9%, compared to Model B. The variation between-schools at initial status is .055, $z = 2.726$, $p = .01$, corresponding to a PRE of nearly 65%, compared to Model B. The AIC dropped to 962.89 improving the model fit improved by 24.7% when compared to Model A.

Summary

We conclude the model building process with a review of the underlying statistical assumptions on which this study was designed. Again, although this study modeled growth over a 5-year period, an alternative metric for time was used to predict the effects of retention before and after policy implementation. However, within each policy period the trend within time was modeled. Pre-policy contained three time points and because the trend before policy was relatively stable, it met the assumption. Again, post-policy was dichotomous and linear de facto. Normality and homoscedasticity was assessed at multiple points during the model building process and remained unchanged from the assessment prior to the model building.

Table 4-1. Results of fitting a taxonomy of multilevel models for change to the log percent retained data, Models A through E ($n = 102$)

Model-	A	B	C	D	E
Intercept	1.66***	1.17***	.463***	.469***	.465***
Pre/Post_Policy		1.14***	1.13***	1.10***	1.11***
sesmean			.014***	.014***	.014***
sesdiff				.022**	.020**
twt1					.034
twt2					-0.226**
Residual	3.72***	1.81***	1.82***	1.78***	1.76***
Intercept	.064~	.155***	.056**	.054**	.055**
Between	.064	.156	.056	.054	.055
Within	3.723	1.807	1.816	1.780	1.755
TOTAL	3.787	1.963	1.872	1.834	1.810
Proportionate reduction in error					
PRE-Residual		.515	-0.005	.015	.029
PRE-Intercept		-1.430	.637	.656	.648
Relative To Model-		A	B	B	B
Intra-class correlation coefficient					
ICC-Between	.017	.079	.030	.029	.030
ICC-Within	.983	.921	.970	.971	.970
Goodness of Fit					
AIC	1279.43	1012.81	966.34	962.86	962.89
AIC difference		266.62	313.09	316.57	316.54
Fit improvement (%)		20.80	24.50	24.70	24.70
Relative to Model-		A	A	A	A

~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

CHAPTER 5 DISCUSSION

Introduction

The models showed that time (or the policy period) and poverty at the school level played dramatic roles in the retention rates across this sample. This discussion delves into retention trends to address my overarching question “What is the trend of third-grade retention practices before and after implementation of Florida’s third-grade retention policy and how is this trend impacted by other variables?” The model building process (Figure 3-13) guided my study to understand how particular variables such as policy, and level of poverty impacted retention practices over a 5-year period. This chapter discusses the results from this model building process to answer the more specific research questions.

- How do principals’ retention practices change between policy periods?
- How is this change impacted by poverty level?
- What variables reduce this variation?

Then, we turn our attention the policy implications and recommendations, as well as suggestions for future research.

Research Question 1: How Do Principals’ Retention Practices Change between Policy Periods?

The model building process originated with a degenerate model providing a baseline to compare model fit. In considering the baseline model, significant variation was found both between principals and within principals over years that needed

explanation. Over 98% of the variation was found within principals over years. In the final model, we gained insight regarding principals' retention policies between policy periods. The key feature of Model E was its representation of the data. First, it adjusted for between- and within-principal SES. Then, it revealed a significant difference between the two policy periods. Furthermore, although it did not reveal a significant trend between the three pre-policy years, it did reveal a significant difference between the two post-policy years.

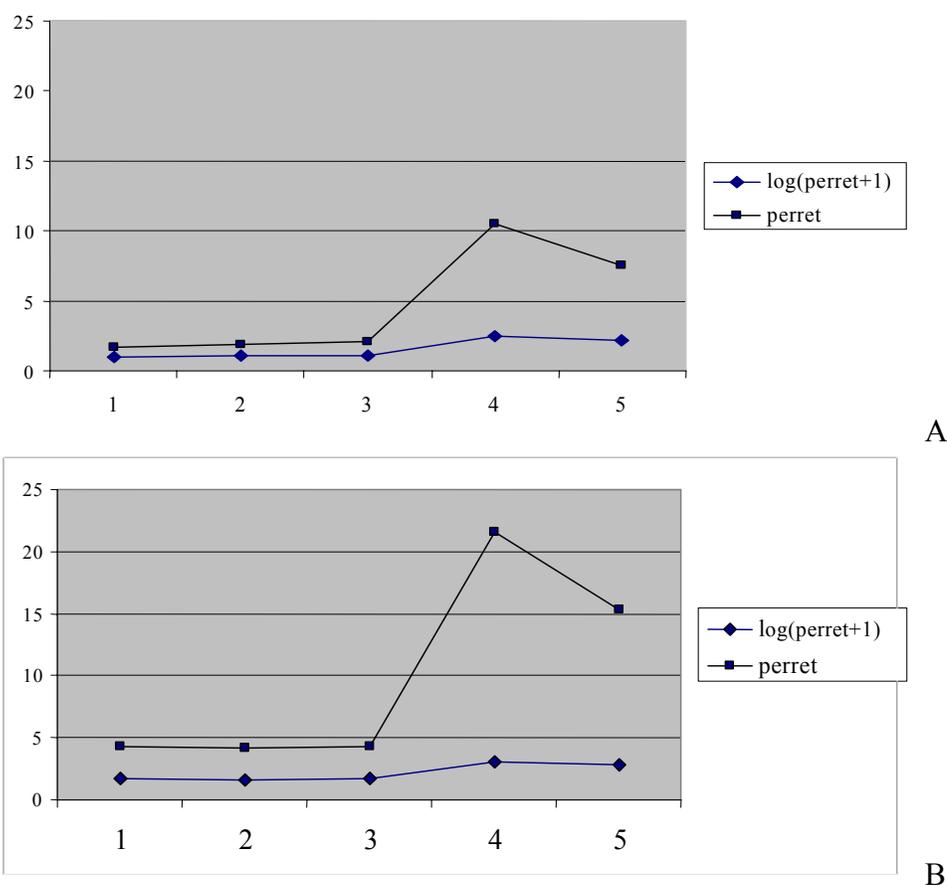


Figure 5-1. Examples of mean predicted retention trends over time. Two lines are shown in each graph. One line represents the predicted log percent retained value ($\log(\text{perret}+1)$), while the other is unlogged showing percent retained units (perret). Because two metrics are shown in one graph, the y-axis is simply one-unit intervals. The x-axis shows time in years, with years 1, 2, and 3 being pre-policy, and years 4 and 5 being post-policy. A) School from mid-size district, mean SES 52%. B) School from same district, mean SES 91%.

The fact that pre-policy was not significant is worth discussion. This means that between the pre-policy years, the change in mean retention rates can be considered equal. Retention practices were stable before implementation of the state retention policy. On the other hand, the post-policy years (years 4 and 5) had mean retention rates that were significantly different from each other. Two examples of the mean predicted values help us understand how retention trends have changed over time (Figure 5-1). Within both examples, on visual inspection, it is clear why the pre-policy trend was not significant, and the post-policy trend was significant.

Research Question 2: How Is This Change Impacted by Poverty Level?

Since the level of poverty is widely regarded as a factor that impacts student achievement, school level of poverty was included as a predictor. Using the mean SES of each school over the 5-year period confirmed that, even in this sample, poverty plays an influential role in retention rates for third graders in Florida. Mean SES rates greatly affect retention practices. Schools with the highest levels of poverty (greater than 67%) consistently retained more students over this 5-year period (Figure 5-2). Principals from schools with the lowest range of poverty (33% or less) retained fewer students. Clearly, these principals have had fewer practical experiences with retention over this 5-year period than principals of higher-poverty schools; however, even they have felt the impact of policy with more students experiencing retention in pre-policy years.

In further examining the pre-post policy trend, I inspected the poverty levels over time. Here we see the relationship between pre- and post-policy (Figure 5-3). As poverty increased, the percent of students retained also increased. Poverty is a great predictor of retention rates and so policy impacts students from higher-poverty schools more greatly. In other words, lower-achieving students in higher-poverty schools were more likely to

be promoted pre-policy and because of this, the mandate impacted more children in higher-poverty schools than students in lower-poverty schools.

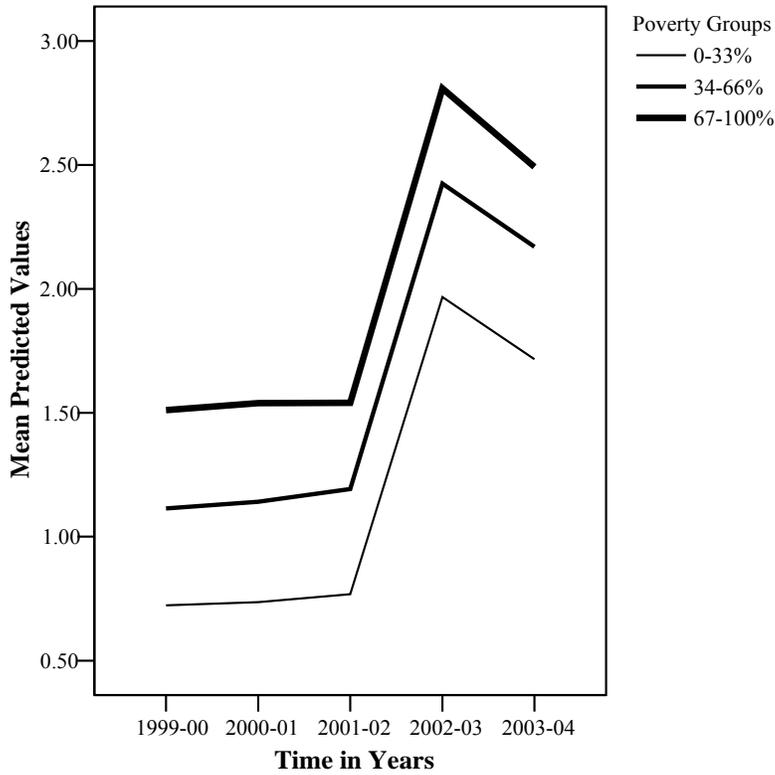


Figure 5-2. Mean predicted third-grade retention rates by poverty level over time

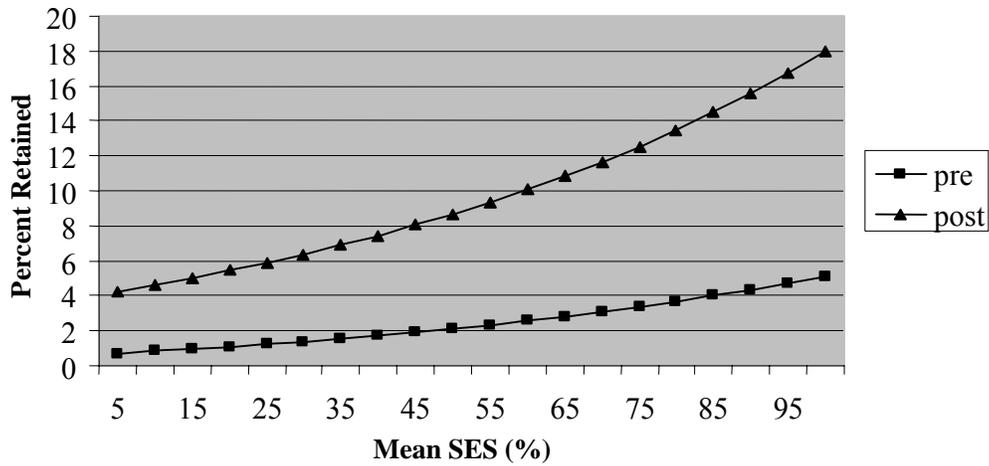


Figure 5-3. Example plotting retention rates using principals' mean SES values before and after policy implementation

Research Question 3: What Variables Reduce This Variation?

With the focus on change over time, it was hoped that the measure for time would help explain much of the variability. Model B, the Pre/Post Policy model, added growth explicitly into the model with the log units of percent retained showing a significant increase from pre- to post-policy. The most notable impact of Pre/Post Policy model is improved model fit compared to the baseline model. It was somewhat surprising that variation within principals over the years decreased by only 8% with the proportion of variance decreasing from 98% to 92%. The Pre/Post Policy Model suggested a marked change in retention rates for most schools in 2002-03 suggesting further analysis was required.

Attempts were made to move the model building forward beyond policy period and level of poverty examining retention belief and the degree of support, however improved model fit was not achieved. Although model fit was not statistically significant, we can examine these patterns descriptively to understand what they reveal in this context (Onwuegbuzie & Leech, 2004). Nevertheless, we first work to understand what these mean predicted values suggest when considering principals' current retention beliefs (Figure 5-4).

Since Florida did not have a formal retention/promotion policy before 2002-03, it is interesting to consider possible explanations for the predicted retention patterns over the policy periods, especially considering retention belief. As learned in Chapter 2, the NCLB Act of 2002 is one of most wide sweeping vehicles of educational reform. The mean predicted retention values showed a pattern shift around the 2001-02 school year when NCLB was enacted. It is certainly possible that NCLB had some sort of impact on the beliefs of principals, especially in light of the stronger focus on accountability. Since

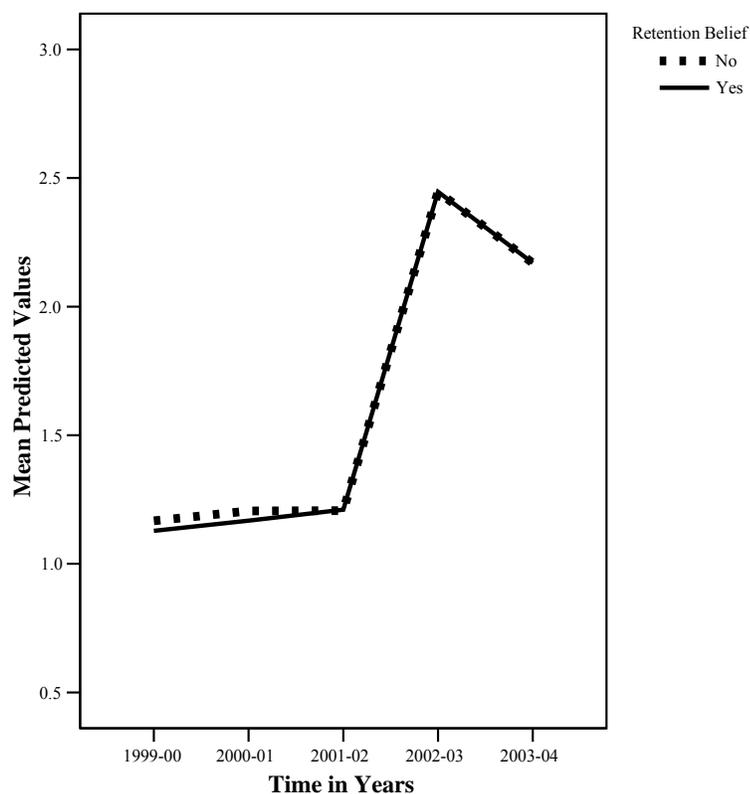


Figure 5-4. Mean predicted third-grade retention rates by principals' current retention belief over time.

the signing of NCLB, the focus on accountability has been prominent. Perhaps principals with the lowest achieving students feel they are in the spotlight. If students in these schools do not meet the AYP mandate, the whole school faces severe penalties, including possible school restructuring and reduced funding. Perhaps Florida principals believe placing their trust in the state mandate to hold back students not reading on grade-level will help their schools achieve AYP. Further information is needed to explain these patterns before and after policy.

In this current policy era, the relationship between retention beliefs and practices yields interesting discussion. Past research of retention beliefs found that 74% of principals favored retention for students who were not meeting the grade-level standard

(Byrnes, 1989); similarly, my study found that 69% of principals believe retention can benefit third-grade students who fail the FCAT. However, as shown in figure 5-4, it appears that the Intent works because the Florida retention policy does not permit principals' retention beliefs to play a factor in retention decisions. My study shows that regardless of a principal's retention belief, students are retained based on their FCAT performance, not a principal's retention belief. In a nutshell, as a result of the policy mandate, retention beliefs as reported and analyzed in my study did not predict retention rates.

Future research may explore why some principals who do not believe in retention, held back more students before the start of the retention policy than those principals reporting they believe in retention. A possible explanation may be that principals' beliefs have swayed since they are now obligated to uphold state policy regardless of their personal beliefs. It could be possible that principals disavowed their personal beliefs so as not to impair their professional judgment. Perhaps, before policy, subtle shifts were the result of teachers or parents overriding the principals' belief to veto retention (Jimerson et al., 1997) or pressure to achieve on high-stakes assessments; whereas this option has been severely limited since the implementation of the retention policy.

It is perhaps more interesting to discuss these mean predicted retention values in relation to poverty level over time to further illustrate the patterns. Figure 5-5 displays four panels, 1 through 4, plotting principals' current retention belief and the mean predicted retention practices of principals over a 5-year period. Panel 1 shows the pattern of all principals using the mean of percent retained (also shown in figure 5-2). Panels 2, 3, and 4 isolate the level of poverty with Panel 2 showing principals who work in schools

with the lowest levels of poverty (lower than 33.9%). Principals of lower-poverty schools have lower retention rates when compared to schools with higher rates of poverty. Panel 3 shows schools in the mid-range of poverty, with between 34.0 to 66.9% students eligible for a free or reduced price lunch, retain more students than their lower poverty counterparts. We can see that principals in these lower- and mid-range poverty schools retain nearly the same numbers of students regardless of their retention belief. This suggests that principals of lower- and mid-range poverty schools are possibly more policy compliant than other types of schools.

Finally, panel 4 shows how principals who work in schools with the highest levels of poverty retain students differently before and after the start of the retention policy. This panel also revealed one of the more intriguing findings of this study. Prior to policy, principals of higher-poverty schools who report not believing in retention held back more students than retention believers. That is, the model predicted that non-retention believing principals of the higher-poverty schools retained more students than retention believers. Perhaps these are the principals who do not believe in retention based on results from high-stakes assessments, but do believe that students can benefit developmentally from more time, or who believe they are giving the “gift of time.” This “gift of time” refers to the conventional wisdom that providing children with more time to learn will give them the opportunity to increase their academic potential (Ames, 1986; Graue, Kroeger, & Brown, 2003).

In response to variables affecting retention practices, we shift our discussion to the supports offered to students at risk for retention as disaggregated by degree of support

(Figure 5-6). First, I will discuss predicted third-grade retention rates and current degree of support over time. Schools currently using nearly all the supports (5 or 6 out of 6)

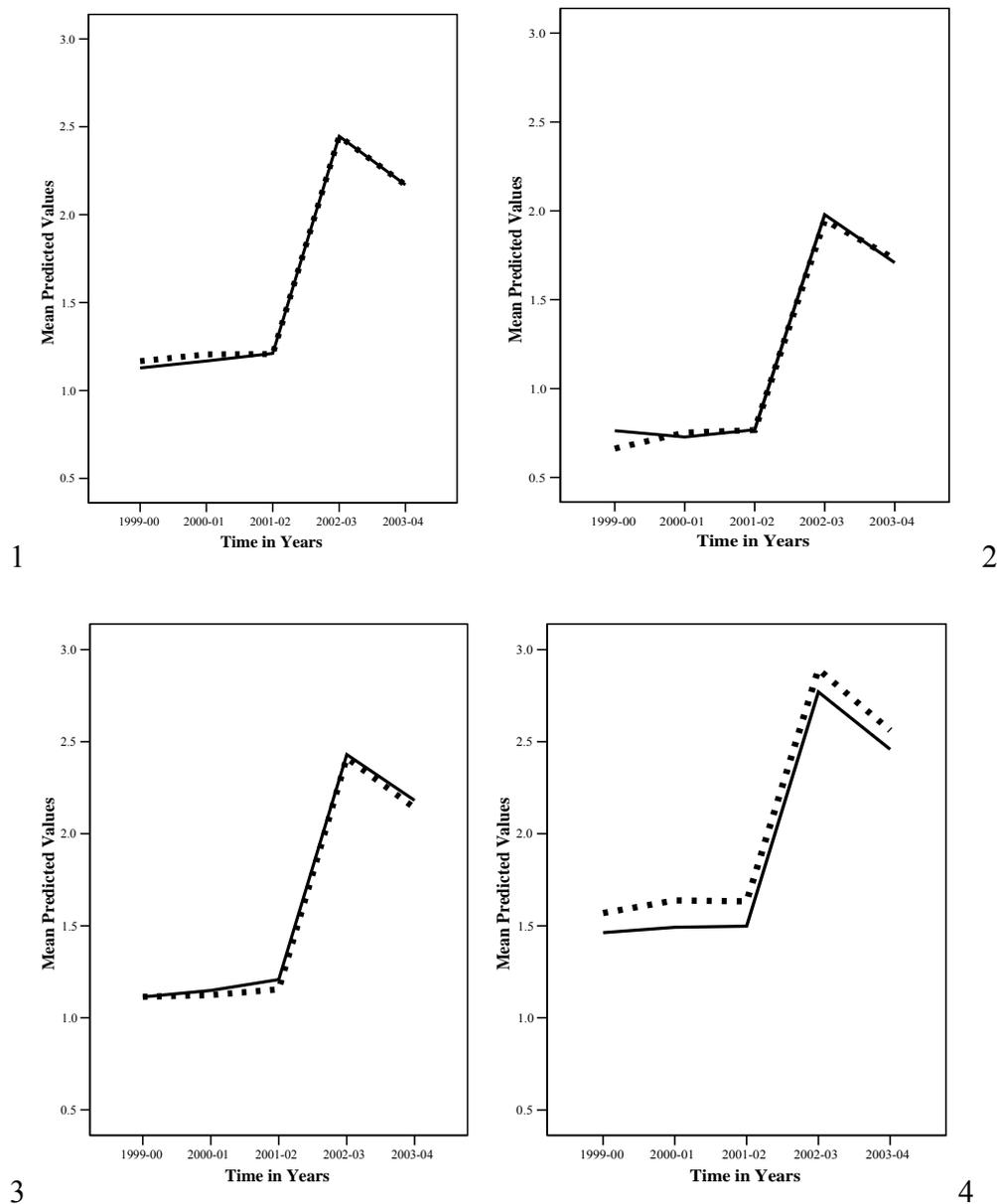


Figure 5-5. Principals' retention belief by the mean predicted third-grade retention rates over time disaggregated by poverty group. 1) Mean level of poverty for all sample schools. 2) Sample schools with less than 33.9% poverty. 3) Sample schools with between 34.0 to 66.9% poverty. 4) Sample schools with more than 67.0% poverty. (Solid lines equal retention believer; dashed lines equal non-retention believer)

have the highest retention levels. When we compare the current degree of support to past retention rates, it could be said that schools with higher rates of retention now provide as many supports as possible to prevent retention. Since we do not know whether these schools used a similar number of supports in years past, we cannot draw conclusions about whether more or fewer supports will benefit students. However, this variable may permit future research to make such determinations.

Interestingly, when including poverty level and policy period in the model building process, degree of support did not account for much of the variation, nor did it achieve significant *p*-values. Figure 5-7 displays four panels, 5 through 8. Panel 5 shows the pattern of all principals using the mean predicted percent retained (also shown in figure 5-6). Panels 6, 7, and 8 isolate the level of poverty with Panel 6 showing principals who work in schools with the lowest levels of poverty (lower than 33.9%). As we have seen

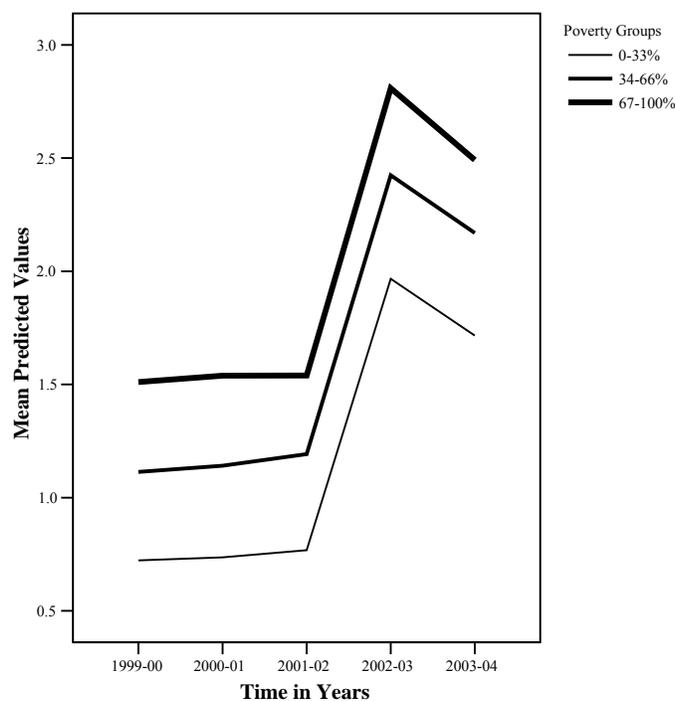


Figure 5-6. Mean predicted third-grade retention rates by degree of support over time

fluctuate; however, what is worth noting here are the similarities in retention patterns regardless of the degree of support provided within the school. This is especially evident in Panels 6, 7 and 8 where depending on the poverty level of the school, similar patterns of students were retained at each time point, regardless of the number of supports used. Based on the past retention patterns of this sample, there does not appear to be an explicit pattern explaining why particular schools use differing amounts of support. It is worth mentioning that principals of the highest-poverty schools did not support using only 1 or 2 supports in their schools; high-poverty schools reported using a minimum of 3 supports. This is important because the pressure of accountability mandates may make principals of higher-poverty schools feel more obligated to offer their students as many supports as possible.

Considering the current supports in use, it was interesting to examine the past retention practices of these schools. Some may believe that the quantity of support leads to success by proxy; however, we do not have data to support this assumption. From these data, we can see that a larger quantity of support does not necessarily raise achievement for students—more is not necessarily better. Poverty level greatly influences the mean predicted retention patterns. Panel 5 suggested that “doing less better” works; however, disaggregating by poverty level loses that inference.

Policy Considerations

The third-grade retention policy appears to be working as intended. As FLDOE press releases from 2004 and 2005 each report, there are: “More 3rd graders reading at or above grade level than ever before” (Florida Department of Education, 2004b, p. 1; 2004c, p. 1). But are they? Third-grade students began taking the FCAT in 1999-00,

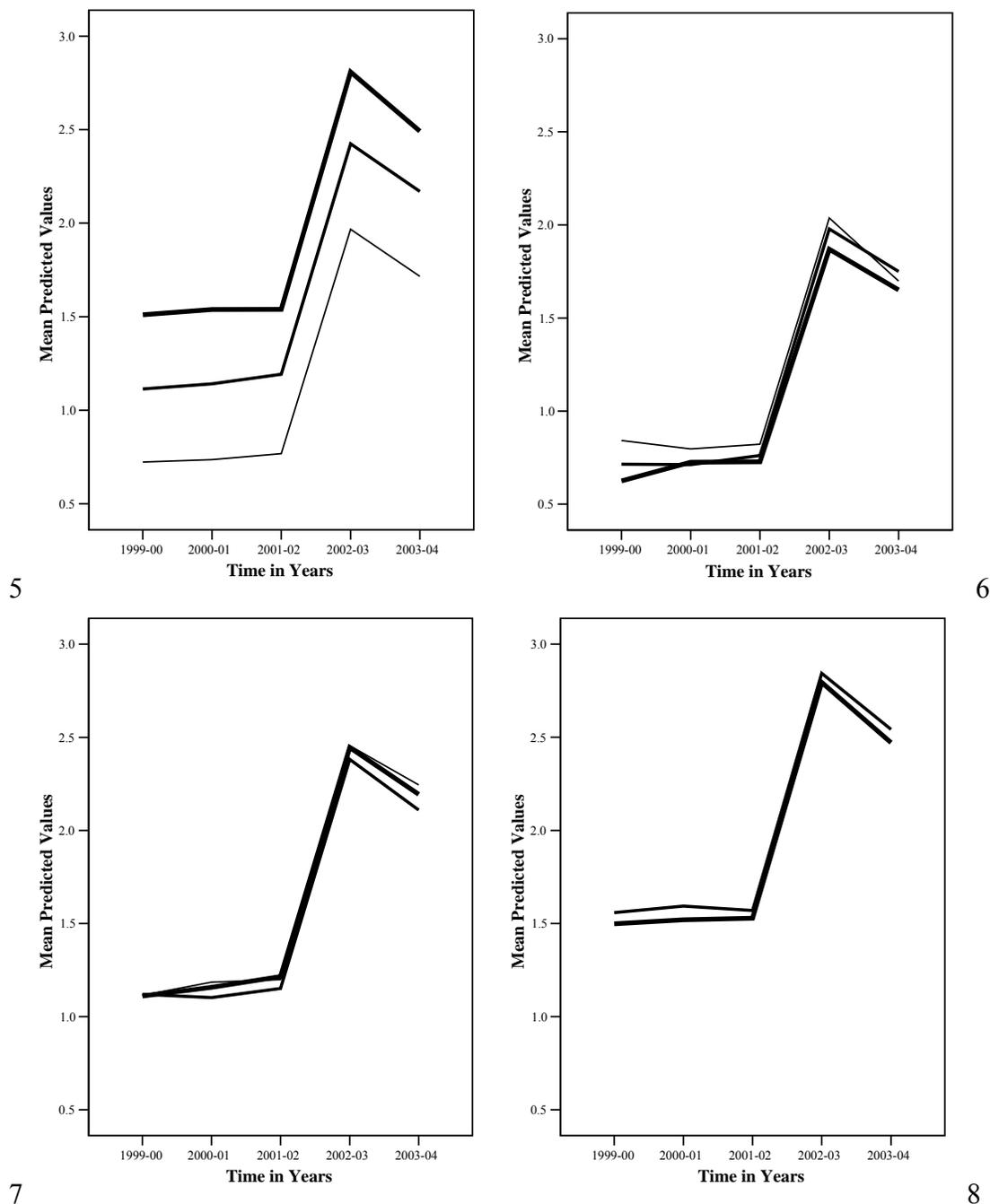


Figure 5-7. Degree of support by mean predicted third-grade retention rates over time disaggregated by poverty group. 5) Mean level of poverty for all sample schools. 6) Sample schools with less than 33.9% poverty. 7) Sample schools with between 34.0 – 66.9% poverty. 8) Sample schools with more than 67.0% poverty. (Thickest lines equal any 5 or 6 supports; medium thickness lines equal any 3 or 4 supports; thinnest lines equal any 1 or 2 supports)

however their scores were not publicly reported during the first two years. Table 5-1 shows the percent of third graders scoring at Level 1 since 2001-02, the year prior to the start of the state policy. Following the 2001-02 administration of the FCAT, before the retention policy, 29% of third graders achieved Level 1 (Florida Department of Education, 2004c, 2005d). Although the percentage of students scoring Level 1 has decreased three percent, dropping from 23% to 20% since the policy start, the number of third graders has also grown. We need to remember why the number of third graders has grown in post-policy years—retained third-grade students have bloated enrollment figures, so proportionately it does appear that achievement is on the rise. Appearances can be misleading.

With the release of the Greene and Winters (Center for Civic Innovation at the Manhattan Institute, 2004) evaluation of Florida’s retention policy then-U.S. Secretary of

Table 5-1. Third-grade FCAT reading achievement Level 1

School year	Level 1 (%)	Number of third graders scoring Level 1	Number of tested third-graders
1999-2000	—	—	—
2000-2001	—	—	—
2001-2002	29	53,980	186,139
2002-2003	23	43,265	188,107
2003-2004	22	45,416	206,435
2004-2005	20	40,595	202,975

Note: Author’s own calculations using publicly available data from Florida Department of Education (<http://www.firn.edu/doe/sas/fcat.htm>).

Education Rod Paige issued a statement praising the report: “I applaud schools in Florida and across the nation that are working to help our children by stopping the dangerous practice of social promotion” (U.S. Department of Education, 2004a). Regardless of the

actual impact on reading achievement, the law works as it has eliminated social promotion in third grade.

Even with these accolades, several concerns are masked within this policy. First, how are third-grade students who score Level 2 supported as they progress through the grades? These are the students on the fringe. Though they did not fail the FCAT in the eyes of the policy, Level 2 students scored with limited success, and technically have not passed the assessment. While it is not advocated here to hold back all students who do not pass the FCAT, we must seriously consider the consequences for both those students who are retained and those who just make it to fourth grade.

As we learned in the review of the Florida third-grade retention policy, the FCAT has had an objective impact on retention practices. Prior to the policy, if retention was chosen, it was a subjective decision, including some combination of the classroom teacher, principal and parent to make a retention decision. Some schools may have relied upon test scores and classroom performance; however, these decisions were decided independently, varying from district-to-district, or even school-by-school. With the retention policy, when a third-grade student does not pass the FCAT, the score guides this decision. Teachers and principals have tangential involvement unless a promotion for good cause, where applicable, goes into effect. We need to remember that over 27,000 of Florida's third-grade students were subject to the retention policy after the first year. After experiencing retention, 41% of retained students again failed the FCAT after their second year in third grade (Florida Department of Education, 2004c).

I uncovered several unintended consequences when promotions are the result of a good cause exemption. Local school districts maintain control over how many of these

good cause exemptions are implemented. However, in the first year of the retention policy, there were generous amounts of flexibility where districts could exercise some degree of freedom in how they exercised the good cause exemptions. Let us revisit the good cause exemption data presented in table 2-1 and possible implications from the promotional patterns.

The newly implemented good cause portfolios were designed to help students who had demonstrated consistent readiness to be promoted to fourth grade; however, due to poor test scores, those students could not be promoted. The classroom teacher with the approval of the school principal and later the district superintendent were able to use a good cause portfolio as a means to by-pass retention. Nearly 1,500 retained third graders were promoted to the fourth grade after the first year of the policy with a good cause portfolio, while after the second year of the policy more than double the number (almost 3,500 retainees) were promoted under the good cause exemption. There are several possible reasons for this increase; however, the most likely may simply be timing. In the first year of the policy, the FLDOE announced the portfolio option only weeks before the close of the first policy year. It is possible that there were students who may have been eligible for this exemption, but due to circumstances, did not receive this option. In the second year of the policy, schools were able to lead a more coordinated effort to begin the good cause portfolio process for at-risk students closer to the beginning of the year.

Three times the number of students ($n = 1924$) were promoted to fourth grade because they had already been retained *twice* with two or more years of remediation from the first to the second year of the policy. Because the 2004-05 school year is the third school year since this policy start, it is anticipated that there will be a substantial increase

again with this good cause exemption. This is because third graders who were retained and subsequently re-retained have now met the third-grade retention limit and will automatically be promoted to fourth grade, regardless of FCAT performance. Because of the high stakes involved in these retention decisions, research needs to be expanded to the primary grades (K to 3) to study Florida's retention patterns. Are primary grade students being retained in earlier grades to thwart third-grade retention? Or are lower-achieving primary grade students being socially promoted because the new state policy ensures they will be retained in third grade unless their achievement improves? Retention decisions for students in grades K to 2 remain at the discretion of the individual school and district.

Also startling is the use of the exemption for SWD who have been retained one time. You may recall that for students to be promoted for good cause to fourth grade, in most cases, the student must be retained at least one time before third grade. The number of SWD who were promoted for good cause more than doubled, increasing from about 3,600 students to nearly 8,000. Prior to the Florida retention policy, research indicated an expanded use of placement in special education programs as a means of meeting accountability pressures (Figlio & Getzler, 2002). Other studies conducted outside Florida have examined retention trends to demonstrate that increasing numbers of students are being placed in special education programs to avoid accountability scrutiny (Allington & McGill-Franzen, 1992; Roderick & Nagaoka, 2005; Roderick et al., 2000). However, we need to further expand research in Florida to understand whether this good cause promotion is used to "game the system" (Figlio & Getzler, 2002).

Even still, national policy has a rule, often called the 2% rule, limiting the percent of students eligible for alternative assessments as part of AYP (Elmore, 2002). Schools

or districts who might be tempted to circumvent the state retention policy by placing more students in special education programs; yet, in the end, all sub-groups of students, including those placed in special education, will have their achievement score disaggregated in AYP calculations. To meet AYP, schools and districts must test, on average, 95 percent of the population to meet the criteria for these calculations (U.S. Department of Education, 2004b).

Interestingly, one of the least widely used good cause exemptions is the ability for the retained students to demonstrate proficiency on an alternative assessment. After the first year, about 3,300 retainees were promoted using this option, while just over 3,800 were promoted the following year based on their performance on an alternative assessment. Only students who attended summer school were eligible to take the alternative assessment. As explained in chapter 2, for students who take the SAT-9, they must achieve at the 51st percentile¹ (something that their Level 2 scoring peers were not held accountable to pass).

As retained students progress through the state system, we must continue to study the long-term effects of this policy, including dropout patterns. This retention policy may have an interesting impact on dropout rates in the years to come. By removing a segment from the 2002-2003 cohort of third-grade students (who were first impacted by the state policy), we await Florida's graduation rates for 2012. This is the year that students first affected by the third-grade retention policy would have been scheduled to graduate from high school. Time will tell whether we will see a marked improvement in graduation rates by shifting students who failed the third-grade FCAT to a different cohort.

¹ Student taking the SAT-10 must achieve at the 45th percentile or higher.

Policy Recommendation

Florida's third-grade retention policy (Chapter 2) strived to end social promotion by identifying K to 3 students who lack the proficiency to pass the reading portion of the third-grade FCAT and consequently failing the third graders who do not pass the FCAT. In fact, my study found that Florida's retention policy has successfully eliminated social promotion in third grade producing students who should be prepared to enter fourth grade reading on (or close to) grade level. In other words, the policy was intended to solve the problem of students not reading on grade level; however, data show that policy is not accomplishing that goal since post-policy retention rates increased, on average, 300% since the start of the policy (Florida Department of Education, 2005e).

In a 2004 memorandum issued by the FLDOE, Florida school districts were provided a lengthy list of requirements as outlined in the Intent to help students meet the policy goal (Warford & Openshaw, 2004). While the Intent does support identifying struggling students before third grade in order to intervene in ways that would reduce the likelihood of later mandated retention in grade, one may wonder how these identified students are also supported when school is not in session. That is, schools must offer summer school, but only to third graders who fail the reading portion of the FCAT. It seems that a key element is missing from the Intent menu of options for schools to select in efforts to support their struggling readers: Summer school should be available to targeted students prior to third grade.

Chicago's retention researcher found that students retained in third and sixth grades were actually identified as struggling readers as early as first grade (Roderick & Nagaoka, 2005). Research in Baltimore and Michigan also support increased attendance in summer school programs for struggling readers. In Baltimore, Borman, Benson and

Overman (2005) found a positive relationship between students who regularly attended summer school and student achievement. In other words, students who spent a greater number of weeks in a summer school program showed academic improvements; while in Michigan, researchers studying summer school programs for lower-achieving K to 3 students found significant benefits for these students (Paris et al., 2004).

While Roderick and Nagaoka (2005) suggest that alternative intervention approaches, such as summer school for identified struggling readers, can be an expensive prospect, it may not be as costly in the long run when the expenditure of retaining a student for one or more years is considered. Each time a student is retained there is an added cost to educate that pupil. As of 2003-04, the total per pupil expenditure for regular education students in Florida was about \$4800. That, then, is the cost of retaining each student. Why not seek an early return on the investment? The FLDOE should consider redirecting funding streams to pay for summer school for early primary grade students, which may help alleviate the need to retain students in third grade. Since Florida already has the READ initiative in place to identify students earlier in their schooling, it appears prudent that the State fund summer school for the most vulnerable early primary grade students in an effort to help students get back on track.

As we can see, waiting until students fail before offering summer school as a support option contradicts research (Borman, Benson & Overman, 2005; Paris et al., 2004; Roderick, Jacob, & Bryk, 2004; Roderick & Nagaoka, 2005); however there is another issue on hand, called summer setback, that may also help support the case for extending the school year for those students most at risk for retention. Researchers have found that students from higher-poverty backgrounds may experience summer setback, or

summer reading loss, when after each summer they begin to trail months behind their peers with greater economic advantages (Cooper, 2004; Heyns, 1978; Kim, 2004; McGill-Franzen & Allington, 2001; Phillips & Chin, 2004).

The “faucet theory” offers one possible explanation for summer setback (Entwisle, Alexander, & Olsen, 2001). The academic faucet is turned “on” during the school year when learning opportunities are more abundant; whereas, in the summer, the academic faucet is turned “off” when fewer learning opportunities exist. After comparing test scores to detect gains and losses between higher- and lower-poverty students, Entwisle, Alexander and Olsen (2004) found seasonal achievement differences when comparing assessment scores over time. For students from higher-poverty backgrounds, they tended to lag behind their advantaged peers during the summer months. In fact, they had a tendency to begin each new school year behind students from lower-poverty backgrounds. Therefore, it appears that turning “off” the academic faucet negatively affects students from higher-poverty backgrounds.

Students who lose bits and pieces of the previous years’ learning experience summer setback (Heyns, 1978). However, just how much learning is lost? For some students, summer setback can amount to about two years of loss over a typical school career (Paris et al., 2004). Increasing the amount of reading practice provided *all year long* may help stave off the effects of summer setback and in turn, may help boost student achievement and even reduce retention rates. For some students, it may come down to limited access to print materials or opportunities for enrichment. During the summer students from high-poverty backgrounds may simply lack access to reading materials (McGill-Franzen & Allington, 2003). Students from higher-poverty backgrounds tend to

have fewer opportunities to engage in reading or other academic activities, such as visiting a local library or museum, during the summer months (Phillips & Chin, 2004).

Why not try to circumvent retention by keeping the faucet turned “on” during the summer months for students most at risk for retention? Waiting until students fail is too late to provide this much needed support. Offering summer school at the earliest point possible may help thwart retention. It seems that if schools are working to identify struggling students, then it would logically follow that these students be provided timely interventions. Because we know that Florida students from higher-poverty backgrounds have an increased likelihood to experience a mandated third-grade retention, it seems as important as ever to try to recoup losses due to summer setback by providing a summer school program. While school day efforts to help identified struggling readers is necessary, it may not be enough for some students. Summer school expands learning time, plus it may help minimize summer setback.

Future Research

Many intriguing issues emerged from this research. Arguably, the most relevant was methodological. For research questions to be answered with a multilevel model, it is important to choose the correct model for the distribution and to address the model assumptions. My study underwent several analyses to determine the correct transformation of the data. Because these data clearly suggest a nonlinear trend, future research needs to consider analyzing a truly nonlinear model in which the model is adjusted. For this study, the data was adjusted to fit the model using the logarithmic transformation plus 1 unit. Further improvements might be made to address the high incidence of zero-values pre-policy.

In Florida, double retention in third grade is a relatively new phenomenon. Research can continue to monitor promotional patterns as student's progress through their schooling. Like Chicago (Allensworth, 2005; Roderick et al., 2000; Roderick & Nagaoka, 2005), research in Florida can study the impact retention has on struggling students learning how they are supported, especially as they reach middle and high school. After promotion to fourth grade, in what ways does the educational system continue to support retainees and double-retainees ensuring equal access to guidance and support throughout their schooling?

The *Florida Principal Survey* provided principals with the opportunity to report upon current practices within their schools. This information serves as a baseline for future research to follow after the state releases the 2004-05 retention rates. We have learned from these analyses about the quantity of the supports available to at risk students. Perhaps we may be able to determine whether we can predict which types of supports work in certain kinds of schools.

Other avenues require more research beyond the scope of the data collected for this study. Considering that principals' retention beliefs did not impact the multilevel model building in this study suggests that further study in the area of beliefs, or underlying assumptions, is needed, especially related to retention policy enactment. One approach could be to develop a construct to determine the underlying perceptions principals have about retention. Insights from such a construct may help glean reasons why believers and nonbelievers retained similar numbers of students before and after policy implementation.

Conclusion

Florida's retention policy has eliminated social promotion in third grade. After reviewing Florida's legislative intent, we know the policy established third grade as the gateway year. The cornerstone of the policy ending social promotion for third-grade students was identifying those students who need a stronger literacy foundation and providing them with extra time and support. However, we must take heed of the numerous supports suggested for schools to implement and consider each student as an individual to determine the best way to support student learning. While most educators and policymakers alike strive to find ways to raise reading achievement and develop stronger readers, we must show prudence in trying anything that has been shown to work in research, as there is a chance that similar results may not be universal. Simply because an intervention has been proven to work with one population of students does not ensure that it will work with *all* students. As should be plainly obvious, all schools, teachers, and students therein are unique and bring different strengths to the forefront. It is important that we remember the individualism that makes each school, teacher, and student unique.

Furthermore, as Elmore (2002) once stated, "low-performing schools, and the people who work in them, don't know what to do. If they did, they would be doing it already" (pp. 33-34). By examining the degree of support in Florida schools, we took stock of the practices of schools, so that future research may assess its benefit.

Ultimately, what matters most is ensuring students, regardless of their principals' retention belief, or the schools' poverty level, receive access to the support they need to be successful. As with many challenges in education, there are no easy answers or "quick fixes" (Allington & Walmsley, 1995); however, we must persist to understand

what makes a difference for each student. Because there are so many variables, it is difficult to isolate what truly makes a difference. It is for this reason that multilevel modeling was selected for my study. It allowed me to focus on testing the best fit of particular variables, as opposed to merely testing hypotheses. Nevertheless, we, researchers and educators alike, must continue to persevere examining how to best support students in the high-stakes policy environment in which we live. I close with a passage from *Laggards in our schools* (Ayers, 1909)

It does not make much difference what we have to do, whether it is a great thing or a little thing, so long as we feel it is possible for us and that we can do it if we try. There are few more hopeless things in the world than to have it borne in upon us that we are driving against a thing that we cannot do. Yet this is the sort of training that we are giving a large part of our children (p. 220).

Ayers was a visionary. We must consider his cautions as we continue studying how to best support those students who struggle with reading.

APPENDIX A
FLORIDA PRINCIPAL SURVEY



UNIVERSITY OF FLORIDA

Florida Principal Survey

Thank you for participating in this survey. The information gathered will help us analyze how principals interpret and implement the third grade retention policy enacted in 2003. Your opinions are important to us.

The information you provide in this survey will remain strictly confidential and will NOT be shared with your school district or the State. We will not refer to specific individuals or school districts by name. School districts will be referred to as small, mid-size, or large. Only members of our research team will have access to your responses. Your participation in this survey is voluntary. A final report will be available next summer.

This survey should take approximately 10 - 15 minutes to complete. If you are have any questions or need your ID number, please e-mail Courtney Zmach (czmach@ufl.edu) for assistance.

Thank you for joining with us in this project.

To help with our record keeping, please answer the following questions:

1. If you have the Identification Number we provided to you earlier, please enter it below. If not, you may leave this blank. _____
2. Your School District _____
3. Your name (last name, first name) _____
4. Your School _____

By completing and mailing this survey, you are stating that you have read the enclosed informed consent letter, and voluntarily agree to participate in this survey.

Directions. Please answer the following questions based on your current interpretation of the third grade retention policy. Some questions provide a space for an "other" answer on a short line. Feel free to use short words or phrases here. Other questions use a rating scale. If at any time during the survey you need to take a break, for one reason or another, the survey is designed such that you should be able to pick up where you left off. For this, you must be on the same computer. Again, your time is most appreciated!

Sources of information. Please respond to the following questions.

5. How useful are the following in providing assistance for implementing the retention policy? (Check only one per row)

	Not useful	Somewhat useful	Mostly useful	Perfectly useful	N/A
Florida DOE					
Local School District					
Support from Assistant Principal					
Support from third grade teachers					
Other Media Source (Internet, News, Radio)					

6. How important is it for you to stay abreast of legislative changes to retention policies?

- Very Important
- Somewhat Important
- Slightly Important
- Not Important

Your experiences. Think about your personal experiences when implementing this policy.

7. Have you felt resistance to the retention policy from the following groups?

	Yes	No
Community Groups (Church, Business, etc.)		
Your Teachers (excluding third grade teachers)		
Third Grade Teachers		
Other School Based Personnel (Asst. Principal, Curriculum Specialist)		
Other School Based Personnel (Teaching Assistants, Paraprofessionals)		
District Staff		
Parents		

8. In your opinion, who should have the "final say" in retention decisions?

- Parent
- Teacher
- Principal
- Superintendent
- Other (please specify) _____

9. Knowing that individual circumstances can exist, what is your response to the following statement? "On the whole, retention in grade can benefit 3rd grade students who score Level 1 on the FCAT"

- Yes, retention is a beneficial option for students who fail the FCAT.
- No, retention is not a beneficial option for students who fail the FCAT.

Good cause exemption portfolios. Think about the good cause exemption portfolios that were decided upon at the end of the 2003-2004 school year.

10. Was there enough time for you to review your students' "good cause" portfolios?

- Yes
- No
- N/A - I did not have ANY good cause portfolios to review.
- N/A - I was not a principal at this school last year.

11. What model did you use for your good cause exemption portfolio?

- A portfolio was developed by the State.
- A portfolio was developed by our school district.
- We use a portfolio developed by another school district.
- We use a portfolio developed by another school in our district.

- We developed our own portfolio at our school.
- Individual teachers decide what to include in the portfolio.
- Not sure

How much is this like you?

12. Please respond to the following questions using the scale provided. Where applicable, use the time-frame provided to help gauge your response.

	1 - Not at all like me	2	3 - Somewhat like me	4	5 - Very much like me
I rely on other people to help me understand the retention policy.					
I believe ability grouping is important to help monitor our retained and re-retained third grade students.					
I am comfortable with the speed in which decisions are made regarding the retention policy (at the state level).					
I am comfortable with the speed in which decisions are made regarding the retention policy (at the district level).					
I am comfortable with the speed in which decisions are made regarding the retention policy (at the school level).					
I consult with the FCRR (FL Center for Reading Research @ FSU) website regularly (at least 2 times per month).					
District officials follow-up with me regularly (at least 1 time per month) to ensure implementation of the retention policy.					
I feel I am part of the decision making process for this policy.					
I consult the FLDOE website regularly (at least 2 times per month) to stay current with new retention laws and regulations.					
I read recommendations from professional journals, such as Reading Teacher, Reading Research Quarterly, Florida Reading Quarterly					

I attend state and national conferences, such as IRA (International Reading Assoc.) or ASCD (Assoc. for Supervision and Curriculum Development).					
--	--	--	--	--	--

Resources for at-risk students.

13. In 2004-2005, how often does your school plan to assess the reading ability of K-3 students?

- Weekly
 Bi-weekly
 Monthly
 Quarterly
 Other (please specify) _____

14. Think about programs or initiatives currently in place at your school designed to help at-risk students.

	Yes	No	Not Sure
Is the Reading Enhancement and Acceleration Development (READ) Initiative available to students in your school?			
Does your school emphasize early intervention programs in K - 2 to prevent retention in 3rd grade more now than in the past?			
Do ALL retained 3rd grade students in your school have a current AIP?			
Do at-risk 3rd graders have access to tutors or mentors trained in reading instruction?			
Is a transitional class (a class designed for the retained and/or re-retained) available for retained students?			

15. Since the retention policy began in 2003, would you say that you are retaining fewer, about the same, or more students in Grades K - 2 to help prepare students for the 3rd grade FCAT?

	Fewer than before policy	Same since policy enacted	More than before policy	Have not compared
Kindergarten				
First				
Second				

16. Since the retention policy began in 2003, would you say that you are retaining fewer, about the same, or more students in Grade 3?

Fewer than before policy	Same since policy enacted	More than before policy	Have not compared

Summer reading camps. Please think back to the summer reading camp held in your district.

17. Who decided the Summer Camp reading curricula for the retained third graders? (Check all that apply.)

Summer School site principal	Team of District Principals	District curriculum leaders	Individual teachers	Not Sure

18. Were you a Summer Reading Camp site principal?

- Yes
 No

Support for the retained.

19. Which, if any, of the following are used as strategies with retained third grade students? (Check all that apply)

- A mentor or tutor with specialized reading training
 Extended school day (such as After School Program, Saturday School, or Extended School Year) as defined in the NCLB legislation
 Ongoing portfolios that meet state portfolio requirements
 None of the above

20. How often does your school plan to review AIP's (Academic Improvement Plan) for retained 3rd graders?

- Weekly
 Bi-Weekly
 Monthly
 Quarterly
 Not Yet Determined
 Other (please specify) _____

21. Are there instructional options that you would prefer to use with retained students, but have not been able to implement for one reason or another? If yes, please share your ideas below.

Third grade teachers. Please respond to the following questions regarding your current third grade instructional staff.

22. What factors would you use to determine whether a 3rd grade teacher is "high-performing"? (Check all that apply)

- District identified "high-performing" teachers.
- Student performance data
- Above-satisfactory on performance appraisals
- Teacher has completed the Reading Endorsement.
- Teacher has completed a Reading Certification program.
- Not Sure
- Other (please specify) _____

23. Think about your current third grade teachers. How many 3rd grade teachers are NOT TENURED (have four years or less teaching experience)? Please enter the number in the space provided.

24. How many third grade teachers does your school currently have? Please enter the number in the space provided.

Funding. Think about the amount of funding your school receives and the adequacy to meet the financial needs to support your retained students.

25. How are decisions made in your school on how to fund new initiatives and programs for the retained? Who is involved in the decision making process?

26. Who do you believe should be involved in the funding decision making process?

Information. Please write the best response to the following questions about your educational background and the number of students "re-retained".

27. How many years have you been in education? Please include all teaching and administrative experiences.

28. How many years have you been a principal?
(Including 2004-2005 school year)

29. How many years have you been principal at your current school?
(Including 2004-2005 school year)

30. How many students were re-retained (retained for the second time) in third grade?

Please enter a numeral in the space provided. _____
If you are unsure, please leave this question blank.

Thank you! Thank you again for completing this survey. We will be contacting some of you for telephone interviews. Wishing you, your staff and your students a wonderful school year!

31. OPTIONAL: Please feel free to add any additional comments or feedback about the retention policy that you feel are important.

APPENDIX B REVIEW OF SURVEY PILOT STUDY

A pilot study of the closed- and open-ended questions established face, content and logical validity, thus ensuring that the content asked accurately reflected the issues that principals believe to be important, and that the questions and wording were clear to principals (Litwin, 1997). Face validity was established by sharing drafts of the survey with colleagues. Their consultation provided the researcher multiple perspectives and unexpected advice. After carefully considering their input, modifications resulted in reducing the number of items asked on the survey, prompting reflection about rating scales and reducing the number of open-ended questions found in early survey drafts. Content and logical validity were established with practicing principals. Two “think-aloud” interviews were conducted with principals while they took the survey (Desimone & Le Floch, 2004); three volunteers completed the survey and provided written feedback after taking the survey.

After analyzing the pilot data, the sample responses were considered in conjunction with the research questions. It was determined that the survey required modification to eliminate questions related to the topic of retention, but not directly related to the research questions of this study. Some questions were reworded using less complicated language. After considering the multitude of reactions from colleagues and pilot participants, including “think-aloud” interviewees, changes were made to enhance the face, logical and content validity of the survey instrument. This process increased the chance that when principals took the finalized version of the survey their responses

would be more closely aligned with their intended meaning. Overall, pilot respondents communicated positive attitudes toward the survey.

APPENDIX C
DATA SOURCES

Table C-1. Sources of data

School & District Level Data	Source	School Year
Third-grade enrollment by school and district	Florida Survey 2 Florida Department of Education (FLDOE)	1999-00, 2000-01, 2001-02, 2002-03, 2003-04
Third-grade retention/ non-promotion by school and district	Florida Survey 5 (FLDOE)	1999-00, 2000-01, 2001-02, 2002-03, 2003-04
Free/reduced lunch by school	FLDOE	1999-00, 2000-01, 2001-02, 2002-03, 2003-04
Enrollment by school and district	FLDOE Schoolresults.org ^a	1999-00, 2000-01, 2001-02, 2003-04 2002-03
School Type/enrollment	NCES/Common Core of Data	2002-03 ^b
Title I recipients	FLDOE website	2004-05 ^b
Reading First recipients	Just Read FL! office	2004-05 ^b

^a Information related to Florida posted at www.schoolresults.org were provided by the FLDOE.

^b Most current available information from these sources was used in my study.

APPENDIX D
SAMPLE FROM STATISTICAL SYNTAX

```
if time=1 twt1 = -1.1  
if time=2 twt1 = 0.  
if time=3 twt1 = 1.  
if time=4 twt1 = 0.  
if time=5 twt1 = 0.
```

```
if time=1 twt2 = 0.  
if time=2 twt2 = 0.  
if time=3 twt2 = 0.  
if time=4 twt2 = -.5.  
if time=5 twt2 = +.5.
```

execute.

```
title "MODEL A: RA / Unconditional Means Model (Degenerate)".  
MIXED logper_ret_plus1  
  /CRITERIA = CIN(95) MXITER(100) MXSTEP(5) SCORING(1)  
SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE)  
  PCONVERGE(0.000001, ABSOLUTE)  
  /FIXED =  
  /METHOD = REML  
  /PRINT = G SOLUTION TESTCOV  
  /RANDOM INTERCEPT | SUBJECT(New_ID) COVTYPE(VC)  
  /REGWGT = lgenrol .
```

```
title "MODEL B: Unconditional Growth Model (UGM)".  
MIXED logper_ret_plus1 WITH PrePost_Policy  
  /CRITERIA = CIN(95) MXITER(100) MXSTEP(5) SCORING(1)  
SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE)  
  PCONVERGE(0.000001, ABSOLUTE)  
  /FIXED = PrePost_Policy | SSTYPE(3)  
  /METHOD = REML  
  /PRINT = G SOLUTION TESTCOV  
  /RANDOM INTERCEPT | SUBJECT(New_ID) COVTYPE(VC)  
  /REGWGT = lgenrol .
```

¹ Syntax created using SPSS 12.0 for Windows.

APPENDIX E
RESULTS TABLE FOR ALL MODELS

Table E-1. Results of fitting a taxonomy of multilevel models for change to the log percent retained data, Models A through F2 (n = 102)

	A	B	C	D	E	F1	F2
Intercept	1.66***	1.17***	.463***	.469***	.465***	.496***	.521**
Pre-Post Policy		1.14***	1.13***	1.10***	1.11***	1.11***	1.11***
sesmean			.014***	.014***	.014***	1.37***	1.397***
sesdiff				.022**	.020**	1.95**	1.952**
tw1					.034	-0.226	.033
tw2					-0.226**	-0.047**	-0.226**
Q9						.035	
suppsys3grp							-0.029
Residual	3.72***	1.81***	1.82***	1.78***	1.76***	1.75***	1.75***
Intercept	.064~	.155***	.056**	.054**	.055**	.056**	.056**
Between	.064	.156	.056	.054	.055	.056	.056
Within	3.72	1.81	1.82	1.78	1.76	1.76	1.76
TOTAL	3.79	1.96	1.87	1.83	1.81	1.81	1.81
Proportionate reduction in error							
PRE-Residual		.515	-0.005	.015	.029	.029	.029
PRE-Intercept		-1.430	.637	.656	.648	.642	.641
Relative to Model-		A	B	B	B	B	B
Intra-class correlation coefficient							
ICC-Between	.017	.079	.030	.029	.030	.031	.031
ICC-Within	.983	.921	.970	.971	.970	.969	.969
Goodness of Fit							
AIC	1279.43	1012.81	966.34	962.86	962.89	947.36	948.03
AIC difference		266.62	313.09	316.57	316.54	332.07	331.40
Fit improvement (%)		20.8	24.5	24.7	24.7	25.95	25.90
Relative to Model -		A	A	A	A	A	A

~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

LIST OF REFERENCES

- Alexander, K. L., Entwisle, D. R., & Olsen, L. S. (2004). Schools, achievement, and inequality: A seasonal perspective. In G. D. Borman & M. Boulay (Eds.), *Summer learning: Research, policies and programs* (pp. 25-51). Mahwah, NJ: Erlbaum.
- Allen, J., Cary, M., & Delgado, L. (1995). *Exploring blue highways: Literacy reform, school change, and the creation of learning communities*. New York: Teachers College Press.
- Allensworth, E.M. (2005). Dropout rates after high-stakes testing in elementary school: A study of the contradictory effects of Chicago's efforts to end social promotion. *Educational Evaluation and Policy Analysis*, 27(4), 341-364.
- Allington, R. L. (2001). Does state and federal policymaking matter? In T. Loveless (Ed.), *The great curriculum debate: How should we teach reading and math?* (pp. 268 - 298). Washington, D.C.: Brookings Institution Press.
- Allington, R. L. (2002a). Research on reading/learning interventions. In A. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 261-290). Newark, DE: International Reading Association.
- Allington, R. L. (Ed.). (2002b). *Big brother and the national reading curriculum: How ideology trumped evidence*. Portsmouth, NH: Heinemann.
- Allington, R. L., & McGill-Franzen, A. (1992). Unintended effects of educational reform in New York State. *Educational Policy*, 6(4), 397-414.
- Allington, R. L., & McGill-Franzen, A. (2000). Looking back, looking forward: A conversation about teaching reading in the 21st century. *Reading Research Quarterly*, 35(1), 136-153.
- Allington, R. L., McGill-Franzen, A., & Schick, R. (1997). How administrators understand learning difficulties. *Remedial and Special Education*, 18(4), 223-232.
- Allington, R. L., & Walmsley, S. A. (Eds.). (1995). *No quick fix: Rethinking literacy programs in America's elementary schools*. New York: Teachers College Press.
- American Educational Research Association. (2000). *AERA position statement concerning high-stakes testing in PreK-12 education*. Retrieved November 12, 2002, from <http://www.aera.net/policyandprograms/?id=378>

- Ames, L. B. (1986). Ready or not. *American Educator*, 10(2), 30-33.
- Ayers, L. P. (1909). *Laggards in our schools*. New York: Russell Sage Foundation.
- Ball, D. L. (1990). Reflections and deflections of policy: The case of Carol Turner. *Educational Evaluation and Policy Analysis*, 12(3), 247-259.
- Balow, I. H., & Schwager, M. (1992). Retention in grade: A failed procedure. In R. Reitz (Ed.), *Retention in grades: Looking for Alternatives* (pp. 9-49). Bloomington, IN: Phi Delta Kappa Center for Evaluation, Development and Research.
- 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.
- Baumann, J. F., & Bason, J. J. (2004). Survey research. In N. K. Duke & M. Mallette (Eds.), *Literacy methodologies* (pp. 287-307). NY: Guilford.
- Borman, G.D., Benson, J., & Overman, L.T. (2005). Families, schools, and summer learning. *Elementary School Journal*, 106(2), 131-150.
- Bracey, G. W. (2003). Poverty issues get short shrift. In *On the death of childhood and the destruction of public schools* (pp. 45- 48). Portsmouth, NH: Heinemann.
- Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assess change. *Psychological Bulletin*, 101, 147-158.
- Bush, G. W. (2001). *No Child Left Behind Act of 2001*. Retrieved February 5, 2003, from United States Department of Education Web site:
<http://www.ed.gov/nclb/overview/intro/execsumm.html>
- Byrnes, D. A. (1989). Attitudes of students, parents, and educators toward repeating a grade. In L. A. Shepard & M. L. Smith (Eds.), *Flunking grades: Research and policies on retention* (pp. 132-150). London, England: Falmer Press.
- Camilli, G., Wolfe, P., & Smith, M. L. (in press). Meta-analysis and reading policy: Perspectives on teaching children to read. *Elementary School Journal*.
- Center for Civic Innovation at the Manhattan Institute (2004). *An evaluation of Florida's program to end social promotion*. (Education Working Paper No. 7). Retrieved January 2, 2005, from http://www.manhattan-institute.org/html/ewp_07.htm#01
- Chase, J. A. (1968). A study of the impact of grade retention on primary school children. *Journal of Psychology*, 70, 169-177.

- Clinton, W. J. (1999, January 20). 1999 State of the Union Address. *Washington Post*, A12. Retrieved September 13, 2004, from <http://www.washingtonpost.com/wp-srv/politics/special/states/docs/sou99.htm>
- Cooper, H. (2004). Is the school calendar dated? Education, economics, and the politics of time. In G. D. Borman & M. Boulay (Eds.), *Summer learning: Research, policies and programs* (pp. 3-23). Mahwah, NJ: Erlbaum.
- Cooper, H., & Hedges, L.V. (Eds.). (1994). *The handbook of research synthesis*. New York: Russell Sage Foundation.
- Darling-Hammond, L. (1990). Instructional policy into practice: "The power of the bottom over the top". *Educational Evaluation and Policy Analysis*, 12(3), 339-347.
- Desimone, L. M., & Le Floch, K. C. (2004). Are we asking the right questions? Using cognitive interviews to improve surveys in education research. *Educational Evaluation and Policy Analysis*, 26(1), 1-22.
- Dillman, D. A. (2000). *Mail and internet surveys: The tailored design method* (2nd ed.). New York: Wiley.
- Dorn, S. (2004). *Alternatives for Florida's assessment and accountability system—Policy brief*. Tempe, AZ: Education Policy Studies Laboratory, Arizona State University.
- Elmore, R. (2002). Unwarranted intrusion. *Education Next*, 2, 31-35. Retrieved September 23, 2003, from <http://www.educationnext.org/20021/30.html>
- Entwisle, D. R., Alexander, K. L., & Olsen, L. S. (2001). Keep the faucet flowing: Summer learning and home environment. *American Educator*, 25(3), 10-15, 47.
- Erpenbach, W. J., Forte-Fast, E., & Potts, A. (2003). *Statewide educational accountability under NCLB: Central issues arising from an examination of state accountability workbooks and U.S. Department of Education reviews under the No Child Left Behind Act of 2001*. Washington, DC: Council of Chief State School Officers.
- Figlio, D.N. & Getzler, L.S. (2002). Accountability, ability, and disability: Gaming the system (NBER Working Papers No. 9307). Retrieved on January 10, 2006, from <http://www.nber.org/papers/w9307.pdf>
- Flinders, D. J., & Thornton, S. J. (Eds.). (1997). *The curriculum studies reader*. New York: Routledge.
- Florida Department of Education. (2000a). *FCAT achievement levels*. Retrieved March 20, 2003, from <http://www.firn.edu/doe/sas/fcat/pdf/fcatlevl.pdf>

- Florida Department of Education. (2000b). *Technical assistance paper: Standards for implementation of student assessment programs in Florida school districts*. Retrieved June 3, 2005, from <http://www.firn.edu/doe/sas/pdf/ethics.pdf>
- Florida Department of Education. (2003a). *Q&A technical assistance paper DPS #03-024: Student progression requirements for 2002-2003*. Retrieved March 1, 2003, from <http://info.fldoe.org/docushare/dsweb/Get/Document-435/TAP17702b1.PDF>
- Florida Department of Education. (2003b). *Total number of planned purchases for all districts, reading, grade K-5*. Retrieved February 17, 2004, from Florida Instructional Materials Reporting Web site: <http://www.fimweb.org/reports/default.asp>
- Florida Department of Education. (2003c). *Understanding FCAT reports 2003*. Retrieved February 20, 2004, from <http://www.firn.edu/doe/sas/fcat/fcinfo.htm>
- Florida Department of Education. (2004a). *2004 legislation CS/SB 364: Reading deficiencies chart*. Retrieved June 9, 2004, from <http://info.fldoe.org/dscgi/ds.py/Get/File-2130/04-132.pdf>
- Florida Department of Education. (2004b). *Press release: Governor Bush and Education Commissioner Winn announce plan to assist schools affected by hurricanes*. Retrieved September 29, 2004, from http://www.fldoe.org/news/2004/2004_09_21-2.asp
- Florida Department of Education. (2004c). *Press release: Governor Bush, Lt. Governor Jennings and Commissioner Horne announce preliminary FCAT results for 3rd and 12th grade*. Retrieved April 19, 2004, from http://www.fldoe.org/news/2004/2004_04_19.asp
- Florida Department of Education. (2004d). *Technical assistance paper #12315: Good cause exemptions (FY2004-17)*. Retrieved June 9, 2005, from <http://www.firn.edu/doe/bin00014/pdf/y2004-17.pdf>
- Florida Department of Education. (2005a). *Florida school indicators report*. Retrieved May 30, 2005, from <http://data.fldoe.org/fsir/default.cfm>
- Florida Department of Education. (2005b, May 9). *Governor Bush announces Reading First grant recipients*. Retrieved June 20, 2005, from http://www.fldoe.org/news/2005/2005_05_09.asp
- Florida Department of Education. (2005c). *Welcome page: Third-grade reading promotion and retention*. Retrieved June 8, 2005, from <http://www.firn.edu/doe/commhome/progress/proghome.htm>

- Florida Department of Education. (2005d). *Press release*. Retrieved May 25, 2005, from <http://fcet.fldoe.org/mediapacket/pdf/05gr3gr12pressrelease.pdf>
- Florida Department of Education (2005e). *Non-promotions in Florida's public schools, 2003-04*. Retrieved September 12, 2005, from <http://www.firn.edu/doe/eias/eiaspubs/pdf/nonpromotions.pdf>
- Florida Statutes. (2002). *Florida K - 20 Education Code: Assessment and accountability: TITLE XLVIII, Chapter 1008.25*, Retrieved March 10, 2005, from The Florida Senate Web site: http://www.flsenate.gov/statutes/index.cfm?mode=View%20Statutes&SubMenu=1&App_mode=Display_Statute&Search_String=&URL=CH1008/Sec25.HTM
- Fullan, M. (2001). *The new meaning of educational change*. New York: Teachers College Press.
- Fullan, M. (2002). The change leader. *Educational Leadership*, 59(8), 16-21.
- Good, R.H., & Kaminski, R.A. (Eds.) (2002). *Dynamic Indicators of Basic Early Literacy Skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.
- Goodlad, J. I. (1952). Research and theory regarding promotion and nonpromotion. *Elementary School Journal*, 53(3), 150-155.
- Graue, M. E., Kroeger, J., & Brown, C. (2003). The gift of time: Enactments of developmental thought in early childhood practice. *Early Childhood Research and Practice*, 5(1). Retrieved September 25, 2004, from <http://ecrp.uiuc.edu/v5n1/graue.html>
- Grissom, J. B., & Shepard, L. A. (1989). Repeating and dropping out of school. In L. A. Shepard & M. L. Smith (Eds.), *Flunking grades: Research and policies on retention* (pp. 34-63). London, England: Falmer Press.
- Guthrie, J. T. (2002). Preparing students for high-stakes test taking in reading. In A. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (pp. 370-391). Newark, DE: International Reading Association.
- Hall, W. F., & Demarest, R. (1958). Effects on achievement scores of a change in promotional policy. *Elementary School Journal*, 58(4), 204-207.
- Hart, G. K. (1996). A policymaker's response. *Journal of Literacy Research*, 28, 596-601.
- Heyns, B. (1978). *Summer learning and the effects of schooling*. New York: Academic Press.

- Hiebert, E. H. (2002). Standards, assessments and text difficulty. In A. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (pp. 337-369). Newark, DE: International Reading Association.
- Holmes, C. T. (1989). Grade-level retention effects: A meta-analysis of research studies. In L. A. Shepard & M. L. Smith (Eds.), *Flunking grades: Research and policies on retention* (pp. 16-33). London, England: The Falmer Press.
- Holmes, C. T., & Matthews, K. M. (1984). The effects of nonpromotion on elementary and junior high school pupils: A meta-analysis. *Review of Educational Research*, 54(2), 225-236.
- Hox, J. J. (2000). Multilevel analyses of grouped and longitudinal data. In T. D. Little, K.U. Schabel & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 15-32). Mahwah, NJ: Erlbaum.
- Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Erlbaum.
- International Reading Association. (1999). *Position statement: High-stakes assessments in reading*. Retrieved September 20, 2002, from http://www.reading.org/resources/issues/positions_high_stakes.html
- Isaac, S., & Michaels, W. B. (1995). Instrumentation and measurement. In *Handbook in research and evaluation* (pp. 105-158). San Diego, CA: Educational and Industrial Testing Services.
- Jackson, G. B. (1975). The research evidence on the effects of grade retention. *Review of Educational Research*, 45(4), 613-635.
- 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., 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.
- Kim, J. (2004). Summer reading and the ethnic achievement gap. *Journal of Education of Students Placed At-Risk*, 9(2), 169-189.
- Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.
- Light, R.J. & Pillemer, D.B. (1984). *Summing up: The science of reviewing research*. Cambridge, MA: Harvard University Press.

- Lindblom, C. E., & Woodhouse, E. J. (1993). *The policy-making process* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.
- Linn, R. L. (2004). Accountability models. In S. H. Fuhrman & R. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 73-95). New York: Teachers College Press.
- Litwin, M. S. (1997). *How to measure survey reliability and validity*. Thousand Oaks, CA: Sage.
- Lobdell, L. O. (1954). Results of a nonpromotion policy in one school district. *Elementary School Journal*, 54(6), 333-337.
- Massell, D. (2001). The theory and practice of using data to build capacity: State and local strategies and their effects. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states* (pp. 148-169). Chicago, IL: University of Chicago Press.
- McElroy, E. (2001). *Biology 283 lab: Brain size*. Retrieved June 28, 2005, from Western Kentucky University, Department of Evolutionary Biology, Genetics, and Biostatistics Web site:
<http://bioweb.wku.edu/faculty/McElroy/BIOL283/283-5.htm>
- McGill-Franzen, A. (1987). Failure to learn to read: Formulating a policy problem. *Reading Research Quarterly*, 22(4), 475-490.
- McGill-Franzen, A. (2000). Policy and instruction: What is the relationship? In M. L. Kamil, P. B. Mosenthal, P. D. Pearson & R. Barr (Eds.), *Handbook of reading research* (Vol. III, pp. 889-908). Mahwah, NJ: Erlbaum.
- McGill-Franzen, A., & Allington, R. L. (1993). Flunk'em or get them classified: The contamination of primary grade accountability data. *Educational Researcher*, 22(1), 19-22.
- McGill-Franzen, A., & Allington, R. L. (2001). Summer reading. (importance of giving children access to books throughout summer months). *Reading Today*, 18(6), 10.
- McGill-Franzen, A., Zmach, C. C., Solic, K., & Love Zeig, J. (in press). The confluence of two policy mandates: Core reading programs and third-grade retention. *Elementary School Journal*.
- McLaughlin, M. W. (1987). Learning from experience: Lessons from policy implementation. *Educational Evaluation and Policy Analysis*, 9, 171-178.

- Moskowitz, D. S., & Hershberger, S. L. (Eds.). (2002). *Modeling intraindividual variability with repeated measures data: Methods and applications*. Mahwah, NJ: Erlbaum.
- Musella, D. F. (1989). Problems in policy implementation. In M. Holmes, K. A. Leithwood & D. F. Musella (Eds.), *Educational policy for effective schools* (pp. 93-111). Toronto, Ontario, Canada: Ontario Institute for Studies in Education Press.
- Nagaoka, J., & Roderick, M. (2004). *Ending social promotion: The effects of retention*. Chicago, IL: Consortium on Chicago School Research.
- National Center for Education Statistics. (2002). *School locale codes 1987-2000 (NCES 2002-02)*. Retrieved September 3, 2004, from <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=200202>
- National Center for Education Statistics. (2004). *Common core of data: Build a table*. Retrieved October 27, 2004, from <http://nces.ed.gov/ccd/bat/>
- National Commission of Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Retrieved September 20, 2004, from <http://www.ed.gov/pubs/NatAtRisk/index.html>
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.
- National Research Council. (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Academy of Sciences.
- Oakes, J. (1999). *Promotion or retention: Which one is social?* Retrieved June 14, 2004, from Harvard Education Letter Web site: <http://www.edletter.org/past/issues/1999-jf/promotion.shtml>
- Onwuegbuzie, A.J., & Leech, N.L. (2004). Enhancing the interpretation of "significant" findings: The role of mixed methods research. *The Qualitative Report*, 9(4), 770-792. Retrieved March 1, 2005 from <http://www.nova.edu/ssss/QR/QR9-4/onwuegbuzie.pdf>
- Paris, S. G., Pearson, P. D., Cervetti, G., Carpenter, R., Paris, A. H., DeGroot, J., et al. (2004). Assessing the effectiveness of summer reading programs. In G. D. Borman & M. Boulay (Eds.), *Summer learning: Research, policies and programs* (pp. 121-161). Mahwah, NJ: Erlbaum.

- Phillips, M., & Chin, T. (2004). How families, children and teachers contribute to summer learning and loss. In G. D. Borman & M. Boulay (Eds.), *Summer learning: Research, policies and programs* (pp. 255-278). Mahwah, NJ: Erlbaum.
- Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.
- Reitz, R. (Ed.). (1992). *Retention in grade: Looking for alternatives*. Bloomington, IN: Phi Delta Kappa Center for Evaluation, Development and Research.
- Roderick, M. (1994). Grade retention and school dropout: Investigating the association. *American Educational Research Journal*, 31(4), 729-759.
- Roderick, M., Bryk, A. S., Jacob, B. A., Easton, J. Q., & Allensworth, E. (1999). *Ending social promotion: Results from the first two years*. Chicago, IL: Consortium on Chicago School Research.
- Roderick, M., & Engel, M. (2001). The grasshopper and the ant: Motivational responses of low-achieving students to high-stakes testing. *Educational Evaluation and Policy Analysis*, 23(3), 197-227.
- Roderick, M., Jacob, B. A., & Bryk, A. S. (2004). Summer in the city: Achievement gains in Chicago's summer bridge program. In G. D. Borman & M. Boulay (Eds.), *Summer learning: Research, policies and programs* (pp. 73-102). Mahwah, NJ: Erlbaum.
- 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.
- Roderick, M., Nagaoka, J., Bacon, J., & Easton, J. Q. (2000). *Update ending social promotion: Passing, retention, and achievement trends among promoted and retained students*. Chicago, IL: Consortium on Chicago School Research.
- Ruhl, K. L. (1984). *A primer on retention: History, research and a checklist for making the decision*. Gainesville, FL: Multidisciplinary Diagnostic and Training Program at the University of Florida.
- Scott, B. A., & Ames, L. B. (1969). Improved academic, personal, and social adjustments in selected primary-school repeaters. *Elementary School Journal*, 69, 431-439.
- Shepard, L. A. (1989). A review of research on kindergarten retention. In L. A. Shepard & M. L. Smith (Eds.), *Flunking grades: Research and policies on retention* (pp. 64-78). London, England: Falmer Press.

- Shepard, L. A., & Smith, M. L. (Eds.). (1989). *Flunking grades: Research and policies on retention*. London, England: Falmer Press.
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.
- Spillane, J. P. (1998). A cognitive perspective on the role of the local educational agency in implementing instructional policy: Accounting for local variability. *Educational Administration Quarterly*, 34(1), 31-57.
- Tomchin, E. M., & Impara, J. C. (1992). Unraveling teachers' beliefs about grade retention. *American Educational Research Journal*, 29(1), 199-223.
- Torgesen, J. (2004). *Information relevant to the challenge of providing reading instruction that will accelerate the development of reading skills in students who are retained in 3rd grade as a result of performing at Level 1 on the FCAT*. Unpublished manuscript, Florida Center for Reading Research.
- Tremor, M., & Butler, M. J. (2004). *Technical assistance paper related to third grade student progression (FY2005-03)*. Retrieved April 5, 2005, from Florida Department of Education Web site:
<http://info.fldoe.org/dscgi/ds.py/Get/File-2343/TAP.pdf>
- Tye, B. B. (2000). *Hard truths: Uncovering the deep structure of schooling*. New York: Teachers College Press.
- U.S. Department of Education. (1999). *Taking responsibility for ending social promotion: A guide for educators and state and local leaders*. Retrieved September 1, 2004, from <http://www.ed.gov/pubs/socialpromotion/index.html>
- U.S. Department of Education. (2002). *No Child Left Behind of 2001*. Retrieved February 5, 2003, from <http://www.ed.gov/nclb/landing.jhtml?src=pb#>
- U.S. Department of Education. (2004a). *Press release: Paige issues statement on report on social promotion*. Retrieved May 29, 2005, from
<http://www.ed.gov/news/pressreleases/2004/12/12082004.html>
- U.S. Department of Education. (2004b). *Press release: Secretary Paige issues new policy for calculating participation rates under No Child Left Behind*. Retrieved January 29, 2006, from <http://www.ed.gov/news/pressreleases/2004/03/03292004.html>
- U.S. Department of Education. (n.d.). *Title I, part A program*. Retrieved June 19, 2005, from <http://www.ed.gov/programs/titleiparta/index.html>

- Valencia, S. W., & Wixson, K. K. (2004). Literacy policy and policy research that make a difference. In R. B. Ruddel & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 69-92). Newark, DE: International Reading Association.
- Warford, J. (2004a). *Memorandum: 2004 required instruction and school closures due to hurricanes (DPS #05-041)*. Retrieved September 30, 2004, from Florida Department of Education Web site:
<http://info.fldoe.org/dscgi/ds.py/Get/File-2455/05-041.pdf>
- Warford, J. (2004b). *Memorandum: Reporting requirements for third grade student progression (DPS # 04-098)*. Retrieved May 20, 2004, from Florida Department of Education Web site:
<http://info.fldoe.org/dscgi/ds.py/Get/File-1978/DPS04-098.pdf>
- Warford, J., & Openshaw, M. L. (2004a). *Memorandum: 2004 legislative changes regarding students with reading deficiencies and student progression (DPS # 04-125)*. Retrieved May 21, 2004, from Florida Department of Education Web site: <http://info.fldoe.org/dscgi/ds.py/Get/File-2073/DPS04-125.pdf>
- Warford, J., & Openshaw, M. L. (2004b). *Memorandum: 2004 legislative requirements regarding students with reading deficiencies and student progression—CS/SB 364 (DPS # 04-132)*. Retrieved June 9, 2004, from Florida Department of Education Web site: <http://info.fldoe.org/dscgi/ds.py/Get/File-2129/04-132.pdf>
- Warford, J., & Openshaw, M. L. (2004c). *Memorandum: Technical assistance paper for implementation of the READ initiative and intensive acceleration class—CS/SB 364 (DPS #04-140)*. Retrieved June 23, 2004, from Florida Department of Education Web site:
<http://info.fldoe.org/dscgi/ds.py/Get/File-2156/DPS04-140.pdf>
- Warford, J., & Openshaw, M. L. (2005). *Memorandum: Third progression technical assistance document (K-12: 2005-60)*. Retrieved May 26, 2005, from Florida Department of Education Web site:
http://info.fldoe.org/dscgi/ds.py/Get/File-2994/k12_05-60.pdf
- Wothke, W. (2000). Longitudinal and multigroup modeling with missing data. In T. D. Little, K. Schabel & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples*. Mahwah, NJ: Erlbaum.

BIOGRAPHICAL SKETCH

Courtney Caldwell Zmach received the University of Florida's highest graduate student award, the Alumni Graduate Fellowship, before beginning her work toward a Ph.D. in curriculum and instruction, specializing in reading education at the University of Florida, Gainesville. She earned her Bachelor of Arts degree in social sciences from St. Bonaventure University (St. Bonaventure, New York) where she was inducted into two international honor societies: Alpha Kappa Delta and Pi Gamma Mu. Courtney also completed her initial teacher preparation training at St. Bonaventure. She earned her Master of Education degree in curriculum, teaching and learning at the University of Toronto/Ontario Institute for Studies in Education (OISE).

Courtney has held elementary education teaching certifications in the Province of Ontario, Canada, as well as the states of Pennsylvania and Florida. Before beginning her doctoral program, she taught for several years in Ontario and Florida. Courtney has also worked as a consultant, across Florida, conducting professional-development sessions for teachers and school administrators. At the University of Florida, she worked as a graduate research assistant and taught reading courses at the undergraduate and graduate levels. Currently, Courtney works as a research analyst, specializing in reading research, at the American Institutes for Research (AIR) in Washington, D.C.