9-29-2005

# Third Grade Retention And The Florida Comprehensive Assessment Test: An Exploratory Study

Heather A. Powell
*University of South Florida*

# Third Grade Retention And The Florida Comprehensive Assessment Test:

# An Exploratory Study

by

Heather A. Powell

A thesis submitted in partial fulfillment
Of the requirements for the degree of
Education Specialist
Department of Psychological and Social Foundations
College of Education
University of South Florida

Major Professor: Michael Curtis, Ph.D.
George Batsche, Ph.D.
Constance Hines, Ph.D.

Date of Approval:
September 29, 2005

Keywords: reading, achievement, remediation, progression, policy

Table of Contents

List of Tables

THIRD GRADE RETENTION AND THE FLORIDA COMPREHENSIVE ASSESSMENT TEST: AN EXPLORATORY STUDY

Heather A. Powell

## ABSTRACT

Literacy is a growing national concern, resulting in federal legislation (e.g., No Child Left Behind Act) instituting higher accountability for states and schools with regard to reading instruction and remediation. As a result, Florida's statewide measure of achievement, the Florida Comprehensive Assessment Test (FCAT) is now tied to retention decisions in third grade as part of the new pupil progression plan in the state. In its first year of implementation (2003), 23% of third-grade students failed the FCAT and over 28,000 were retained.

Research has consistently shown retention to be a negative experience for children; even when academic gains are made, within two to three years, their achievement is equal to or lower than that of both same-grade and same-age regularly promoted students. However, these findings cannot be generalized to the current student progression plan in Florida, which mandates specific remediation activities during the retention year. Therefore, holding negative beliefs about grade retention in Florida is premature as no research exists to date evaluating the outcomes of the plan.

The present study examined the student progression plan in Florida at it relates to high-stakes testing and mandated third-grade retention. More specifically, this study

examined the relationship between the effects of retention and various student demographic characteristics (e.g., gender, SES, race/ethnicity), as well as the future performance on the FCAT-Reading of low-performing students who were promoted through good cause exemptions.

Descriptive analyses revealed that of 20,617 third-grade retainees, 38% again scored at Level 1 in 2004. In addition, future success of retainees was significantly associated with gender, race, and SES. With regard to students who were promoted due to a good cause exemption, findings indicated that a higher percentage of those who demonstrated reading proficiency through an alternative procedure (65%) achieved success in fourth grade compared to those who did not demonstrate proficiency (23%).

This study contributes to the literature by examining the outcomes of grade retention within a context of high-stakes testing and mandated remediation activities. Implications for future research include controlling for the quality of interventions and identifying strategies that target specific populations of at-risk students.

# CHAPTER ONE

## Introduction

Literacy is a growing national concern, justified by statistics reporting that only 32% of our nation's fourth grade children read proficiently (United States Department of Education, 2002). In addition, there is a growing achievement gap, such that the performance of top achievers has increased over time while that of the poorest achievers has declined (United States Department of Education, 2002). These developments have been alarming, provoking new federal legislation (e.g., No Child Left Behind Act), the effects of which have proven to be far-reaching (United States Department of Education, 2003a).

Since the No Child Left Behind Act was enacted on January 8, 2002, initiatives have been undertaken across the country to restructure many aspects of the American educational system (NCLB, 2002). The law was designed to address the country's reading crisis; its overriding goal is for every child to be reading at a proficient level by the 2013-2014 school year (United States Department of Education, 2003b). To this end, the law includes strict guidelines regarding school and state accountability, including procedures to measure student progress. The law mandates demonstration of Adequate Yearly Progress (AYP) in student performance by every state; thus, each state is required to set annual goals for its schools and propose a way to systematically measure progress toward those goals (United States Department of Education, 2003c). In order to address

the widening achievement gap, states are no longer allowed to aggregate achievement data across economic background, race/ethnicity, English proficiency or disability status; schools now must disaggregate performance data and demonstrate that all of their students are learning and that their disadvantaged or at-risk students are not being "left behind".

Although the Act has been criticized in some quarters for the failure to financially support its mandates, funds have been allocated to support some specific programs. For example, funds are available to assist states in the implementation of empirically-based reading programs. States apply for the funds through the established "Reading First" program and funding is granted based on the number of children aged 5-17 years who are considered low-income in each state (United States Department of Education, 2003d).

The acquisition of Reading First funds is a large part of Florida's response to the No Child Left Behind Act. The state has established goals which state that 31% of its students are to be reading proficiently by the 2003-04 school year, 48% by 2006-07, 65% by 2009-10, 82% by 2012-13, and 100% by 2013-14 (Florida Department of Education, 2003a). With Reading First funds ($52 million annually, for six years), schools around the state will be instructing children using research-based reading programs (Florida Department of Education, 2003b).

In addition to preventative programs such as Reading First, the State of Florida also enacted a new retention policy, effective January 7, 2003 (The Florida Senate, 2003). Designed to align Florida policies with federal law, this new policy places heavy emphasis on remediation of reading deficiencies and is strongly tied to the standardized test used to quantify student performance, the Florida Comprehensive Assessment Test

(FCAT). The reading portion of the FCAT is administered every year in grades 3-10 and purports to measure students' progress towards attainment of predefined academic benchmarks in reading according to five levels, defined according to the following scaled scores: 1) 100-258, 2) 259-283, 3) 284-331, 4) 332-393, and 5) 394-500 (Florida Department of Education, 2003c). With regards to reading, it is imperative that students master the basic skills at the lower grade levels (grades K-3), as in fourth grade and beyond students are expected to utilize reading skills in all academic areas. Without these basic skills, the chances of student success are severely limited, since the approach to teaching reading shifts at this critical juncture (Just Read, Florida, 2003). The state contends that third grade students who are achieving at Level 1 have not mastered the most basic reading skills, indicating that they are not ready to move on to the more challenging material of the fourth grade; consequently, students who score at Level 1 on the FCAT and are not eligible for promotion according to predefined "good cause exemptions" are to be retained in third grade (The Florida Senate, 2003). Though perceived by some as pejorative, retention in this context is viewed as an effort towards the remediation of current academic difficulties and the prevention of additional difficulties in the students' future academic careers.

The 2003 FCAT was administered several months after the new retention policy went into effect, resulting in a total of 28,028 (14.6%) third grade students being retained in grade for the 2003-2004 school year (Florida Department of Education, 2003d). The tremendous numbers of students being affected by the policy combined with longstanding assumptions that grade retention is overwhelmingly detrimental to students'

academic progress call for a careful re-examination of the literature regarding grade retention, as well as empirical examination of the impact of this new policy.

The practice of grade retention is not a new phenomenon; in fact, retention literature spans several decades. Despite prevailing negative beliefs, some studies (e.g., Mantzicopoulos, 1997; Alexander, Entwisle, & Dauber, 1994; Gottfredson, Fink, & Graham, 1994) have reported positive effects of retention, to varying degrees. In some cases, small academic gains may be seen (Mantzicopoulos, 1997), particularly when comparing retainees with younger, same-grade peers (Alexander, et al., 1994). However, a larger number of studies have found benefits to be short-term; within two to three years, the achievement of retainees is equal to or lower than that of both same-grade and same-age peers (Jimerson, 1999; Jimerson, Carlson, Rotert, Egeland, & Sroufe, 1997; Mantzicopoulos, 1997, Mantzicopoulos & Morrison, 1992; McCoy & Reynolds, 1999). In addition, it is interesting to note that increases in performance are most often found only in math; reading remains a consistent deficit in retained children (Jimerson et al., 1997; Mantzicopoulos, 1997).

Despite a few studies in which positive results have been reported, the overwhelming majority of research indicates that grade retention does not have a remedial effect; rather, it operates to produce the direct opposite of desired effects. Reported outcomes have included lower reading achievement as measured by standardized tests (McCoy & Reynolds, 1999; Jimerson, 1999), increased social and behavior problems (Pagani, Tremblay, Vitaro, Boulerice, & McDuff, 2001), and increased dropout rates (Jimerson, 1999). These types of results were consistent across many studies using a variety of designs and analyses.

Studies describing negative outcomes and warning against the practice of retention are plentiful; however, it is imperative that we interpret these outcomes carefully, as it is very difficult to design a methodologically sound study when looking at outcomes of grade retention. The grade retention literature abounds with methodological problems, including the absence of comparison groups (McCoy & Reynolds, 1999; Meisels & Liaw, 1993), a lack of consideration of socio-economic variables (Mantzicopoulos, 1997; Mantzicopoulos & Morrison, 1992), and a lack of longitudinal data (Buntaine & Costenbader, 1997). Perhaps most relevant to the current study, all previous studies have failed to indicate what, if any, remediation efforts were made in conjunction with grade retention (e.g., Armistead, Kempton, Lynch, & Forhand, 1992; Gottfredson et al., 1994; Jimerson, 2001). This raises questions about the application of past results to the educational climate in Florida, as Florida's policy does not allow a child to repeat a grade without additional, intensive remediation in place. State law mandates remediation in the form of an Academic Improvement Plan (AIP) for each child who is retained in grade. These plans consist of instructional modifications as well as clear and measurable academic goals that relate to individual skill deficiencies. Examples of instructional modifications reflected in AIPs include pull-out services, one-on-one tutor instruction, peer tutors, and the employment of reading coaches. AIPs include periodic evaluation to determine if retained students are making progress towards their academic goals. It should be noted that while state law is quite clear regarding remediation requirements, specific data verifying treatment integrity is lacking at this time. Regardless, due to remediation requirements such as AIPs, holding negative beliefs

about the efficacy of the current student progression plan in Florida is premature as no research exists to date evaluating the outcomes of the plan.

*Purpose of the Study*

The purpose of this study was to evaluate the effects of retention practices in the State of Florida, as measured by scores on the state mandated standardized test (FCAT-Reading). Previous research has not examined retention within the context of a state mandated remediation program, as is part of Florida's policy. More specifically, it examined the relationship between the effects of retention and various student demographic characteristics, namely gender, race/ethnicity, and SES. In addition, the future performance of low-performing third grade students who were promoted to the fourth grade through good cause exemptions was investigated.

*Research Questions*

In order to explore the effectiveness of third-grade retention, this study utilized FCAT-Reading scores from the 2003 and 2004 administrations to address the following research questions:

1) What proportion of Florida students who scored at Level 1 and were retained at the end of the third grade in 2003 scored at Level 2 or higher on the 3rd grade FCAT-Reading in 2004 a) at the overall state level, and b) by school district?
2) For Florida students who scored at Level 1 and were retained at the end of third grade in 2003, what is the relationship between reading performance levels on the 3rd grade FCAT-Reading in 2004 and a) gender, b) race/ethnicity, and c) SES?

3) For Florida students who scored at Level 1 on the third grade FCAT-Reading in 2003, but were promoted due to good cause exemptions, what proportion for each exemption scored at Level 2 or higher on the fourth grade FCAT-Reading in 2004 a) at the overall state level, and b) by school district?
4) What is the likelihood that a 2003 third grade retainee's attainment of Level 2 or higher on the 2004 third grade FCAT-Reading is related to gender, SES, and race/ethnicity?

*Significance of Study*

A review of the literature reveals that no study exists in which academic modifications implemented during the retained year are clearly identified or delineated. It is unknown whether the students included in previous studies received instruction during the retained year that was quantitatively and/or qualitatively different from that of the previous year. Similarly, past research has not considered variables involving remediation strategies in the evaluation of retention outcomes; therefore, it is not possible to understand the effects of goal setting or monitoring during the retained year. Because the State of Florida mandates a systematic plan of remediation for every child who is retained in the third grade (Florida Department of Education, 2003f), the current study will examine the outcomes of retention among students who are required to receive intense remedial instruction during the retained year. The findings of this study could make an important contribution to the literature on retention in addressing an area that has heretofore gone unexamined.

*Definition of Terms*

*Florida Comprehensive Assessment Test (FCAT)*. The FCAT is a standardized test designed to measure student performance progress in reading, writing, and mathematics in accordance with the Sunshine State Standards. It also represents the foundation of the federally mandated and state created accountability system (Florida Department of Education, 2003c).

*Sunshine State Standards*. The Sunshine State Standards are a set of benchmarks developed by the Florida Department of Education and adopted in 1996. These standards identify the academic skills that the State of Florida wants its students to have attained at each grade level (Florida Department of Education, 2003c).

*Retention in grade*. This refers to the act of repeating a grade level. This study focuses only on students who were retained in third grade after failing to score Level 2 or above on the FCAT. In Florida, retention is to be accompanied by intensive remediation efforts.

*Academic Improvement Plans (AIP)*. An AIP is a set of formalized instructional modifications and related goals that are designed to address specific skill deficits in any child who is not meeting academic benchmarks in a timely manner; AIPs are state mandated in Florida (Florida Department of Education, 2002).

*Good cause exemptions*. These exemptions are six guidelines indicating the circumstances under which a student scoring Level 1 on the FCAT can still be promoted to the fourth grade. They include "[a]limited English proficient (LEP) students with less than two years in an English for Speakers of Other Languages (ESOL) program, [b]students with disabilities whose individual educational plan (IEP) indicated that

participation in the FCAT was not appropriate, [c]demonstration of an acceptable level of performance on an alternate assessment…or scoring at the 51st percentile or higher on the norm referenced test portion of the FCAT, [d]demonstration of proficiency in Sunshine State Standards through a student portfolio, [e]students with disabilities who participate in the FCAT…(but)…still demonstrate a deficiency in reading after more than two years of intensive remediation and were previously retained in kindergarten, first, second, or third grade, and [f]students who still demonstrate a deficiency in reading after two or more years of intensive remediation and were previously retained in kindergarten, first, second, or third grade for a total of two years." (Florida Department of Education, 2003e)

*No Child Left Behind Act*. The No Child Left Behind Act is a piece of federal legislation passed on January 8, 2002, directly addressing the reading crisis in this country; it established strict guidelines for school and state accountability. The Act wants every child to be reading proficiently by the year 2013-2014 (NCLB, 2002).

*Adequate Yearly Progress*. States are required by NCLB to set annual goals for adequate progress for their schools and to measure progress toward the attainment of those goals in a systematic way each year through standardized tests; funding is secured on the basis of states' success in meeting those goals (United States Department of Education, 2003c).

*Reading First*. Reading First is a federal program established through NCLB that distributes additional funds to assist states in the implementation of empirically-based reading programs. Funding is granted based on the number of children aged 5-17 years

who are considered low-income in each state (United States Department of Education, 2003d).

# CHAPTER TWO

## Literature Review

In this information age consisting of constant technological discoveries, it is astonishing to realize that only 32% of America's fourth graders read proficiently (United States Department of Education, 2002). Since the initial passage of the Elementary and Secondary Education Act in 1965, taxpayers have invested over $321 billion in public education; yet the average reading scores for 17-year-olds have not shown improvement since the 1970's (United States Department of Education, 2002). While scores of the top achievers have improved over time, those of the poorest performers have declined (United States Department of Education, 2002). Most of the children who are nonproficient in reading are minority students who live in poverty, creating a disturbing achievement gap. Recent federal legislation (e.g., No Child Left Behind Act) is reflecting acknowledgement of, as well as concern over this fact, by creating and implementing standards regarding student progression (United States Department of Education, 2003a). They are not only more rigorous for every child but also include new accountability procedures designed to ensure that substandard performance by any child is not tolerated as it was in the past.

### *Enactment of Federal and State Legislation*

The No Child Left Behind Act was enacted on January 8, 2002 and has changed the face of education across the country (NCLB, 2002). The law directly addresses the

country's reading crisis by establishing strict guidelines for both student performance as well as school and state accountability. The overriding goal of the law is for every child to be reading at a proficient level by the 2013-2014 school year (United States Department of Education, 2003b). To this end, each state is required to demonstrate Adequate Yearly Progress (AYP) in educating its students, though states are allowed some freedom in the adoption of procedures used to monitor this progress. States are required to set annual goals for adequate progress for their schools and measure progress toward the attainment of these goals in a systematic way each year through standardized tests; funding will be secured on the basis of states' success in meeting those goals. In addition, schools that do not meet these academic standards will be identified as needing improvement; parents of children enrolled in these schools can, with the aid of district funds, choose to send their child to a better performing public school or secure additional tutoring (United States Department of Education, 2003c).

Monitoring of academic progress is not a new concept in this country; however, the No Child Left Behind Act has identified and addressed some accountability issues. In the past, schools as well as states could report aggregated data across the total population of students; high scores could balance out very low scores, giving the impression that the school was effectively teaching all of its students. However, many schools were only effectively reaching a certain subset of students; minority, poor and children with disabilities or limited English proficiency (LEP) were consistently being "left behind". In response to this problem, data can no longer be aggregated across students, schools, districts or states; test scores are required to be disaggregated and reported by subgroups based on economic background, race/ethnicity, English proficiency, and disability status.

It is hoped that with these procedures in place, the achievement gaps that exist between disadvantaged students and other students will be diminished (United States Department of Education, 2003a).

In addition to lofty achievement goals and stringent accountability plans, the Act allows for additional federal funds to be distributed to states for the purposes of implementing scientifically proven, empirically based reading programs. The National Reading Panel (2000) has identified five basic skills important to early reading success, including phonemic awareness, phonics, fluency, vocabulary and comprehension. States can apply for funds through the federal "Reading First" program and funding will be granted them based on the number of their children aged 5-17 years who are considered low-income (United States Department of Education, 2003d). The funds will be used to teach the five basic skills in a systematic and evidence-based manner, and the students will be monitored closely (yearly state assessments) to ensure that they are moving toward success (United States Department of Education, 2003d).

Since each state can establish and enforce its own standards and design its own tests, it is helpful to look more specifically at Florida's response to the No Child Left Behind Act. The State of Florida is committed to compliance with the new federal law and has set intermediate goals to this end. The state has established goals according to which 31% of its students are to be reading proficiently by the 2003-04 school year, 48% by 2006-07, 65% by 2009-10, 82% by 2012-13, and 100% by 2013-14 (Florida Department of Education, 2003a). The State has applied for and received $52 million in annual funding (over $300 million across six years) in Reading First funds and plans to use the money in schools around the state with large numbers of at-risk children (Florida

Department of Education, 2003b). These schools must demonstrate through a competitive application process that they are committed to providing not only the required research-based reading programs, but an additional block of time devoted only to the use of these techniques as well. Children in these schools are to receive nearly twice the amount of reading instruction as children in typical schools. The hope is that intensive reading intervention will have a preventative effect.

Prevention programs such as these have received a tremendous amount of attention and are argued to be key in solving the reading crisis in the State. In addition to prevention, the State of Florida has implemented what some would consider to be rather drastic intervention strategies to address other current educational issues. New retention policies went into effect on January 7, 2003, the consequences of which continue to sweep the state (The Florida Senate, 2003). These policies include mandatory retention at the end of grade 3 for students whose reading deficiencies are not successfully remediated, as measured by performance on the Florida Comprehensive Achievement Test (FCAT). This policy is designed to align Florida policies with federal law and indicate that the educational system in this state will not let any child fail to learn to read; it will instead continue its efforts in assisting all children to achieve academic success. Inherent in the retention policy is an emphasis on prevention and/or remediation of reading deficiencies, which will be addressed later, as well as strong ties to the standardized test used to quantify those deficiencies, the FCAT.

In looking at Florida's standardized test, the standards that it purports to measure cannot be ignored. The Sunshine State Standards, developed by the Florida Department of Education and adopted in 1996, identify the academic skills that the State of Florida

wants its students to have attained at each grade level (Florida Department of Education, 2003c). For each grade, a series of benchmarks have been identified which, when met, move a child toward meeting the standards for that grade. Within the State of Florida, the FCAT represents the foundation of the federally mandated and state created accountability system; it purports to directly measure students' progress in reading, writing, and mathematics as outlined in the Sunshine State Standards. The reading portion of this test is administered every year in grades 3-10. Students receive a scale score for each subject, and from that scale score they are categorized as performing at one of five levels, with Level 1 being the lowest and Level 5 being the highest. A great deal of attention is given to the lowest two levels; at Level 1, students are having "little success with the challenging content of the Sunshine State Standards." Students who achieve Level 2 are assumed to be having "limited success" with the same content. Statewide policy mandates that unless a "good cause exemption" can be demonstrated, all students achieving at Level 1 will be retained at the third grade level (The Florida Senate, 2003). These students are assumed to be struggling with learning to read such that they have not mastered the most basic skills at the most basic levels. Without these basic skills, chances of success in subsequent grades are severely limited, as the approach to teaching reading shifts at this critical juncture. Prior to and through the third grade year, students are learning to read; beginning in fourth grade, students are instead expected to utilize reading skills in all areas of their schooling. In essence, they are now "reading to learn" (Just Read, Florida, 2003).

A wave of concern flooded the state as the 2003 FCAT results were disseminated. According to the scores reported, almost one out of four (23%) of Florida's third grade

students were reading at Level 1, indicating potential retention for the following school year of approximately 47,000 students (Florida Department of Education, 2003d). Six guidelines, or "good cause exemptions" indicated those circumstances under which a student scoring Level 1 on the FCAT could still be promoted to the fourth grade. These included "[a]limited English proficient (LEP) students with less than two years in an English for Speakers of Other Languages (ESOL) program, [b]students with disabilities whose individual educational plan (IEP) indicated that participation in the FCAT was not appropriate, [c]demonstration of an acceptable level of performance on an alternate assessment…or scoring at the 51st percentile or higher on the norm referenced test portion of the FCAT, [d]demonstration of proficiency in Sunshine State Standards through a student portfolio, [e]students with disabilities who participate in the FCAT…(but)…still demonstrate a deficiency in reading after more than two years of intensive remediation and were previously retained in kindergarten, first, second, or third grade, and [f]students who still demonstrate a deficiency in reading after two or more years of intensive remediation and were previously retained in kindergarten, first, second, or third grade for a total of two years" (Florida Department of Education, 2003e).

Historically, grade retention has been viewed in education as the exception, rather than the norm, yet has been a widespread practice despite persisting empirical evidence that it is not typically a successful intervention strategy (Jimerson, 2001; Phelps, Dowdell, Rizzo, Ehrlich, & Wilczenski, 1992). Despite the promotion of 12,403 (6.5%) students due to "good cause", a total of 28,028 (14.6%) third grade students were retained after the 2002-2003 school year due to Level 1 scores on the FCAT (Florida Department of Education, 2003d). In view of the tremendous numbers of affected children, it is

important to return to the research to examine these issues, looking past the longstanding assumptions, and objectively re-examining the literature with regards to the intents, functions and outcomes of retention. The body of retention literature is extensive and spans several decades; however, much of the earliest research is saddled with substantial methodological flaws, including a lack of comparison group (Simmons & Blyth, 1987), an insufficient sample in terms of size and diversity (Jackson, 1975). In addition, much of the earlier work in this area was unpublished. In an early review (Holmes, 1989) of 63 retention studies, only 20 had gone through the rigorous publication process. This speaks to the quality of the earliest studies, as the scientific quality could not be monitored. Due to these issues, this review will not attempt to be exhaustive; rather, it will focus on the most recent and relevant research to date and will provide a synthesis of the findings. This review is organized by the following categories: (a) characteristics of retained children, (b) evidence of positive outcomes of retention, (c) evidence of negative outcomes of retention, (d) factors that determine/contribute to retention outcomes, (e) methodological issues to consider, and (f) future directions in retention research.

*Characteristics of Retained Children*

While low academic performance has been assumed to be the primary reason to retain a child in grade, significant differences in achievement between retained students and low-achieving-but-promoted students have not been found, suggesting that social and/or behavioral factors influence decisions to retain (Jimerson, Carlson, Rotert, Egeland, & Sroufe, 1997). Over the years, certain factors have been identified that consistently characterize retained students, enabling us to predict who is more likely to be retained in the future. These factors include being male, an ethnic minority, and socially

immature. Additionally, having poor emotional health, maladaptive or aggressive behaviors, and reading or math problems have been found to put a child at higher risk (Jimerson, 1999; Jimerson et al., 1997; McCoy & Reynolds, 1999; Reynolds, 1992). Children also are more likely to be retained if they live in poverty, change schools frequently, or have parents who are not involved in school, do not value education or have a low level of education themselves (McCoy & Reynolds, 1999; Reynolds, 1992). It is unclear whether these factors, particularly the demographic characteristics, are themselves decreasing the odds of academic success, or if they are serving as mediators on an unidentified variable to put those children at a higher risk of being retained.

As retained children generally lag behind their peers academically and often face numerous other social and behavioral disadvantages, it follows logically that one goal of retention would be remediation. Therefore, children who are retained should show an improvement in academic performance following additional time in the same grade; without this improvement, it would be hard to justify such a financially, as well as emotionally costly method of intervention. Though outcome research has yielded mixed results in this area, with findings generally indicating negative effects, several studies indicate that early retention can benefit the future academic careers of low achieving students.

*Evidence of Positive Outcomes of Retention*

A recent study by Mantzicopoulos (1997) serves as an example of positive outcomes resulting from retention. The 40 children in this study were selected in the spring of their kindergarten year, based on high inattention status (as measured by teacher rating using the Revised Problem Behaviors Checklist) from a matched group of 62

children who had participated in a previous study (Mantzicopoulos & Morrison, 1992). This larger group had been matched on demographic and academic variables including sex, age, at-risk status and reading/math achievement. Of the 40 children included in this study, 25 had been retained in kindergarten and 15 were promoted; all children were followed through the second grade. Measures included the SEARCH instrument, which is composed of 10 subscales designed to assess spatial and temporal orientation skills deemed necessary to beginning reading, the Revised Problem Behavior Checklist (RPBC) to measure the severity of problem behaviors, and the Total Reading and Total Math subtests from one of two standardized achievement tests (Stanford Achievement Test or California Test of Basic Skills). Children were given the SEARCH screening during their kindergarten year and were rated with the RPBC each year. Academic achievement tests were administered near the end of each school year.

Repeated measures multivariate analyses of covariance were performed for both same-age and same-grade comparisons. The results for the same-grade comparisons indicated a significant main effect for math achievement, favoring the retained group ($F = 5.63$, $p<.05$). In addition, the adjusted mean differences showed that retained children performed above the national mean in math while promoted children remained below the national mean at first and second grades ($M_{Adj} = -0.25$ and $-0.18$), respectively. No significant differences were found in reading performance.

Same-age comparisons yielded similar results; retained children performed significantly better than promoted children in math, while the reading performance of both groups showed a slight decline over time, with comparable levels at the end of second grade. The author concludes that retention is not a beneficial educational

intervention due to the lack of consistent findings in reading; however, for the purposes of this review, the clear increase in math performance cannot be overlooked.

Alexander, Entwisle, & Dauber (1994) offer more conclusive evidence of retention as a positive intervention tool. Though never achieving at the same level as never-retained students, the retainees in their study show marked improvements over time. The authors utilized a stratified random sample of 800 Baltimore children entering first grade and monitored their school progress for eight subsequent years. The design was longitudinal, leading to subject groups that were too complex to be separated into distinct retained versus promoted groups based on one point of time. For instance, 127 first graders were retained in year 1 of the study. Over the next eight years, because of erratic grade-level progressions such as double retentions (49), triple retentions, (3), mid-year retentions (1), mid-year promotions (17), and double promotions (12), only nine of the original 127 students who had been retained in year 1 were on grade level after eight years. In a similar vein, by year 8 of the study, 142 children were in the seventh grade, one year behind the progression schedule that would be expected if there were no deviations; however, only 22 were actually seventh grade repeaters; the vast majority are a year behind because of retention(s) earlier in their school careers. This design resulted in a tremendous amount of long-term data, garnered mainly from school records (including test scores and grade reports), one-on-one interviews with the children and their parents, and self-administered questionnaires by teachers. This study was unique in comparison to other recent retention research in that it did not employ any type of matched control group. As will be discussed later, an appropriately matched comparison group is very difficult to construct and define consistently across studies, leaving the

possibility that an undetermined variable could skew the results; these authors chose to circumvent these issues by using statistical methods to adjust for prior performance, as well as demographic risk factors such as low socioeconomic status, low parent education, minority status, and low school readiness. The statistical adjustments allowed them to make appropriate comparisons between retainees, those who performed similarly but were promoted, and all remaining never-retained children. In addition, the researchers compared students' post-retention performance to their own pre-retention performance in order to examine a change in trajectory. To measure achievement, the researchers administered two subtests from the California Achievement Test battery (reading and math) twice annually. The reading portion consists of 20-40 items (20 items in first and second grade versions, 27 items in third grade version, and 40 items for all subsequent versions), while the math subtest consists of 36-45 items (36 items in first grade, 40 items in second grade, and 45 thereafter). The authors report scale scores that have been calibrated to account for all versions of the test, allowing for meaningful interpretations across grade levels. With regard to the CAT-R (reading), children who were retained at the end of first grade ($n = 127$) lagged significantly behind their same-age, never-retained peers at the beginning as well as at the end of the first grade (33.5 points lower in the fall; 59.5 points lower in the spring; $p<.01$), but when compared to those same peers after the retained year, had made up seven points in reading and four in math ($p<.01$). Gains were larger when comparing the children retained in first grade to younger, same-grade peers; they scored 17 points higher in math ($p<.01$). Unfortunately, these gains were not maintained past this first follow-up (i.e., after one year of retention in first grade); these children's scores in later years remained similar to what they had been in the pre-

retention year. The authors noted that retention in the first grade seems qualitatively different from retention in later grades; they remain unclear as to whether this is due to the first grade retention itself or whether first grade retainees' academic difficulties are so severe. They found that for children retained in the second and third grades, results were more positive. Second grade children were farther behind their peers at the end of the year than the first grade retainees had been; however, after retention for one year, they had improved their standing relative to same-age peers. At the time of retention, they scored an average of 52.0 points lower than their peers; after one year of retention in second grade, they scored only 34.7 points lower ($p<.01$). In addition, it appears that those gains were long-term, as by the seventh year of the study they were only 24.8 points lower than never-retained peers ($p<.01$). This pattern of improvement was even stronger for children retained in third grade, who lagged 28.4 points behind their same age peers at the point of retention, but only 9.7 points behind by year 7 of the study ($p<.01$). The authors are quick to point out that these retainees are still lagging far behind their never-retained peers, seen in the adjusted scores, but especially in the unadjusted numbers. They feel it is crucial, however, to remember the retained children's pre-retention problems. Whatever the problem causing poor achievement, it was present before the retained year, leading these authors to conclude that the retention year actually improved a previously charted course. They hypothesized, based on early test scores, that without retention, these children would have continued to fail, and that the gap between their scores and those of never-retained children would have continued to widen. While this explanation makes conceptual sense, it lacks empirical basis, as this study lacks the data to pinpoint specific variables involved in school failure, as well as detailed

information regarding which aspects of retention made a positive impact on children's academic careers.

To summarize, these studies and a small number of others (Gottfredson, Fink, & Graham, 1994; Peterson, DeGracie, & Ayabe, 1987) have reported positive results to varying degrees. Unfortunately, results such as these appear to be in the minority, as past research has consistently shown retention to be an overwhelmingly negative experience for children. Small academic gains may be seen in certain retainees (Mantzicopoulos, 1997), especially when comparing them with younger, same-grade peers (Alexander et al., 1994); however, these gains are apparently temporary and as these and other studies have shown, within two to three years, their achievement is equal to or lower than that of both same-grade and same-age regularly promoted students (Jimerson, 1999; Jimerson et al., 1997; Mantzicopoulos, 1997; Mantzicopoulos & Morrison, 1992; McCoy & Reynolds, 1999). Interestingly, studies that have found modest, temporary increases in performance most often find them in math; reading is a consistent and distinct deficit in retained children (Jimerson et al., 1997; Mantzicopoulos, 1997). This finding is of particular relevance when considering that the retention policy in the State of Florida establishes reading remediation as its primary goal. It should also be noted that according to the existing literature, socio-emotional and behavior deficits such as high inattention, immature or aggressive behavior, low self-esteem, and poor peer relationships tend to be exacerbated rather than improved by retention (Jimerson, 1999; Jimerson et al., 1997; Mantzicopoulos, 1997). The direct effects of socio-emotional/behavioral health on academic performance are not the focus of this review; however, issues such these are likely to have a negative overall effect on a child's school experience.

*Evidence of Negative Outcomes of Retention*

Numerous studies have attempted to study both behavioral and academic factors in retention (Ferguson et al., 2001; Gottfredson et al., 1994; Mantzicopoulos, 1997; Mantzicopoulos & Morrison, 1992). Pagani, Tremblay, Vitaro, Boulerice, and McDuff (2001) maintained this dual focus in their research efforts. Their 1,830 subjects were selected from a larger sample (6,397) of kindergarten students who had been randomly selected for a separate study and followed through age 12 (Vitaro, Desmariais-Gervais, Tremblay, & Gagnon, 1992). The independent variable was grade retention (one year maximum) and the dependent variables were teacher ratings of children's global academic performance as well as scores on the Social Behavior Questionnaire, which is based on teacher rating and comprised of items from the Preschool Behavior Questionnaire and the Prosocial Behavior Questionnaire. In order to control for differences in naturally occurring trajectories expected from children of differing achievement levels, the researchers used a statistic relatively unique to retention research, a basic autoregressive model. Testing separately for boys and girls, results indicated that grade retention negatively affected children's development, regardless of prior characteristics. With regards to academics, for example, 10-year-old boys experienced negative deviations from the expected trajectory after being retained between the ages of 6 and 8 years (path = -.12, $p<.01$). This negative effect was still evident at age 12. In girls as well, early grade retention had a profound and lasting negative effect on later academic performance (measured at ages 10 and 12), as their trajectories had diverted negatively from the expected path at age 10 (path = -.07, $p<.01$) and remained off course at age 12 (path = -.07, $p<.05$). This model also evidenced negative behavioral outcomes,

particularly in boys, for whom a persistent and worsening of disruptive behavior was reported. The authors hypothesized that these behavioral outcomes and the emotional component that accompanied them could have interfered with the learning process and could, therefore, have been acting upon the academic outcomes that were observed. While their hypothesis is worth further examination, the most salient information involves the authors' evidence-based conclusions that retention leads to long-lasting negative effects.

McCoy and Reynolds (1999) provided another illustration of negative retention outcomes in the findings from their Chicago Longitudinal Study. This study examined school achievement of children who were graduates of a government funded kindergarten program and who were at risk for underachievement due to poverty. Participants were seventh and eighth grade students (depending on retention status) at the time of enrollment, thus eliminating the opportunity to match subjects based on pre-existing achievement or behavior characteristics known to predict retention. Of 1,164 children enrolled, 315 had been retained in grade at least once. Students were assessed using the reading comprehension and mathematics subsections of the Iowa Test of Basic Skills (ITBS, Level 13 or 14; $\bar{x}$=145.06, SD=22.38) and hierarchical multiple regression was used to analyze the results. When comparing same-age retained versus non-retained children, grade retention was found to consistently predict reading and mathematics performance ($R^2$ = .47, $R^2$ = .57, respectively) as measured by this standardized test. Retention was significantly and negatively correlated with reading achievement scores (10.6 standard score points lower, $p<.01$), even after the inclusion of sociodemographic factors and early adjustment indicators in the analysis. The same finding was true of

mathematics achievement (9.7 standard score points lower, $p<.01$), although fewer sociodemographic characteristics were related. The authors also conducted analyses using same-grade comparisons; again, retention was found to predict negative outcomes. Retained seventh grade students scored significantly lower than promoted seventh graders in reading achievement ($\beta = -4.6$ standard score points, $p<.001$); they also scored lower in mathematics ($\beta = -1.3$ standard score points), but the difference was not significant. The achievement gap between retained and promoted (same grade) students was found to have widened four years later, despite the stated goal of retention as decreasing the gap. The authors concluded, therefore, that retention was ineffective as an intervention and even harmful to students' future academic achievement.

Such unequivocal conclusions that retention outcomes have a negative impact on children are not unique; still, caution needs to be utilized when interpreting results, as methodological issues abound in retention research. Because of several limitations, findings from the McCoy and Reynolds (1999) study necessitate further corroboration. The authors used a correlational design with a low-income sample, limiting the generalizability; however, it is the retrospective nature of the study that is most problematic. Enrolling subjects after the retention occurred disallowed an examination of predisposing factors, as well as examination of the trajectories of experimental and control groups. In 1997, partly as a response to these and other design deficits, Jimerson et al. attempted to examine retention effects by utilizing a prospective longitudinal design. Since matching is not typically feasible when looking at students only after retention has occurred, Jimerson et al. utilized subjects whose mothers had been participating in another longitudinal study prior to the birth of their children. These

children were identified prenatally as being at-risk for social and emotional development problems due to maternal characteristics. They were followed through age 16 and placed into one of three research groups: retained once in grades K-3 (n = 29), low-achieving promoted (n = 50) and control (n = 100). The children in the low-achieving promoted group served as a matched comparison group, as they did not differ from the retained children in terms of academic achievement or intellectual functioning as measured by standardized assessments (Peabody Individual Achievement Test, Wechsler Preschool and Primary Scale of Intelligence, Wechsler Intelligence Scale for Children-Revised, Woodcock-Johnson Achievement Test-Revised). Interestingly, they did display fewer maladjusted behaviors in the classroom ($p<.05$) and higher levels of emotional health ($p<.01$) than did retained students, based on teacher interview, the Child Behavior Checklist-Teacher Form, and emotional health/self-esteem measures developed by the authors. Short-term effects were examined one year following retention; retained students did not appear to display an advantage over the comparison students in academic functioning or school adjustment. When prior achievement levels were covaried out, the retained group showed a slight advantage in math achievement ($p<.01$). Students were assessed in the sixth grade as well as at 16 years of age to examine long-term effects of retention. Results favored the comparison group on measures of emotional health ($p<.05$), but the two groups did not differ in academic achievement. Essentially, despite an extra year of instruction, the retained students were comparable to low-achieving promoted students in the short term as well as years later. Though this is not as negative a conclusion as the outcomes found by other studies, neither is it establishing retention as a positive, successful intervention for underachieving students.

In an effort to extend these results, Jimerson (1999) followed these same students into their 21st year of life. The students were assessed in 11th grade and again at age 19, in an effort to answer questions regarding academic adjustment (as measured by grade point average, credits earned, behavioral problems and attendance), dropout status, receipt of a certificate of high school completion and postsecondary education. In reference to academic adjustment in 11th grade, the retained group of students achieved significantly lower levels as compared to both the comparison ($F = 6.59$, $p<.01$) and the control ($F = 13.95$, $p<.001$) groups. The latter two groups were not significantly different from each other. By age 19, a larger percentage of retained children had dropped out of high school as compared to both the comparison ($\bar{x} = 3.57$, $p<.05$) and control groups ($\bar{x} = 13.79$, $p<.001$), with the comparison group having a significantly greater number of dropouts than the control group ($F = 3.77$, $p<.05$). By age 20, a significantly lower percentage of the retained students had received a certificate of high school completion as compared to both the comparison ($\bar{x} = 5.44$, $p<.01$) and control ($\bar{x} = 23.66$, $p<.001$) groups. As expected, the number of retained students who eventually enrolled in postsecondary education was significantly lower than the number of control students ($\bar{x} = 8.74$, $p<.01$). Incidentally, as adults, retained students were less likely than were comparison students to have a full-time job, be a full-time student, or be involved in a combination of work and school ($F = 4.62$, $p<.05$). They also earned lower wages ($F = 4.09$, $p<.05$) and were rated as less competent in the workplace ($F = 4.39$, $p<.05$). Low-achieving-but-promoted students were more comparable to the control group than to the retained group in regards to all previously described employment outcomes.

*Factors that Determine/Contribute to Retention Outcomes*

It appears then, that both positive as well as negative outcomes are possible for students retained in early elementary school. Consequently, it is reasonable to ask which factors, if any, can be used to first determine why some children experience success after a retention year, in the hopes that the information can be utilized to manipulate outcomes for all children. Ferguson, Jimerson & Dalton (2001) posed this question and examined within group variables in an attempt to answer it. These authors looked at the performance of retained (only once in grades K-2; n = 58) and promoted (n = 48) students on several school readiness (Brigance Kindergarten and First Grade Screen, Gesell School Readiness Test, Metropolitan Readiness Tests) and standardized achievement instruments (Science Research Associates' Survey of Basic Skills, 2nd and 5th grades; Stanford Achievement Test 8th Edition, 8th and 11th grades) as well as on academic measures such as grade point average (GPA) and teacher ratings, which were administered at various points from 2nd through 11th grades. They defined "successful" retained students as those having a mean seventh grade GPA of 3.2 and a mean eighth grade SAT (Stanford Achievement Test) score in the 53rd percentile. These students represented 25% of all retained students, had scored higher on early school readiness measures ($p<.05$), were significantly younger ($p<.05$), and were less aggressive ($p<.01$) than unsuccessful retained peers. They were from families with a higher level of maternal education; their parents also placed a higher value on education for their child. These results suggest that the outcomes of retention, positive or negative, are dependent on variables beyond the retention itself. Perhaps criticism that focuses solely on retention should be replaced with investigation of mediating factors that may influence outcomes.

*Methodological Issues to Consider*

The potential impact of mediating variables on retention effects is only one of many reasons to reexamine longstanding assumptions that retention is harmful to students. For example, the retention literature has been characterized by various and longstanding methodological issues, which underlie the difficulties in making accurate interpretation of results. Design deficiencies include, but are not limited to, the absence of comparison groups (McCoy & Reynolds, 1999; Meisels & Liaw, 1993), a lack of consideration of socio-economic variables (Mantzicopoulos, 1997; Mantzicopoulos & Morrison, 1992), and a lack of longitudinal studies (Buntaine & Costenbader, 1997; Dennebaum & Kulberg, 1994; Mantzicopoulos & Morrison, 1992). Studies conducted in the last decade have been markedly improved over their predecessors with regard to methodology; however, there is still considerable variability in methodological quality. For example, in the studies reviewed there is evidence of a lack of consistency in composition of comparison groups, with some studies making same grade comparisons others making same age comparisons (Mantzicopoulos, 1997; McCoy & Reynolds, 1999; Reynolds, 1992). There is also a lack of adequate matching in control groups (Mantzicopoulos & Morrison, 1992), and a lack of student information prior to retention (McCoy & Reynolds, 1999).

*Future Directions in Retention Research*

Finally, it should be noted that *all* retention studies to date have examined retention as intervention (Armistead, Kempton, Lynch, & Forhand, 1992; Gottfredson et al., 1994; Jimerson, 2001; Phelps et al., 1992). There is no mention of remediation efforts in conjunction with, or in lieu of, retention. In addition, the literature thus far has

neglected to examine individual components of retention in an attempt to determine exactly which variables are associated with the negative effects. These are perhaps the most significant characteristics of the available literature and the overriding presence of such methodological limitations render the application of findings to the present student progression plan in the State of Florida inappropriate.

The State of Florida has been clear in its emphasis on remediation. That retention is part of its remediation plan has caused alarm; however, systematic efforts to address academic difficulties are mandated throughout the plan. Academic Improvement Plans, or AIPs, (Florida Department of Education, 2002) are a set of formalized instructional modifications and related goals, both of which are mandated in state law and designed to address specific skill deficits in any child who is not meeting academic benchmarks in a timely manner. The purpose of an AIP is to delineate the specific instructional and environmental variations that, when implemented as designed, are intended to encourage remediation of a student's academic deficiencies. Examples of AIP interventions include 1) time spent with a peer tutor, 2) one-to-one instruction with a reading coach, and 3) extra time (30 minutes per day) devoted to reading instruction. Important features of AIPs are their provision of a direct link to previous assessment as well as their inclusion of objective, measurable goals; both are critical for effective progress monitoring, which is required for every AIP. Children are reevaluated at specific intervals to determine if they are making progress toward specified goals; these goals may also be re-evaluated and revised if necessary. AIPs are available to any struggling student, regardless of age or retention status, as the goal is to remediate problems before retention becomes necessary. If a child's reading performance is determined to be at Level 1 on the FCAT at the end of

third grade and no cause for exemption is found, he will then be retained; still, remediation efforts will continue and likely intensify (Florida Department of Education, 2002). It should be noted that while state law is quite clear regarding remediation requirements, specific data verifying treatment integrity is lacking at this time.

Regardless, conclusions about the virtues of Florida's retention policy based on past research alone may not be valid in the current educational atmosphere of reform and accountability. The State has chosen a relatively radical response to dramatic changes in federal policy. This is a time of transition, but it affords the opportunity to add to the body of literature in a meaningful way. Retention has been long considered an inadequate intervention strategy for low-achieving students; the emergence of new policies and practices involving retention calls for a new generation of literature investigating this model of "retention plus remediation".

# CHAPTER THREE

## Method

### *Purpose of the Study*

The purpose of this study was to evaluate the effects of third-grade retention practices in the State of Florida, as measured by scores on the state mandated standardized test (FCAT). Previous research has not examined retention within the context of a state mandated remediation program, as is part of Florida's policy. In addition, this study aimed to determine the relationship between the effects of retention and various student demographics (gender, SES, race/ethnicity) as well as examine the future performance of low-performing students who were promoted through good cause exemptions.

### *Participants*

Participants included all students whose reading proficiency was classified as Level 1 (scaled score between 100 and 258) on the reading portion of the 2003 FCAT and were either retained in the third grade or promoted to the fourth grade based on a good cause exemption. Students who met the criteria but had missing data for any of the variables of interest (gender, race, SES) were excluded. Students who attended a lab school or attended school at home were also excluded. The retained students included in the study failed to demonstrate reading achievement thought to be necessary in order for them to experience success in the fourth grade. Consequently, these students completed

the third grade curriculum for a second time during the 2003-2004 school year, this time according to an individualized intervention plan, or Academic Improvement Plan (AIP). The students who were promoted due to a good cause exemption, despite scoring at Level 1 on the 2003 FCAT Reading, experienced a regular fourth-grade curriculum.

To examine the impact of retention more closely, each student was coded on several characteristics, including race, gender and socioeconomic status. Race distinctions included Caucasian, African American, Asian-Pacific Islander, Hispanic and Alaskan Native/Native American. Socioeconomic status was categorized by each student's eligibility for Free and Reduced Lunch, as determined by the income parameters set by the State of Florida.

*Florida Comprehensive Assessment Test*

As it is the dominant factor in the State of Florida's policy for the evaluation of the academic progression of its students, this study utilized scores from the Florida Comprehensive Assessment Test (FCAT) in order to examine the effectiveness of retention on low achieving third-grade students. According to the Florida Department of Education (2003), the FCAT is a criterion-referenced test developed by a panel of curriculum specialists. Its intent is to assess student achievement of the higher-order cognitive skills represented in the Sunshine State Standards (SSS). The FCAT reading SSS reports scores in four areas including; (1) main idea, plot and purpose, (2) words and phrases in context, (3) comparisons of cause/effect, and (4), reference and research. Included in the FCAT are literary passages, and informational passages. Another portion of the FCAT reading section is used normatively, comparing the students of Florida with the rest of the nation. The FCAT contains both multiple-choice and performance

questions; the multiple-choice questions are scored by computer, while the performance tasks are hand scored. During the months of February and March, the FCAT is administered to over 1.5 million students across the state. The completed tests are then sealed and sent to the Florida Department of Education

Scoring of the FCAT is based on item response theory (IRT, Lord & Novick, 1968). IRT theory assumes that student responses to individual questions are directly related to underlying achievement in a given content area. Cronbach's reliability coefficient estimates of the fourth-grade reading portion of the FCAT scores were reported by total score ($r = .88$), literary text ($r = .79$) and informational text ($r = .79$) (Florida Department of Education, 2002b).

Five levels have been defined for the FCAT reading achievement test, based on scaled scores ranging from 100 to 500. Level 1 represents the lowest level of reading achievement; each subsequent level represents an increasing level of reading proficiency. In 2003, the five levels were defined according to the following scaled scores: Level 1: 100-258, Level 2: 259-283, Level 3: 284-331, Level 4: 332-393, and Level 5: 394-500. In reading, it is assumed that a student who achieves a scaled score falling in the category of Level 1 is experiencing little success with the content of the Sunshine State Standards (Florida Department of Education, 2003c).

*Procedures*

Upon the approval of the IRB, the primary investigator collected the following data from the Florida Department of Education on students who scored Level 1 on the 2003 FCAT-Reading: retention status, reading scores, race, gender, socioeconomic status, and good cause promotion status. The data collection procedure was repeated with the

data from the 2004 FCAT-Reading administration to examine changes in FCAT scores of the students who were either retained or promoted due to a good cause exemption. The data was then screened; students who did not meet established criteria or who had missing data were excluded.

*Data Analysis*

Descriptive analyses were performed on all the variables of interest and the results describe the participant pool in terms of gender, race, and SES.

To answer each research question data were subjected to appropriate statistical methods, as described below.

*Research Question 1: What proportion of Florida students retained at the end of the third grade in 2003 scored at Level 2 or higher on the 3rd grade FCAT-Reading in 2004 a) at the overall state level, and b) by school district?* For students who scored Level 1 and were retained in 2003, the number and percent who then scored Level 1, Level 2, and Level 3 or higher in 2004 was computed. The data were examined at both the state level and by district. The alpha significance level for the present study was established at .05.

*Research Question 2: For Florida students retained at the end of third grade in 2003, what is the relationship between their reading performance levels on the 3rd grade FCAT-Reading in 2004 and a) gender, b) race/ethnicity, and c) SES?* For each demographic variable (gender, SES, race/ethnicity), data were subjected to a Chi-square test of association to see if there is differential performance across groups for each of these variables. The alpha significance level was established at .05.

*Research Question 3: For Florida students who scored at Level 1 on the 3rd grade FCAT-Reading in 2003 but were promoted due to good cause exemptions, what proportion for each and for any exemption scored at Level 2 or higher on the fourth grade FCAT-Reading in 2004 a) at the overall state level, and b) by school district?* For students who scored Level 1 in 2003 but were promoted due to good cause exemptions, the number and percent of students who then scored Level 1, Level 2, and Level 3 or higher on the fourth grade FCAT in 2004 were computed. The data were examined at both the state level and by district. The alpha significance level for the present study was established at .05.

*Research Question 4: What is the likelihood that a 2003 third grade retainee's attainment of Level 2 or higher on the 2004 third grade FCAT-Reading is related to gender, SES, and race/ethnicity?* Data for retained students were entered into a logistic regression model to test for associations of gender, race, and SES with academic performance as measured by scores on the 3rd grade FCAT-Reading in 2004. The outcome variable, success, was defined as achievement at Levels 2-5 on the 2004 FCAT-Reading, and was coded as 1 = Successful and 0 = Unsuccessful. The model controlled for previous performance, measured by scaled scores from 2003 (SSR). The predictor variables were coded as follows: gender was coded as 1 = Female and 0 = Male, with female students acting as the referent, race/ethnicity was dummy coded for African American (1 = AA, 0 = Not AA), Hispanic (1 = Hispanic, 0 = Not Hispanic), and Asian (1 = Asian, 0 = Not Asian), with Caucasian students serving as the referent, and SES was coded as 1 = Low SES and 0 = Not low SES, with Low SES providing the reference. The alpha significance level was established at .05.

# CHAPTER FOUR

## Results

As stated previously, this study was intended to determine the relationship between the effects of retention in third grade and various student demographic characteristics (gender, SES, race/ethnicity) as well as examine the future performance of low-performing students who were promoted through good cause exemptions. This chapter begins with a description of the sample. Then the results of the analyses are provided for each of the four research questions addressed in this study. The data were analyzed using SAS Version 9.1. An alpha level of .05 was set for all statistical tests.

### *Description of Sample*

The sample for this study consisted of 29,317 students who scored at Level 1 on the 2003 third-grade Florida Comprehensive Assessment Test-Reading (FCAT-Reading) and were either retained or promoted to fourth grade through a good cause exemption for the 2003-2004 academic year. Students who attended a lab school or who received schooling at home were excluded from all analyses. In addition, students with missing data on two or more variables were excluded. Finally, each individual analysis excluded any students who were missing data on any one of the variables of interest.

Race and socioeconomic status were defined in this study according to the data made available by the State of Florida, Department of Education (FLDOE). The FLDOE defines race using six categories (Caucasian, African-American, Hispanic, Asian, Native

American, Mixed Race); however, the categories of Native American and Mixed Race were collapsed in this study due to small sample sizes and are included in the category labeled "Other". Florida defines SES by a student's eligibility status for receiving lunch at free or reduced prices; therefore, in this study, students who were eligible to receive free or reduced lunch (FRL) were coded as "Low SES" and those who were not eligible were coded as "Not low SES".

A breakdown of the final sample in terms of retention/promotion status at the end of third grade in 2003 by gender, race/ethnicity, and SES is reported in Table 1.

Table 1

*Students Scoring Level 1 on the 2003 FCAT-Reading and Retention/Promotion Status by Gender, Race/Ethnicity, and SES.*

| | 2003 FCAT-Reading Level 1 | | Retained in Third Grade | | Promoted due to Good Cause Exemption | |
|---|---|---|---|---|---|---|
| | n | % | n | % | n | % |
| Race/Ethnicity | | | | | | |
| Caucasian | 8524 | 29 | 5459 | 26 | 3065 | 35 |
| African American | 10891 | 37 | 8387 | 41 | 2504 | 29 |
| Hispanic | 9036 | 31 | 6200 | 30 | 2836 | 33 |
| Asian | 269 | 1 | 160 | 1 | 109 | 1 |
| Other | 591 | 2 | 408 | 2 | 183 | 2 |
| Gender | | | | | | |
| Male | 17294 | 59 | 11998 | 58 | 5296 | 61 |
| Female | 12022 | 41 | 8619 | 42 | 3403 | 39 |
| SES | | | | | | |
| Low | 23335 | 80 | 16713 | 81 | 6622 | 77 |
| Not low | 5860 | 20 | 3905 | 19 | 1955 | 23 |
| Total | 29317 | | 20618 | 70 | 8699 | 30 |

*Research Questions*

*Question 1. What proportion of Florida students retained at the end of the third grade in 2003 scored at Level 2 or higher on the 3rd grade FCAT-Reading in 2004 a) at the overall state level, and b) by school district?*

The number and percent of 2003 third grade retainees performing at Levels 1-5 on the 2004 FCAT-Reading were computed. After examination of the data, Levels 3-5 were collapsed due to small cell sizes. The resultant performance levels are reported in Table 2. As is shown, 62% (n = 12,806) of the third grade retainees scored at Level 2 or higher on the 2004 FCAT-Reading. More specifically, 21% (n = 4,424) of the retained students scored at Level 2, while 41% (n = 8,382) scored at Levels 3-5. Thirty-eight percent (n = 7,812) of the retained students scored at Level 1.

Table 2

*Performance of Retained Students on the 3rd grade FCAT-Reading in 2004 by Level*

| Reading Level | Number | Percent |
|---|---|---|
| 1 | 7812 | 38 |
| 2 | 4424 | 21 |
| 3-5 | 8382 | 41 |
| Total | 20618 | 100 |

Data were examined by district and no clear pattern of achievement on the FCAT-Reading was identified. Breakdown of student performance by district is reported in Appendix A.

*Question 2. For Florida students retained at the end of third grade in 2003, what is the relationship between their reading performance levels on the 3rd grade FCAT-Reading in 2004 and a) gender, b) race/ethnicity, and c) SES?*

Chi-square tests of Association were performed to test for relationships between each of these variables and performance on the 3rd grade FCAT-Reading in 2004.

*Gender and performance on the 2004 FCAT-Reading*. A total of 20,617 (58% male) students were included in the sample to examine performance of retainees by gender; one student was excluded due to missing data. The obtained Chi-square was statistically significant, $X^2$ (2, N = 20614) = 88.15, $p$<.0001, indicating that there was a significant relationship between performance of retained students and gender, as is shown in Table 3. Specifically, with regard to male retainees, a greater number than statistically expected scored at Level 1 (4,854 vs 4,546), while fewer than expected scored at Levels 3-5 (4,749 vs 4,877). Regarding female retainees, fewer than expected scored at Level 1 (2,957 vs 3,265), while a greater number than expected scored at Levels 3-5 (3,633 vs 3,504). However, the strength of the association between gender and performance on the 2004 FCAT-Reading was small (Cramer's V = .07).

Table 3

*Number and Percent of Retained Students of Different Performance Levels on the 3rd grade FCAT-Reading in 2004 by Gender.*

| | | Performance Level | | | | | |
|---|---|---|---|---|---|---|---|
| | N | Level 1 | | Level 2 | | Levels 3-5 | |
| Gender | | n | % | n | % | n | % |
| Male | 11998 | 4854 | 62 | 2395 | 54 | 4749 | 57 |
| Female | 8619 | 2957 | 38 | 2029 | 46 | 3633 | 43 |
| Total | | 7811 | 100 | 4424 | 100 | 8382 | 100 |

$\chi^2$ (2, N = 20617) = 88.15, p < .0001

*Race/ethnicity and performance on the 2004 FCAT-Reading.* A total of 20,614 (26% Caucasian, 41% African American, 30% Hispanic, 1% Asian, and 2% Other) students were included in the sample to examine performance of retainees by race/ethnicity; four students were excluded due to missing data. The obtained Chi-square was statistically significant, $X^2$ (8, N = 20614) = 161.23, $p$<.0001, indicating there was a significant relationship between performance of retained students and race/ethnicity (see Table 4). Specifically, with regard to African American retainees, a greater number than statistically expected scored at Level 1 (3,432 vs 3,177), while fewer than expected scored at Levels 3-5 (3,022 vs 3,410). Regarding Caucasian retainees, fewer than expected scored at Level 1 (1,829 vs 2,068), while a greater number than expected scored at Levels 3-5 (2,523 vs 2,219). However, the strength of the association between gender and performance on the 2004 FCAT-Reading was small (Cramer's V = .06).

Table 4

*Number and Percent of Retained Students of Different Performance Levels on the 3rd Grade FCAT-Reading in 2004 by Race/Ethnicity*

| | | Performance Level | | | | | |
|---|---|---|---|---|---|---|---|
| | N | Level 1 | | Level 2 | | Levels 3-5 | |
| Race/Ethnicity | | n | % | n | % | n | % |
| Caucasian | 5459 | 1829 | 23 | 1107 | 25 | 2523 | 30 |
| African-American | 8387 | 3432 | 44 | 1933 | 44 | 3022 | 36 |
| Hispanic | 6200 | 2362 | 30 | 1259 | 28 | 2579 | 31 |
| Asian | 160 | 58 | 1 | 38 | 1 | 64 | 1 |
| Other | 408 | 129 | 2 | 87 | 2 | 192 | 2 |
| Total | | 7810 | 100 | 4424 | 100 | 8380 | 100 |

$\chi^2$ (8, N = 20614) = 161.23, p < .0001

*Socio-economic status and performance on the 2004 FCAT-Reading.* A total of 20,404 (83% Low SES, 17% Not-low SES) students were included in the sample to examine performance by socioeconomic status (SES) as defined by free/reduced lunch (FRL) eligibility; 214 students were excluded due to missing data. The obtained Chi-square was statistically significant, $\chi^2$ (2, N = 20404) = 102.28, $p$<.0001, indicating that there was a significant relationship between performance of retained students and SES (see Table 5). Specifically, with regard to retainees with Low SES status (i.e., students eligible for FRL), a greater number than statistically expected scored at Level 1 (6,669 vs 6,422), while fewer than expected scored at Levels 3-5 (6,660 vs 6,890). Regarding retainees who were classified as Not low SES (i.e., students not eligible for FRL), fewer

than expected scored at Level 1 (1,056 vs 1,302), while a greater number than expected scored at Levels 3-5 (1,628 vs 1,397). However, the strength of the association between gender and performance on the 2004 FCAT-Reading was small (Cramer's V = .07).

Table 5

*Number and Percent of Retained Students of Different Performance Levels on the 3rd Grade FCAT-Reading in 2004 by SES Status*

| | | Performance Level | | | | | |
|---|---|---|---|---|---|---|---|
| | N | Level 1 | | Level 2 | | Levels 3-5 | |
| SES | | n | % | n | % | n | % |
| Low SES | 16964 | 6669 | 86 | 3635 | 83 | 6660 | 20 |
| Not low SES | 3440 | 1056 | 14 | 756 | 17 | 1628 | 80 |
| Total | 20404 | 7725 | 100 | 4391 | 100 | 8288 | 100 |

$\chi^2$ (2, N = 20404) = 102.28, p < .0001

*Question 3. For students who scored at Level 1 on the 3rd grade FCAT-Reading in 2003 but were promoted due to good cause exemptions, what proportion for each type of exemption scored at Level 2 or higher on the fourth grade FCAT-Reading in 2004 a) at the overall state level, and b) by school district?*

Altogether, 30% (N = 8,699) of the third-grade students who scored at Level 1 on the 2003 FCAT-Reading were promoted to the fourth grade due to good cause exemptions. Table 6 displays a breakdown of this group of students by type of exemption.

Table 6.

*Number and Percent of Students Promoted to 4th Grade due to Good Cause Exemption by Type of Exemption*

| Exemptions | N | % |
|---|---|---|
| 1. Proficiency demonstrated through alternative standardized reading assessment | 2189 | 26 |
| 2. Proficiency demonstrated through student portfolio | 1209 | 14 |
| 3. Limited English Proficient (LEP) | 1897 | 22 |
| 4. Did not participate due to Individualized Educational Plan (IEP) | 120 | 1 |
| 5. Previous retention (1 year) + IEP | 2852 | 33 |
| 6. Previous retentions (2 years) | 432 | 5 |
| Total | 8699 | 100 |

These promoted students took the 4th grade FCAT-Reading in 2004. Their performance levels on this test by type of good cause exemption are reported in Table 7. Levels 3-5 were collapsed due to small cell sizes.

Table 7

*Level 1 Promoted Students' Performance on the 2004 4th grade FCAT-Reading by Good Cause Exemption*

| | Level 1 | | Level 2 | | Levels 3-5 | |
|---|---|---|---|---|---|---|
| Type of Exemption | n | % | n | % | n | % |
| 1. Proficiency demonstrated through alternative standardized reading assessment | 687 | 31 | 742 | 34 | 760 | 35 |
| 2. Proficiency demonstrated through student portfolio | 491 | 41 | 392 | 32 | 326 | 27 |
| 3. Limited English Proficient (LEP) | 1308 | 69 | 301 | 16 | 288 | 15 |
| 4. Did not participate due to Individual Education Plan (IEP) | 98 | 82 | 13 | 11 | 9 | 8 |
| 5. Previous retention (1 year) + IEP | 2376 | 83 | 297 | 10 | 179 | 6 |
| 6. Previous retentions (2 years) | 338 | 78 | 64 | 15 | 30 | 7 |
| Total | 5298 | 61 | 1809 | 21 | 1592 | 18 |

Overall, 39% of the third-grade students promoted in 2003 due to a good cause exemption scored at Level 2 or higher on the 2004 FCAT-Reading, while 61% scored at Level 1. Looking more closely at the 40% of students who demonstrated proficiency through an alternative standardized reading assessment (Exemption 1) or through student portfolio (Exemption 2), data reveals a different pattern of performance than that of

students who were promoted without demonstrating proficiency in reading (Exemptions 3-6). Specifically, the percentages of students scoring at Level 2 or Levels 3-5 on the 4th grade FCAT-Reading in 2004 were quite high for those who were promoted due to Exemption 1 (69%, n = 1,502) and Exemption 2 (59%, n = 718) when examining this smaller subset of students. In contrast, fewer students promoted due to Exemptions 3-6 scored at Level 2 or Levels 3-5 in 2004.

In addition, the performance of all students who were promoted due to a Good Cause Exemption and who subsequently passed the 2004 FCAT (Levels 2-5) was examined and the results are displayed in Table 8.

Table 8

*Students Promoted through Good Cause Exemption in 2003 Scoring at Levels 2-5 on the 2004 FCAT-Reading by Proficiency Status*

| Proficiency Status | N | % |
|---|---|---|
| Proficient (Exemptions 1-2) | 2220 | 65 |
| Non-Proficient (Exemptions 3-6) | 1181 | 35 |
| Total | 3401 | 100 |

When considering only those students promoted through good cause exemption who scored at Level 2 or higher on the 4th grade FCAT-Reading in 2004, 65% had been promoted the previous year due to one of the two exemptions that require demonstration of reading proficiency through an alternative mechanism. In contrast, only 35% (n = 1181) had been promoted due to an exemption that did not require a demonstration of reading proficiency.

Data were examined by district and no clear pattern of achievement on the FCAT-Reading was identified. Breakdown of student performance by district is reported in Appendix B.

*Question 4: What is the likelihood that a 2003 third grade retainee's attainment of Level 2 or higher on the 2004 third grade FCAT-Reading is related to gender, SES, and race/ethnicity?*

To test the relationship between the likelihood that a third-grade retainee is successful on the 2004 FCAT-Reading and his or her gender, race/ethnicity, SES and 2003 FCAT-Reading score, data were subjected to a logistic regression analysis. The outcome measure "Success on the 2004 FCAT-Reading" was operationally defined as scoring at Level 2 or higher on that test and was treated as a dichotomous variable (success = 1, failure = 0). Four explanatory variables were entered into the model: 2003 FCAT-Reading scores, gender, race, and SES. A total of 20,001 students were included in the analysis; 617 were excluded due to missing data. Results are shown in Table 9.

Table 9

*Logistic Regression Analysis of Retained Students Achieving Success on the 2004 FCAT-Reading*

| Predictor | *B* | SEB | Wald's $\chi^2$ | *df* | *p* | $e^B$ (odds ratio) |
|---|---|---|---|---|---|---|
| Constant (SSR) | -4.760 | .1138 | 1749.9543 | 1 | <.0001 | NA |
| Gender (1 = Female) | .1217 | .0332 | 13.4127 | 1 | .0002 | 1.129 |
| Race | | | | | | |
| African American | -.2624 | .0432 | 36.8578 | 1 | <.0001 | .769 |
| Hispanic | -.0335 | .0457 | .5378 | 1 | .46 | .967 |
| Asian | -.2459 | .1855 | 1.7569 | 1 | .19 | .782 |
| SES (1 = Low SES) | -.2683 | .0485 | 30.5880 | 1 | <.0001 | .765 |
| Test | | | $\chi^2$ | *df* | *p* | |
| Overall model evaluation | | | | | | |
| Likelihood ratio test | | | 4170.8078 | 6 | <.0001 | |
| Score test | | | 3940.3166 | 6 | <.0001 | |
| Wald test | | | 2962.6594 | 6 | <.0001 | |
| Goodness-of-fit test | | | | | | |
| Hosmer & Lemeshow | | | 464.0075 | 8 | <.0001 | |

Note. Cox and Snell $R^2$ = .188. Nagelkerke $R^2$ = .256. Goodman-Kruskal Gamma = .558. *c*-statistic = 77.8%. All statistics reported herein use 4 decimal places in order to maintain statistical precision. SSR = 2003 standard score. NA = not applicable.

The results of the logistic regression analysis showed that

Predicted logit (SUCCESS) = -4.96 + 0.122*FEMALE + (-0.262*AFRICAN AMERICAN) + (-0.034*HISPANIC) + (-0.246*ASIAN) + (-0.268*LOW SES).

Based on the likelihood ratio test, the model with the four factors in the equation was found to be significantly more effective than a constant-only model, $\chi^2$ (6, N = 20,001) = 4170.81, $p < .0001$, indicating that the predictors, as a set, reliably distinguished between students who achieved success and those who did not. The score and Wald tests support this result. The Hosmer-Lemeshow goodness-of-fit test is significant but not consistent with the other tests of model fit. The strength of the prediction using Nagelkerke's $R^2$ was .256, while Cox's and Snell's $R^2$ was a more conservative .188. Measures of association are reported, which indicate the degree to which predicted probabilities agree with actual outcomes. The Goodman-Kruskal's Gamma statistic, which accounts for ties on both the outcomes and predictor variables (as are present in these data), is .558 (see Table 3). This is interpreted as 56% fewer errors made in predicting which of two students would achieve success on the FCAT-Reading by using the estimated probabilities than by chance alone. Another example is the c statistic, which for this model is .778, meaning that for 77.8% of all possible pairs of students – one successful and the other unsuccessful – the model correctly assigned a higher probability of success to the student who was successful. This indicates that the model is better at assigning outcomes than one that randomly assigns probabilities to observations. In addition to the four measures of association, a measure of classification was conducted, and results are displayed in Table 10.

Table 10

*Observed and Predicted Frequencies for Success by Logistic Regression with a Cutoff of .50*

| | Predicted | | |
|---|---|---|---|
| Observed | Successful | Unsuccessful | % Correct |
| Successful | 10950 | 1452 | 88.3 |
| Unsuccessful | 4212 | 3387 | 44.6 |
| Overall % Correct | | | 71.7 |

Note. Sensitivity = 10950/(10950+1452)% = 88.3%. Specificity = 3387/(3387+4212)% = 44.6%. False positive = 4212/(4212+10950)% = 27.8%. False negative = 1452/(1452+3387)% = 30.0%.

At a .50 probability level, the model correctly predicted 88.3% of the students achieving success, 44.6% of students not achieving success, and 71.7% of students overall. The false positive rate (27.8%) measures the proportion of observations misclassified as events while the false negative rate (30.0%) measures the proportion of observations misclassified as nonevents. The overall correction prediction was 71.7%, which is improved from chance.

The individual predictors were tested using the Wald chi-square statistic (see Table 9). Results indicated that even when controlling for previous performance, success was significantly associated with gender, race, and SES. Specifically, African American students were less likely *(B* = -.262) than their Caucasian peers to score at Level 2 or higher on the FCAT-Reading in 2004 (p<.0001). In fact, the odds of an African American scoring at Level 2 or higher were only .77 times that of the odds for a Caucasian student. In addition, students of low SES were less likely *(B* = -.266) to achieve success on the

FCAT-Reading, and had decreased odds (.77) of scoring at Level 2 or higher when compared to peers classified as Not low SES. Female students ($B$ = .122), on the other hand, were more likely to achieve success than male students; the odds of a female scoring at Level 2 or higher in 2004 were 1.13 times that of the odds for a male student ($p<.01$). Thus, results suggest that gender, race and SES, as a set, reliably distinguished between students who achieved success and those who did not. In addition, when considered individually, each of these predictors was associated with students' likelihood of success in reading performance as measured by the FCAT.

# CHAPTER FIVE

## Discussion

The pupil progression plan in the State of Florida mandates grade retention in the third grade for those students who fail to demonstrate adequate reading skills as measured by the statewide reading achievement test (i.e., FCAT-Reading). Students who do not meet this criterion but who qualify for a good cause exemption [e.g., portfolio, other norm-referenced achievement test, previous grade retention(s)] can be promoted to the fourth grade nevertheless. The present study investigated the academic outcomes of the practices of retention and promotion of low-achieving third grade students within a climate of high-stakes testing and state-mandated remediation efforts.

The educational climate in Florida is in a state of flux, as new policies and procedures have been perceived by many educators in the state as being distributed almost continuously since the enactment of the No Child Left Behind (NCLB) federal legislation and subsequent state legislation. The implementation of the policy changes has created an educational climate that is experienced by many educators, as well as students and families as being characterized by pressure and uncertainty at every level. Policy-makers are being asked to apply a major piece of national legislation to the educational system in the State of Florida; as a result of state legislation tying the FCAT to high-

stakes decisions such as retention and grades for schools, administrators are dealing with very large third-grade classrooms and are also faced with the possibility of losing access to funding on which they have counted in the past. Teachers are battling the controversy sparked by pressure that exists to "teach to the test". Finally, students as young as kindergarteners are aware of the FCAT and the ramifications they will face if they fail; many students experience considerable anxiety about being left behind in the third grade. The anxiety that has impacted the entire state has resulted in large part from the recency of the legislative changes, lack of familiarity with or uncertainty about the procedures, and the rapidity with which educators have been compelled to comply. This climate of stress may affect students' scores in these early years; however, it is possible that time will ease the impact of some or all of these factors, and these changes in Florida's educational climate may impact the academic outcomes of students who are retained due to inadequate performance on the FCAT-Reading.

*Summary of Findings*

Findings of the present study indicated that, of the retained students, 62% scored at Levels 2 or higher on the FCAT-Reading (i.e., "passed the FCAT-Reading") the subsequent year. This finding is encouraging, as it suggests that retention was associated with positive academic outcomes; after repeating the third grade curriculum and, at least in terms of mandate by policy, receiving intensive academic supports, almost two-thirds of the retainees were successful in improving their FCAT-Reading score to a passing level as defined by state standards. Upon closer examination, however, it is noted that less than half (41%) of all retained students were considered to be proficient readers (i.e., Levels 3-5 according to the FLDOE) at the end of their retention year. This number is

substantially lower than the 66% of all third-grade students who scored at Levels 2-5 in 2004 (Florida Department of Education, 2005); however, it also represents considerable improvement in this group from the previous year (i.e., all had scored at Level 1 in 2003). These findings also illustrate the lack of consistency in Florida's definitions of "adequate reading skills", which makes the interpretation of such seemingly contradictory results difficult. Students who score a Level 2 on the FCAT-Reading are demonstrating only "limited success" with the curriculum; however, according to the pupil progression plan, these same students are eligible for promotion to the fourth grade. As such, they are considered to be both "adequate" to be promoted as well as "inadequate" as readers. According to state legislation and for the purposes of this study, Levels 2-5 define success; however, it is important to keep in mind the actual limited reading skills of many of the students who have been promoted.

The findings are less equivocal, but also less positive, for students who scored at Level 1 in 2003, but were promoted due to a good cause exemption. Results show that 61% of these students failed the fourth grade version of the FCAT-Reading (i.e., they scored at Level 1) the subsequent year. Since these students took a different version of the FCAT-Reading than did the retained students (i.e., they took the fourth grade test, rather than the third grade test), the results are not directly comparable; however, these findings do indicate that within their respective curricula, the retained Level 1 students were more competent relative to the promoted Level 1 students. In addition, while 39% of Level 1 students promoted through good cause exemption "passed" the 4$^{th}$ grade FCAT-Reading at a Level 2 or higher, only 18% were considered proficient readers (i.e., scoring at Level 3 or higher). It appears that the vast majority of students promoted to

fourth grade through a good cause exemption did not meet with success with the fourth-grade curriculum, a situation predicted by their scores on the third-grade FCAT-Reading. It is important to consider the results for each individual exemption, however, as students who were promoted due to an exemption that required a demonstration of proficiency in reading (i.e., passing an alternate test, student portfolio) fared quite well in fourth grade. Of students promoted for one of the two "proficiency" exemptions, 65% "passed" the FCAT-Reading in fourth grade, scoring at a Level 2 or higher. When examined another way, results indicate the same pattern; of all "good cause exempted" students who "passed" the fourth-grade FCAT-Reading (i.e., Levels 2-5), 65% had been promoted to the fourth grade due to an exemption that indicated proficiency in reading.

Taken together, these findings indicate that there is a distinct difference between students promoted to grade four as a result of demonstrating proficiency in reading and those who were promoted for a reason unrelated to adequate academic skills (e.g., LEP, previous retentions, etc.). It would be unwise to draw conclusions about the appropriateness of promoting students through good cause exemptions without more closely examining the implications of each exemption. Allowing students to display their skills through alternative mechanisms appears to be an asset to the pupil progression plan; such exemptions provide an additional tool with which students who are likely to be successful with the fourth-grade curriculum can be identified. For students who experience test anxiety, or a diminished performance on the FCAT-Reading for any reason other than inadequate skills, these exemptions seem appropriate. The data suggest that the majority of these students possess the skills necessary to be successful in the fourth grade, and provide support for the decision to not retain them in third grade.

Students qualifying for one of the other four good cause exemptions were also promoted and experienced the fourth grade curriculum without adequate basic reading skills; as a result, they have essentially spent another year significantly behind their peers, and most likely, falling even farther behind. Though they are classified as being "promoted due to good cause," this group of students may have very different needs from those who were able to demonstrate reading competency in an alternate manner. For example, students who qualified for exemption through Limited English Proficiency (LEP) due to less than two years of instruction in an English for Speakers of Other Languages program may need extra time to master certain aspects of the basic reading curriculum (e.g., phonemic awareness, phonics) before they can be expected to progress along with their same-age peers for whom English is their primary language. It may be appropriate for policy-makers to re-examine the purpose of promoting these students and to identify where and how their academic needs would best be served.

With regard to the student variables examined, analyses revealed that gender, race, and SES are each independently and significantly ($p<.0001$) associated with repeated academic failure as measured by the FCAT-Reading. A model including all three of these variables while controlling for previous achievement indicated that retained African-American students tended to be less successful than their Caucasian peers. Retained males tended to be less successful than females, and retained students with a low SES tended to be less successful than students who were not low SES. Previous research has established that gender (Jimerson, 1999; Pagani et al., 2001), race (Jimerson, 1999; McCoy & Reynolds, 1999), and SES status (Jimerson, 1999; McCoy & Reynolds, 1999; Fergusen et al., 2001) may act as predictors of retention during the elementary

grades. Specifically, being male, a member of an ethnic minority group, and having a low SES status put children at a higher risk for grade retention. The findings from the current study indicate that these variables continue to impact academic performance during and after the retention year.

*Limitations*

One identified threat to validity in this study is related to the central measure of reading skills used both in this study, as well as in Florida's student progression plan, the FCAT-Reading. There are many issues, too many to delineate here, concerning the use of standardized tests for high stakes decisions such as grade retention. A major concern is the possibility that FCAT-Reading scores may not be a true representation of a student's actual reading skills. Certainly the results of the present study have found that for some students, particularly those who qualified for promotion through an alternate test or portfolio, the FCAT-Reading did not provide the best estimate of their reading skills as the majority managed to be quite successful in fourth grade regardless of their poor performance on the test in the third grade. Although other methods of evaluating reading skills may need to be explored, the State of Florida is currently relying heavily on the FCAT to determine reading skill level; thus, FCAT-Reading scores were used in this study as well.

Another threat to validity in this study involves treatment integrity of state-mandated intervention through Academic Improvement Plans (AIPs) for retained students. The Florida Department of Education has delineated several components of an AIP in reading, including: 1) a description of specific deficiencies in phonemic awareness, phonics, fluency, comprehension, and vocabulary, 2) concrete goals in each

area, and 3) the specific instructional and/or support services provided (Florida Department of Education, 2003g). For students who are retained in third grade due to receiving a Level 1 score on the FCAT-Reading, AIPs should be reviewed at the beginning of the retention year and again as soon as it is determined that the student may fail the FCAT-Reading again (Florida Department of Education, 2003h). This implementation plan attempts to control for the quality of each student's reading remediation; however, the plan is deficient in at least two areas that may have impacted the conclusions drawn in the present study. Although every student who receives a Level 1 score on the FCAT-Reading, regardless of retention/promotion status, should have an AIP, at this point in time, there is no method in place that would evaluate the consistency of AIPs with state requirements or to ensure that each student is actually receiving the intervention as intended. It is highly possible that students' AIPs vary widely both in quality of construction and in implementation, making statistical results difficult to interpret. It is only through detailed documentation regarding the extent to which each student received interventions as mandated that an evaluation of the effectiveness of the interventions, as well as the retention year itself, would be possible. Another issue that impacts the construction and implementation of AIPs is the variability of the remedial activities, which are defined by the state as "effective instructional practices" and "scientifically-based reading instruction." No further details are provided regarding specific activities and interventions, making it incumbent upon practitioners such as school psychologists and teachers to make these decisions according to the specific needs of each student (Florida Department of Education, 2003h). While this affords great individualization, it also implies that actual interventions may vary widely, disallowing

generalizations about their effectiveness. For the reasons mentioned above, examination of the quality of AIPs was considered to be beyond the scope of the present study and was not included.

This study was not able to control for inconsistencies in data coding or entry, factors that also threaten the validity of the findings. Data were entered in thousands of schools across the state by numerous coders using a coding system that was relatively unknown and untested. In addition, the study variable of socioeconomic status was defined using students' free or reduced lunch qualification status. These data were collected by the state and thus this category could not be defined more clearly. While the qualification status for free or reduced lunch as a definition of high or low SES is perhaps overly broad and may not offer a clear picture of students' actual socioeconomic status, it is a widely accepted method of categorization in both research and practice.

Additionally, the present study did not include a control group and the data examined were cross-sectional in nature, which limited the nature of the conclusions that could be drawn. Inclusion of a true experimental control group is not ethically possible in this type of research; however, a longitudinal design would strengthen the study by allowing conclusions to be drawn about the long-term outcomes of early grade retention.

*Delimitations*

The results may be generalized to third grade students who are retained in the State of Florida, or in a similar student progression program which requires a structured plan of remediation for every retained student. In addition to retained students, the study will include participants who performed poorly (Level 1) on the FCAT but were promoted due to one of six good cause exemptions. Results can be generalized to other

third grade students who meet these requirements. Results cannot be generalized to older or younger populations, or to students with special needs not included in the predefined good cause exemptions.

*Implications for Practice and Future Research*

Most educators and researchers concur that neither repeating the same curriculum by repeating a grade nor merely being promoted to the next grade will provide sufficient opportunities for students who are experiencing academic difficulties to experience success; the results of the present study support this position. Instead, struggling students will need consistent, intensive, and empirically-supported remediation efforts to make sufficient academic gains. Florida has instituted AIPs to address this need; however, student outcomes suggest that gaps exist in one or more areas that impede the effectiveness of AIPs, for example, in the development and monitoring of the plans, as well as continuing professional development in these areas. Findings from the present study also suggest that certain subgroups of students (e.g., males, African-Americans, students with a low SES) are at-risk for repeated failure; the AIP system and resulting interventions are not proving as effective for these subgroups. This indicates a need to make appropriate modifications that will target these at-risk students, including the development of new interventions or more frequent monitoring.

The data from this study also raise questions in terms of policy decisions that promote students scoring at Level 1 through good cause exemptions. The purpose of student promotion through good cause is arguably to help these students achieve future academic success; however, current findings indicate that only one exemption provides a viable predictor of success in the fourth grade. Students promoted through Exemption 1

(demonstration of proficiency through an alternative assessment) have a nearly 70% chance of "passing" the FCAT-Reading at Level 2 or higher in the fourth grade. Students promoted through Exemption 2 (demonstration of proficiency through portfolio) have roughly a 50% chance of passing, while students promoted through the other four categories have only a 17-31% chance of passing. The State of Florida has not defined what percentage of students promoted through a good cause exemption must pass the FCAT-Reading in the fourth grade in order for the exemption to be considered a valid reason for promotion. If the purpose of the policy were to promote only those students with a reasonable chance for success, policy-makers would be wise to reexamine students who are qualifying for promotion through Exemptions 3-6. What is contributing to their high rates of academic failure (i.e., scoring Level 1 on the FCAT-Reading in fourth grade) and what can be done to reduce those rates? If the purpose lies in something other than academic achievement (e.g., social or emotional health), further research is needed to determine whether these outcomes can be empirically supported and in turn, whether promoting students through these exemptions can be considered a valid practice.

Future research should also expand on the findings presented in this exploratory study. More information is needed regarding the student (e.g., demographics, family support, etc.) or ecological variables (e.g., teacher variables, school variables, interventions, etc.) that may have contributed to the observed variability in outcomes. The findings reported in the present study raise important questions regarding the effects of treatment integrity on outcomes (Do higher levels of treatment integrity during the retention year result in better academic outcomes?), the use of cutoff scores in determining student progression, and which students will most benefit from which types

of remediation efforts (What is the treatment utility of using a cutoff score to determine student progression, and what are the most valid selection criteria?), variability in performance according to gender, race/ethnicity, and SES (What interventions may be more beneficial for male/African American/poor students?), and longitudinal effects of the retention year (Do the academic gains demonstrated by retained students continue through subsequent grades? If so, how do those gains compare with the achievement of promoted students?).

Another important issue beyond the scope of this study is the social and emotional ramifications of being retained. Research has consistently reported the negative social and emotional effects of grade retention, such as poorer school attendance, social adjustment, and more problem behaviors (Holmes, 1989), as well as peer difficulties (Shepard & Smith, 1990). Grade retention has historically been a stigmatizing experience for many students such that even the large numbers seen in the State of Florida would not necessarily improve the emotional outcomes of being ostracized from one's peer group. Despite this, however, no studies have examined these constructs in the context of large-scale retention as part of a state-wide pupil progression plan.

*Conclusion*

Recognition of the importance of early literacy skills has resulted in a national educational climate that emphasizes reading development in students in grades K through 3. The policies in Florida have reflected this climate; its pupil progression plan includes the use of high-stakes testing and mandatory retention for third-grade students who do not display adequate progress toward the acquisition of reading skills. In addition, third-grade students who are retained in Florida are required through legislation and policy to

receive intensive academic interventions. The literature is replete with studies indicating the potentially harmful effects of grade retention on students' future academic progress; however, previous research has not included an examination of the effects of remediation efforts – such as those that exist in Florida – occurring during the retention year. The present study found that the majority of students who were retained in third grade with intensive reading interventions went on to score at Levels 2 or higher on the FCAT-Reading the following year. Most students (69% and 59%, respectively) who failed the FCAT-Reading but were promoted through a good cause exemption requiring demonstration of proficiency were successful in the fourth grade; however, those promoted through a good cause exemptions not requiring demonstration of proficiency struggled in the fourth grade despite remedial efforts; more than one-half failed the 4$^{th}$ grade FCAT-Reading the following year. Taken together, these results suggest support for retention practices; however, the cross-sectional design and lack of experimental control in this study preclude such conclusions. In addition, several risk factors for non-success the year after retention were identified, including being male, African American, and having a low SES. Future efforts should focus on learning more about these populations, as current interventions appear to be less effective with these students.

References

Alexander, K. L., Entwisle, D. R., & Dauber, S. L. (1994). *On the success of failure: A reassessment of the effects of retention in the primary grades*. New York, NY: Cambridge University Press.

Armistead, L., Kempton, T., Lynch, S., & Forhand, R. (1992). Early retention: Are there long-term beneficial effects? *Psychology in the Schools*, 29(4), 342-347.

Buntaine, R. L., & Costenbader, V. K. (1997). The effectiveness of a transitional prekindergarten program on later academic achievement. *Psychology in the Schools*, 43, 41-50.

Dennebaum, J. M., & Kulberg, J. M. (1994). Kindergarten retention and transition classrooms: Their relationship to achievement. Psychology in the Schools, 31, 5-12.

Ferguson, P., Jimerson, S. R. & Dalton, M. J. (2001). Sorting out successful failures: Exploratory analyses of factors associated with academic and behavioral outcomes. *Psychology in the Schools*, 38(4), 327-341.

Florida Department of Education. (2002). *Student Progression*. Retrieved May 3, 2005 from http://www.firn.edu/doe/commhome/progress/proghome.htm

Florida Department of Education (2002b). Technical report: For operational test Administrations of the 2000 Florida Comprehensive Assessment Test. Retrieved May 3, 2005 from http://www.firn.edu/doe/sas/fcat/pdf/fc00tech.pdf

Florida Department of Education. (2003a). No Child Left Behind: Adequate yearly progress fact sheet. Retrieved May 3, 2005 from http://www.fldoe.org/NCLB/

Florida Department of Education. (2003b). Just Read, Florida!: Reading First. Retrieved May 3, 2005 from http://www.justreadflorida.com/reading_first.asp

Florida Department of Education. (2003c). FCAT reading SSS. *The Florida Comprehensive Assessment Tes*t. Retrieved May 3, 2005 from http://www.firn.edu/doe/sas/fcat.htm

Florida Department of Education. (2003d). Third grade reading – 2003 profiles. Retrieved May 3, 2005 from http://www.firn.edu/doe/bin00014/progress/pro-file.htm

Florida Department of Education. (2003e). Good Cause Exemptions. Retrieved May 3, 2005 from http://www.firn.edu/doe/bin00014/pdf/y2004-17.pdf

Florida Department of Education. (2003f). Third Grade Progression – Requirements for students who were retained. Retrieved May 3, 2005 from ] http://www.firn.edu/doe/commhome/progress/proghome.htm

Florida Department of Education. (2003g). Public School Student Progression. Retrieved May 9, 2005 from DPS_memo_03—024—attach2.pdf

Florida Department of Education. (2003h). Technical Assistance Paper Related to Third Grade Student Progression. Retrieved May 9, 2005 from http://www.firn.edu/doe/commhome/progress/proghome.htm

Florida Department of Education (2005). Reading and Mathematics Scores – GRADE 3 Statewide Comparison for 2001 to 2005. Retrieved October 2, 2005 from http://fcat.fldoe.org/

Gottfredson, D. C., Fink, C. M., & Graham, N. (1994). Grade retention and problem behavior. *American Educational Research Journal*, 31(4), 761-784.

Holmes, C. T. (1989). Grade level retention effects: A meta-analysis of research studies. In L. A. Shepard & M. L. Smith (Eds.), *Flunking grades: Research and policies on retention* (pp. 16-33). London: Falmer Press.

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

Jimerson, S. (1999). On the failure of failure: Examining the association between early grade retention and education and employment outcomes during late adolescence. *Journal of School Psychology*, 37(3), 243-272.

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

Jimerson, S., Carlson, E., Rotert, M., Egeland, B., & Sroufe, A. (1997). A prospective, longitudinal study of the correlates and consequences of early grade retention. *Journal of School Psychology*, 35(1), 3-25.

Just Read, Florida. (2003). Read to Learn: *Intervention*. Retrieved May 3, 2005 from http://www.read-to-learn.org/intervention.html

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental scores*. Reading, MA: Addison-Wesley.

Mantzicopoulos, P. (1997). Do certain groups of children profit from early retention? A follow-up study of kindergartners with attention problems. *Psychology in the Schools*, 34, 115-127.

Mantzicopoulos, P., & Morrison, D. (1992). Kindergarten retention: Academic and behavioral outcomes through the end of second grade. *American Educational Research Journal*, 29, 182-198.

McCoy, A. R., & Reynolds, A. J. (1999). Grade retention and school performance: An extended investigation. *Journal of School Psychology*, 37(3), 273-298.

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

National Reading Panel. (2000). Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. Retrieved May 3, 2005 from http://www.nichd.nih.gov/publications/nrp/smallbook.htm

No Child Left Behind. (2002). Retrieved May 3, 2005 from http://www.ed.gov/nclb/landing.jhtml?src=pb

Pagani, L., Tremblay, R., Vitaro, F., Boulerice, B., & McDuff, P. (2001). Effects of grade retention on academic performance and behavioral development. *Development & Psychopathology*, 13(2), 297-315.

Peterson, S. E., DeGracie, J. S., & Ayabe, C. R. (1987). A longitudinal study of the effects of retention/promotion on academic achievement. *American Educational Research Journal*, 27, 107-118.

Phelps, L., Dowdell, N., Rizzo, F. G., Ehrlich, P., & Wilczenski, F. (1992). Five to ten years after placement: The long-term efficacy of retention and pre-grade transition. *Journal of Psychoeducational Assessment*, 10, 116-123.

Reynolds, A. J. (1992). Grade retention and school adjustment: An explanatory analysis. *Educational Evaluation & Policy Analysis*, 14(2), 101-121.

Simmons, R. G., & Blyth, D. A. (1987). *Moving into adolescence: The impact of pubertal change and school context*. Hawthorne, NY: Aldine de Gruyter.

The Florida Senate. (2003). The 2003 Florida Statutes: Title XLVIII; Chapter 1008. Retrieved May 3, 2005 from http://www.flsenate.gov/Statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=Ch1008/SEC25.HTM&Title=-

United States Department of Education, (2002). *Overview: Why NCLB is Important to America*. Retrieved May 3, 3005 from http://www.ed.gov/nclb/overview/importance/edlite-index.html

United States Department of Education. (2003a). Stronger accountability: Accountability for schools. Retrieved May 3, 2005 from http://www.ed.gov/nclb/accountability/schools/accountability.html

United States Department of Education. (2003b). Introduction: No Child Left Behind Retrieved May 3, 2005 from http://www.ed.gov/nclb/overview/importance/edlite-index.html

United States Department of Education. (2003c). Stronger accountability: The facts about making gains each year. Retrieved May 3, 2005 from http://www.ed.gov/nclb/accountability/ayp/yearly.html

United States Department of Education. (2003d). Proven methods: Reading. Retrieved May 3, 2005 from http://www.ed.gov/nclb/methods/reading/reading.html

Vitaro, F., Desmarais-Gervais, L., Tremblay, R. E., & Gagnon, C. (1992). Social adjustment at the beginning of elementary school and preschool day care arrangements. *Revue des Sciences de l'Education*, 18, 1-15.

Appendix A

Performance of Retained Students on the 3rd grade FCAT-Reading in 2004 by District

| | | Performance Level | | | | | |
|---|---|---|---|---|---|---|---|
| | | Level 1 | | Level 2 | | Levels 3-5 | |
| District | N | n | % | n | % | n | % |
| Alachua | 219 | 78 | 35.6 | 55 | 25.1 | 86 | 39.2 |
| Baker | 41 | 11 | 26.8 | 7 | 17.1 | 23 | 56.1 |
| Bay | 92 | 34 | 40.0 | 16 | 17.4 | 42 | 45.7 |
| Bradford | 32 | 8 | 25.0 | 6 | 18.8 | 18 | 56.3 |
| Brevard | 316 | 99 | 31.3 | 68 | 21.5 | 149 | 47.2 |
| Broward | 1972 | 821 | 41.6 | 435 | 22.1 | 716 | 36.3 |
| Calhoun | 10 | 2 | 20.0 | 1 | 10.0 | 7 | 70.0 |
| Charlotte | 28 | 5 | 17.9 | 3 | 10.7 | 20 | 71.4 |
| Citrus | 54 | 17 | 31.5 | 11 | 20.4 | 26 | 48.6 |
| Clay | 110 | 29 | 26.4 | 20 | 18.2 | 61 | 55.5 |
| Collier | 560 | 221 | 39.5 | 116 | 20.7 | 223 | 39.8 |
| Columbia | 82 | 21 | 25.6 | 17 | 20.7 | 44 | 53.7 |
| Dade | 5377 | 2173 | 40.4 | 1047 | 19.5 | 2157 | 40.1 |
| DeSoto | 44 | 21 | 47.7 | 12 | 27.3 | 11 | 25.0 |
| Dixie | 29 | 14 | 48.3 | 6 | 20.7 | 9 | 31.0 |
| Duval | 763 | 206 | 27.0 | 199 | 26.1 | 358 | 46.9 |
| Escambia | 393 | 136 | 34.6 | 95 | 24.2 | 162 | 41.2 |
| Flagler | 31 | 17 | 54.8 | 6 | 19.4 | 8 | 25.8 |
| Franklin | 2 | 0 | 0.0 | 2 | 100.0 | 0 | 0.0 |
| Gadsden | 34 | 20 | 58.8 | 6 | 17.7 | 8 | 23.5 |
| Gilchrist | 16 | 6 | 37.5 | 1 | 6.3 | 9 | 56.3 |
| Glades | 9 | 5 | 55.6 | 2 | 22.2 | 2 | 22.2 |
| Gulf | 11 | 5 | 45.5 | 3 | 27.3 | 3 | 27.3 |
| Hamilton | 34 | 18 | 52.9 | 7 | 20.6 | 9 | 26.5 |
| Hardee | 56 | 20 | 35.7 | 12 | 21.4 | 24 | 42.9 |
| Hendry | 86 | 33 | 38.4 | 20 | 23.3 | 33 | 38.4 |
| Hernando | 101 | 34 | 33.7 | 17 | 16.8 | 50 | 49.5 |
| Highlands | 84 | 36 | 42.9 | 10 | 11.9 | 38 | 45.2 |
| Hillsborough | 944 | 403 | 42.7 | 217 | 23.0 | 324 | 34.3 |
| Holmes | 32 | 11 | 34.4 | 6 | 18.8 | 15 | 46.9 |
| Indian River | 127 | 46 | 36.2 | 25 | 19.7 | 56 | 44.1 |
| Jackson | 7 | 4 | 57.1 | 1 | 14.3 | 2 | 28.6 |
| Jefferson | 11 | 6 | 54.6 | 1 | 9.1 | 4 | 36.4 |
| Lafayette | 5 | 0 | 0.0 | 1 | 20.0 | 4 | 80.0 |
| Lake | 171 | 67 | 39.2 | 38 | 22.2 | 66 | 38.6 |
| Lee | 560 | 184 | 32.9 | 104 | 18.6 | 272 | 48.6 |
| Leon | 139 | 39 | 28.1 | 46 | 33.1 | 54 | 38.9 |
| Levy | 22 | 9 | 40.9 | 4 | 18.2 | 9 | 40.9 |

Appendix A continued

| | | Performance Level | | | | | |
|---|---|---|---|---|---|---|---|
| | | Level 1 | | Level 2 | | Levels 3-5 | |
| District | N | n | % | n | % | n | % |
| Liberty | 6 | 2 | 33.3 | 0 | 0.0 | 4 | 66.7 |
| Madison | 54 | 26 | 48.2 | 12 | 22.2 | 16 | 29.6 |
| Manatee | 285 | 94 | 33.0 | 56 | 19.7 | 135 | 47.4 |
| Marion | 407 | 132 | 32.4 | 84 | 20.6 | 191 | 46.9 |
| Martin | 55 | 19 | 34.6 | 11 | 20.0 | 25 | 45.5 |
| Monroe | 32 | 13 | 40.6 | 7 | 21.9 | 12 | 37.5 |
| Nassau | 52 | 20 | 38.5 | 6 | 11.5 | 26 | 50.0 |
| Okaloosa | 131 | 45 | 34.4 | 33 | 25.2 | 53 | 40.5 |
| Okeechobee | 58 | 15 | 25.9 | 12 | 20.7 | 31 | 53.5 |
| Orange | 1554 | 599 | 38.6 | 338 | 21.8 | 617 | 39.7 |
| Osceola | 334 | 115 | 34.4 | 80 | 24.0 | 139 | 41.6 |
| Palm Beach | 1238 | 517 | 41.8 | 303 | 24.5 | 418 | 33.8 |
| Pasco | 536 | 187 | 34.9 | 104 | 19.4 | 245 | 45.7 |
| Pinellas | 853 | 329 | 38.6 | 203 | 23.8 | 321 | 37.6 |
| Polk | 803 | 285 | 35.5 | 166 | 20.7 | 352 | 43.8 |
| Putnam | 104 | 36 | 34.6 | 19 | 18.3 | 49 | 47.1 |
| St. Johns | 81 | 20 | 24.7 | 29 | 35.8 | 32 | 39.5 |
| St. Lucie | 336 | 126 | 37.5 | 78 | 23.2 | 132 | 39.3 |
| Santa Rosa | 53 | 18 | 34.0 | 12 | 22.6 | 23 | 43.4 |
| Sarasota | 102 | 41 | 40.2 | 16 | 15.7 | 45 | 44.1 |
| Seminole | 320 | 103 | 32.2 | 81 | 25.3 | 136 | 42.5 |
| Sumter | 85 | 31 | 36.5 | 16 | 18.8 | 38 | 44.7 |
| Suwannee | 69 | 22 | 31.9 | 14 | 20.3 | 33 | 47.8 |
| Taylor | 28 | 8 | 28.6 | 8 | 28.6 | 12 | 42.9 |
| Union | 6 | 3 | 50.0 | 1 | 16.7 | 2 | 33.3 |
| Volusia | 379 | 138 | 36.4 | 80 | 21.1 | 161 | 42.5 |
| Wakulla | 2 | 0 | 0.0 | 0 | 0.0 | 2 | 100.0 |
| Walton | 41 | 5 | 12.2 | 11 | 26.8 | 25 | 61.0 |
| Washington | 10 | 4 | 40.0 | 1 | 10.0 | 5 | 50.0 |

Appendix B

Level 1 Promoted Students' Performance on the 2004 4th grade FCAT-Reading by Good Cause Exemption

| | | Performance Level | | | | | |
|---|---|---|---|---|---|---|---|
| | | Level 1 | | Level 2 | | Levels 3-5 | |
| District | N | n | % | n | % | n | % |
| Alachua | 0 | 0 | 0.0 | 0 | 0.0 | 0 | 0.0 |
| Baker | 24 | 9 | 37.5 | 4 | 16.7 | 11 | 45.8 |
| Bay | 74 | 55 | 74.3 | 13 | 17.6 | 6 | 8.1 |
| Bradford | 22 | 14 | 63.6 | 6 | 27.3 | 2 | 9.1 |
| Brevard | 254 | 177 | 69.7 | 40 | 15.6 | 37 | 14.6 |
| Broward | 777 | 419 | 53.9 | 182 | 23.4 | 176 | 22.7 |
| Calhoun | 9 | 5 | 55.6 | 1 | 11.1 | 3 | 33.3 |
| Charlotte | 6 | 4 | 66.7 | 1 | 16.7 | 1 | 16.7 |
| Citrus | 48 | 30 | 62.5 | 9 | 18.8 | 9 | 18.8 |
| Clay | 56 | 39 | 69.6 | 11 | 19.6 | 6 | 10.7 |
| Collier | 130 | 74 | 56.9 | 33 | 25.4 | 23 | 17.7 |
| Columbia | 43 | 36 | 83.7 | 6 | 14.0 | 1 | 2.3 |
| Dade | 1211 | 687 | 56.7 | 255 | 21.1 | 269 | 22.2 |
| DeSoto | 27 | 18 | 66.7 | 7 | 25.9 | 2 | 7.4 |
| Dixie | 4 | 3 | 75.0 | 0 | 0.0 | 1 | 25.0 |
| Duval | 9 | 5 | 55.6 | 4 | 44.4 | 0 | 0.0 |
| Escambia | 114 | 89 | 78.1 | 16 | 14.0 | 9 | 7.9 |
| Flagler | 8 | 4 | 50.0 | 1 | 12.5 | 3 | 37.5 |
| Franklin | 0 | 0 | 0.0 | 0 | 0.0 | 0 | 0.0 |
| Gadsden | 14 | 5 | 35.7 | 7 | 50.0 | 2 | 14.3 |
| Gilchrist | 24 | 13 | 54.2 | 3 | 12.5 | 8 | 33.3 |
| Glades | 3 | 2 | 66.7 | 1 | 33.3 | 0 | 0.0 |
| Gulf | 7 | 5 | 71.4 | 0 | 0.0 | 2 | 28.6 |
| Hamilton | 3 | 3 | 100.0 | 0 | 0.0 | 0 | 0.0 |
| Hardee | 38 | 28 | 73.7 | 7 | 18.4 | 3 | 7.9 |
| Hendry | 34 | 26 | 76.5 | 4 | 11.8 | 4 | 11.8 |
| Hernando | 72 | 44 | 61.1 | 11 | 15.3 | 17 | 23.6 |
| Highlands | 28 | 23 | 82.1 | 5 | 17.9 | 0 | 0.0 |
| Hillsborough | 1465 | 821 | 56.0 | 374 | 25.5 | 270 | 18.4 |
| Holmes | 9 | 6 | 66.7 | 2 | 22.2 | 1 | 11.1 |
| Indian River | 98 | 41 | 41.8 | 27 | 27.6 | 30 | 30.6 |
| Jackson | 38 | 24 | 63.2 | 7 | 18.4 | 7 | 18.4 |
| Jefferson | 8 | 6 | 75.0 | 2 | 25.0 | 0 | 0.0 |
| Lafayette | 0 | 0 | 0.0 | 0 | 0.0 | 0 | 0.0 |
| Lake | 65 | 50 | 76.9 | 6 | 9.2 | 9 | 13.9 |
| Lee | 168 | 93 | 55.4 | 49 | 29.2 | 26 | 15.5 |
| Leon | 126 | 72 | 57.1 | 34 | 27.0 | 20 | 15.9 |

Appendix B continued

| | | Performance Level | | | | | |
|---|---|---|---|---|---|---|---|
| | | Level 1 | | Level 2 | | Levels 3-5 | |
| District | N | n | % | n | % | n | % |
| Levy | 18 | 11 | 61.1 | 4 | 22.2 | 3 | 16.7 |
| Liberty | 3 | 2 | 66.7 | 1 | 33.3 | 0 | 0.0 |
| Madison | 2 | 2 | 100.0 | 0 | 0.0 | 0 | 0.0 |
| Manatee | 174 | 112 | 64.4 | 25 | 14.4 | 37 | 21.3 |
| Marion | 151 | 88 | 58.3 | 33 | 21.9 | 30 | 19.9 |
| Martin | 30 | 12 | 40.0 | 12 | 40.0 | 6 | 20.0 |
| Monroe | 39 | 19 | 48.7 | 9 | 23.1 | 11 | 28.2 |
| Nassau | 33 | 21 | 63.6 | 6 | 18.2 | 6 | 18.2 |
| Okaloosa | 30 | 22 | 73.3 | 3 | 10.0 | 5 | 16.7 |
| Okeechobee | 58 | 35 | 60.3 | 14 | 24.1 | 9 | 15.5 |
| Orange | 716 | 412 | 57.5 | 161 | 22.5 | 143 | 20.0 |
| Osceola | 195 | 131 | 67.2 | 39 | 20.0 | 25 | 12.8 |
| Palm Beach | 243 | 107 | 44.0 | 76 | 31.3 | 60 | 24.7 |
| Pasco | 208 | 128 | 61.5 | 40 | 19.2 | 40 | 19.2 |
| Pinellas | 429 | 301 | 70.2 | 67 | 15.6 | 61 | 14.2 |
| Polk | 442 | 357 | 80.8 | 38 | 8.6 | 47 | 10.6 |
| Putnam | 56 | 45 | 80.3 | 4 | 7.1 | 7 | 12.5 |
| St. Johns | 77 | 48 | 62.3 | 12 | 15.6 | 17 | 22.1 |
| St. Lucie | 93 | 57 | 61.3 | 19 | 20.4 | 17 | 18.3 |
| Santa Rosa | 23 | 17 | 73.9 | 6 | 26.1 | 0 | 0.0 |
| Sarasota | 79 | 49 | 62.0 | 17 | 21.5 | 13 | 16.5 |
| Seminole | 180 | 122 | 67.8 | 28 | 15.6 | 30 | 16.7 |
| Sumter | 36 | 16 | 44.4 | 10 | 27.8 | 10 | 27.8 |
| Suwannee | 17 | 12 | 70.6 | 0 | 0.0 | 5 | 29.4 |
| Taylor | 5 | 5 | 100.0 | 0 | 0.0 | 0 | 0.0 |
| Union | 14 | 10 | 71.4 | 4 | 28.6 | 0 | 0.0 |
| Volusia | 289 | 192 | 66.4 | 51 | 17.7 | 46 | 15.9 |
| Wakulla | 17 | 14 | 82.4 | 1 | 5.9 | 2 | 11.8 |
| Walton | 21 | 17 | 81.0 | 1 | 4.8 | 3 | 14.3 |
| Washington | 5 | 4 | 80.0 | 0 | 0.0 | 1 | 20.0 |