New Directions in Education Research:
Using Data Mining Techniques to Explore Predictors of Grade Retention

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Deanna Kelley-Winstead
Master of Education
George Mason University, 2005
Bachelor of Arts
George Mason University, 2003
Associate of Arts
Northern Virginia Community College, 2001

Director: Clifton D. Sutton, Associate Professor
Department of Statistics

Spring Semester 2010
George Mason University
Fairfax, VA

Dedication

This is dedicated to each person, place, and bit of knowledge that has shaped me into the woman I am today. I especially want to thank the God of my understanding and my wonderful parents who are the reason I'm alive today. To my loving husband, in-laws, and our growing family I thank you from the bottom of my heart for your unconditional love, support and encouragement.

Acknowledgements

I would like to thank the many colleagues who have been supportive of my educational endeavors. Drs. Clifton Sutton, Anastasia Samaras, Thomas Scruggs, Margo Mastropieri, Dimiter Dimitrov, Earle Reybold, Anastasia Kitsantas, Panagiota Kitsantas, Gary Galluzzo, Frederick Brigham, and other faculty members at George Mason who provided invaluable support throughout my academic career. Above all I thank my higher power and the inspirational experiences I have encountered which have guided me through this journey.

Table of Contents

List of Tables

List of Figures

Abstract

NEW DIRECTIONS IN EDUCATION RESEARCH: USING DATA MINING TECHNIQUES TO EXPLORE PREDICTORS OF GRADE RETENTION

Deanna Kelley-Winstead, Ph.D.

George Mason University, 2010

Dissertation Director: Dr. Clifton D. Sutton

The purpose of this study was to use classification trees and logistic regression to identify subgroups of students more likely to be retained. The National Educational Longitudinal Study of 1988 (NELS:88) was used to identify the sociodemographic, family background and school related factors associated with grade retention. The sample size for this study consisted of 10,140 students, 1,570 of which had been held back. The NELS data were obtained from student questionnaires and surveys with the students' parents, teachers, and school administrators. In order to identify the predictors of students more likely to be held back, models were built using classification trees and logistic regression. Overall, the current study identified the predictive factors of grade retention. Moreover, this study demonstrates the effectiveness of using classification trees in conjunction with stepwise logistic regression in educational research.

1. Introduction

History of Grade Retention

Grade retention is the practice of requiring a student to repeat an entire grade level. Historically, grade retention has been known as the practice of requiring a student who has been in a given grade level for an entire academic year, to remain in the same grade level for another year. Grade retention also refers to delayed entry of a child into kindergarten or first grade. The term "grade retention" is used synonymously with nonpromotion, held back, repeating, and flunking. Although retention can also be referred to as a time to grow, from the perceptive of the retained student, these terms all mean failure (Bocks, 1977; Jackson, 1975).

Historical information pertaining to grade retention dates back to the beginning of the 20th century. Soon after grade promotion was standardized in the early-1900s, concerns began to emerge about the potential negative consequences student retention may have on student intellectual and social development. This concern led to the start of a social promotion policy in the 1930s (Sandoval & Fitzgerald, 1985). From the 1930s to the 1960s, low-achieving students were socially promoted to the next grade to remain with their peers. Although the new policy was widely used, grade retention remained prevalent in some parts of the country (Goodland, 1954; Livingston, 1959).

During the 1960s, the social promotion policy was challenged by educators who contested that declines in student achievement test scores were attributable to social promotion and more lenient academic standards (Rose, Medway, Cantrell & Marus, 1983). Growing criticism among the public gave rise to the back-to-basics movement in the 1970s. This movement focused on promoting students who had achieved grade level mastery (measured by minimum competency tests). This resulted in the adoption of minimum competency tests in the majority of states.

From the 1980s to the present a movement has started that is progressing away from traditional intelligence quotient (IQ)-Achievement testing to determine student eligibility for special education services. Appropriate identification procedures have been a concern ever "since the inclusion of learning disabilities (LD) as a category in Education for All Handicapped Children Act" (Scruggs & Mastropieri, 2002). One of the main problems that exists regarding student identification for special education services is overidentification. An influx of students has been identified with LD in the past two-decades through the use of IQ-Achievement discrepancy. However, since 2005 the implementation of responsiveness-to-intervention has been pursued by the education system in response to changes in the Individuals with Disabilities Education Improvement Act (IDEA) and No Child Left Behind (NCLB; Staskowski & Rivera, 2005). During the 1980s retention rates were estimated at 2.3 million a year. From 1979-1980 1.2 million students were identified as LD and from 1998-1999, 2.8 million students were identified as LD (Vaughn & Fuchs, 2003).

In Chapter 2, numerous research studies that have been conducted to better understand the factors that contribute to grade retention are discussed. These studies explore sociodemographic, family and school related predictors of grade retention. In addition, studies that have explored multiple variables across these three groups were assessed to provide an in-depth review of the impact of retention as an intervention.

Statement of the Problem

Although prior research has been conducted in the area of grade retention, a comprehensive study that takes all previously identified predictors of grade retention into account had not been conducted prior to this study. In order to assess the predictors of grade retention this study employed two quantitative methodologies: classification trees and logistic regression. Logistic regression is commonly used by educational research methodologists to explore the predictors of grade retention, whereas tree-based models have been successfully employed in the health sciences (Raymond, Tafari, Troendle & Clemens, 1994; Zhang & Bracken, 1995; 1996; Zhang & Singer, 1999), in the investigation of medical decision making (Harper, 1996), birth outcomes (Kitsantas, Hollander & Li, 2006; Raymond et al., 1994; Sims et al., 2000; Zhang & Bracken, 1995; 1996), and several other medical and epidemiological studies (Kitsantas, Moore & Sly, 2007; Marshall, 2001; Toschke, Beyerlein & von Kries, 2005; Werneck, De Carvalho, Barroso, Cook & Walker, 1999). In addition, classification trees have been applied to evaluate the motor vehicle industry (Kuhnert, Do, & McClure, 2000) as well as the functionality of new inventions (Buratti, Benedetti, Scampicchio & Pangerod, 2004).

Classification trees and logistic regression work well in conjunction with one another for a number of reasons. First of all both methodologies can identify predictive variables and the percentage of students in each predicted group. The sensibility and accuracy of models can also be assessed by classification trees and logistic regression. Therefore, in the present study, classification trees and logistic regression were used to develop prediction rules that identify subgroups of students who are more likely than average to be held back.

Furthermore, classification trees are known for having several advantages over logistic regression. First, they automatically deal with missing values which leaves more data to work with thus allowing predictions to be made using a larger proportion of cases. Secondly, tree-based models provide a valuable tool for uncovering the hidden interactive structure of variables. Additional benefits of classification trees are discussed in Chapter 3 which also includes a detailed overview of stepwise logistic regression.

Classification trees were selected to be used in conjunction with stepwise logistic regression for a number of reasons. In addition, to the successful application of both methodologies in previous research, and the advantages tree-based models have over logistic regression, classification trees have not been used to predict grade retention. Therefore, this study is intended to provide additional information and insights into the predictors of grade retention by analyzing the previously identified predictors of grade retention using these two predictive methods (classification trees and logistic regression). The overarching purpose of this study is to use classification trees in conjunction with

logistic regression in order to identify subgroups of students more likely than average to be retained.

Significance of the Study

The significance of this study is in its content and methodology. This study contributes to the general research on grade retention through a specific focus on the predictive factors of grade retention, as well as the methodological application of classification trees in educational research. The information gained from this study will encourage educational research methodologists to employ classification trees in conjunction with stepwise logistic regression. Applying both methods will enable educational research methodologists to explore the predictive nature of variables that were otherwise limited, due to the inability of logistic regression to handle missing values. This study was also designed to address the issue of grade retention for students who were ever held back or were retained due to a parental request or school request. In turn, the findings from this study identified the sociodemographic, family background, and school related factors of students who are more likely than average to be retained.

Definitions of Terms

1. Retention: In this study, retention is defined as repeating an entire grade level for a full academic year. The term most often used in this study to indicate retention is *held back*.

2. Social Promotion: This is the practice of promoting a student with their peers without having demonstrated satisfactory academic progress or mastery of the grade-level content (Rummel, 2007).
3. Race and ethnicity: Terminology for race and ethnicity is taken from the National Educational Longitudinal Study (NELS): Asian Pacific Islander, Hispanic, Black (non-Hispanic), White (non-Hispanic), and American Indian.

2. Literature Review

This chapter provides an overview of grade retention research dating back to the early 1900s. This overview includes factors related to grade retention which are grouped by sociodemographic, family background, and school related factors. Multiple variables across these three groups are also discussed in this chapter regarding the impacts of grade retention on a student's academic achievement, sociopsychological adjustment, peer relationships, conduct, and self-concept.

Factors Related to Grade Retention

Researchers have come to identify multiple variables related to grade retention. These variables represent characteristics that are present prior to a child being retained. The variables are grouped by sociodemographic factors, family background, and school related factors. Sociodemographic variables include race and gender. Family background includes parent's marital status, level of education and involvement. School related variables include student academic achievement, as well as the school's retention policy to remediate poor school performance. This section outlines each of these factors, which have been found in grade retention research over the past forty years. A list of the factors related to grade retention along with research studies can also be found in Table 1. Variable definitions are provided in Chapter 3.

Sociodemographic Factors

As indicated in Table 1, studies that have looked at the demographic characteristics of retained students have found that gender and race are related to student retention. Numerous studies have most often found that African American males are more likely to be held back (Abidin, Jr., Golladay, & Howerton, 1971; Alexander, Entwisle, & Dauber, 2003; Gottfredson, Fink, & Graham, 1994; Graue & DiPerna, 2000; Jimerson & Kauffman, 2003; McArthur & Bianchi, 1993; McCoy & Reynolds, 1999; NCES, 2006; Thomas & Knudsen, 1965; Zill, Loomis, & West, 1998). The methods used in these studies to analyze predictors of grade retention include logistic regression, meta analysis, factor analysis, and hierarchical multiple regression. In 1971, Abidin, Golladay, and Howerton studied 85 students that were retained during first or second grade, in comparison with 43 continuously promoted students. Of the students retained, the majority were male (70%), more than half of which were African American (54%). Four years later Jackson (1975) discovered that minority students were three or four times more likely to be held back than white students. In 2004, the National Center for Education Statistics (NCES) found similar results in a nationwide study; a greater percentage of males (13%) than females (6%) and African Americans (16%) than Whites (8%) had been retained (NCES, 2006). These findings are also consistent with the National Household Education Surveys Program (NHES) data which found that African American males were retained more often than any other subgroup (McArthur & Bianchi, 1993; Zill, Loomis, & West, 1998).

Studies have also been conducted that have found only gender *or* race to be predictors of grade retention. These studies offer an alternative explanation to the previously stated research findings. A study conducted among first through fourth graders in Baltimore inner-city schools, found that gender is a predictor of grade retention (Dauber, Alexander, & Entwisle, 1993). However, a statistical significance was not present for race, which the researchers attribute to the overrepresentation of race, socio-economic status and parent dropout rates. Cosden, Zimmer, and Tuss (1993) studied the impact of age, gender, and ethnicity on the retention rates of Latino children. The study found statistically significant results for the overrepresentation of Latino males in one of the three districts of Southern California. However, this was the district where Latinos were in the minority rather than the majority. This explanation could explain why the previously mentioned study by Dauber, Alexander, and Entwisle (1993) did not find race to be a predictor of grade retention in inner-city Baltimore which is predominately African American. In addition to sociodemographic factors, there are other predictors of grade retention associated with a student's family background.

Family background

Family background is believed to contribute to the likelihood of a child being retained in school. Family factors that have been found to contribute to grade retention include parent education level, parent marital status, and parent involvement. Other family factors include socio-economic status, high rates of home transiency, language proficiency and school mobility.

Low socio-economic status, in particular, influences grade retention (Alexander, Entwisle, & Dauber, 2003; Alexander, Entwisle, & Horsey, 1997; Dauber, Alexander, & Entwisle, 1993; Fernandez, Paulsen, & Hirano-Nikanishi, 1989; Grant, 1997; Janosz, LeBlanc, Boulerice, & Tremblay, 1997; Jimerson, 1999; McCoy & Reynolds, 1999; Meisels & Liaw, 1993; Rumberger, 1995; Rumberger & Larson, 1998). In 2004, the National Center for Education Statistics (NCES) found that children from low-income households (16.9%) were more likely to be retained than children from middle and high income homes (NCES, 2006). Furthermore, studies conducted on parents living in poverty that were high school dropouts found that their children were more likely to be retained and uninvolved in school activities (Alexander, Entwisle, & Dauber, 2003; Grant, 1997; McCoy & Reynolds, 1999).

Additionally, low socio-economic status often carries the burden of an impermanent place to live. High rates of home transiency increase a student's risk of grade retention especially when students have to frequently change schools (Alexander, Entwisle, & Dauber, 2003; Grant, 1997; McArthur & Bianchi, 1993; McCoy & Reynolds, 1999; Powell, 2005; Stringer, 1960). Powell (2005) interviewed ten adults who were retained during elementary school about their family life prior to retention. Half of the participants shared that their family moved frequently during that time in their lives or moved right before they were retained.

Multiple family factors pertain to the students' parents including, parent's level of education and academic support parents provide to their children. McCoy and Reynolds (1999) conducted a follow-up of the predictors of grade retention using the same sample

in Reynolds' original study conducted in 1992. The sample consisted of 93% of the original participants including 1,164 low-income, Black (95%) and Hispanic (5%) children from the Chicago Longitudinal Study. Twenty-eight percent of the sample was retained in first through eighth grade. Results indicated that family factors that place children at risk for grade retention include low parent educational attainment (i.e., failure to graduate from high school) and low parental participation in school activities. Each of these family related factors was used in this study in conjunction with the previously mentioned sociodemographic factors and the following school related factors.

School Related Factors

Multiple factors related to the school environment are believed to contribute to student grade retention. These school related factors include academic achievement, socialization, absenteeism, and self concept. In addition, school retention policies and the special instructional programs provided to students at-risk of repeating a grade also influence grade retention (Janosz, LeBlanc, Boulerice, & Tremblay, 1997; Jimerson, Egeland, & Teo, 1999; Roderick, 1994; Rumberger, 1995).

Academic achievement is commonly discussed when determining the reason for student retention, including the progress of retained students before and after retention. Previous research centered on kindergarten and first-grade academic achievement led to the discovery that students who demonstrate poor test performance during the first grading period are more likely to be retained at the end of the school year (Dauber, Alexander, & Entwisle, 1993; McCoy & Reynolds, 1999). In 1992, Reynolds studied the

impacts of early grade retention on reading and math achievement. He found that low reading ability and low math scores were predictors of retention. Findings also indicated that students are more likely to be held back based on low socio-economic status, low parental involvement, and high school mobility.

In addition to these risk factors, special instruction programs are known to help prevent grade retention. For instance, early childhood education has been found to reduce a child's likelihood of being retained. Preschool programs such as Head Start[1] (McArthur & Bianchi, 1993) found that minority males who did not attend a structured preschool program (e.g., Head Start) were more likely to be retained. Researchers recognize the limitations in McArthur and Bianchi's statistical analysis of the NHES data, seeing that the connection to preschool was linked only to minority populations. Similar NHES data collected two years later in 1995 provided researchers with a more general conclusion. Subsequent research indicates that students who attended Head Start or a similar program, regardless of race, were less likely to be retained (Zill et al., 1998). Despite academic gains, students who attend Head Start are still behind the cognitive levels of their peers, making little impact on future academic success (Lee, Brooks-Gunn, & Schnur, 1988). According to Reynolds (1992), students who attend Head Start are more likely to be held back.

[1] Head Start is administered by the Office of Head Start (OHS), Administration for Children and Families (ACF), Department of Health and Human Services (HHS). The program is designed to prepare children from birth to age five, who comes from families with incomes below the poverty level, for kindergarten.

Impacts of Grade Retention

In addition to the previously mentioned studies pertaining to sociodemographic, family background, and school related factors, other studies have looked at multiple variables across these three groups. The subsections that follow provide an in-depth review of the literature related to grade retention factors, including the impact of retention as an intervention.

Impact on Academic Achievement

In regards to academic achievement, several studies have either compared the academic progress of retained students to their promoted peers or have examined the progress of retained students before and after retention. Some studies have found that retention was ineffective as an educational intervention. However, one study conducted in 1954 found the opposite to be true. Coffield (1954) matched 147 pairs of students. The first group consisted of seventh graders who had been retained at least once in elementary school, and the comparison group consisted of students who were never held back. Findings from this study revealed that although retained students demonstrated short-term improvement after retention, they performed at the same level as their peers in seventh grade. Although retention did not have a long-term effect on ensuring the mastery of school subject matter for the retained students, retention allowed students the opportunity to catch up with their peers academically.

Similarly, Skelton (as cited in Holmes, 1989) studied 34 children who had repeated the second grade between 1957 and 1961. Students were matched on IQ,

chronological age, mental age, and the Stanford Achievement Test (SAT). Despite being a year older and having one more year of education, held back students scored significantly lower than their continuously promoted peers. This is one of the first studies to challenge the notion that being held back provides an opportunity for low-achieving students to academically catch up with their peers.

Dobbs and Neville (1967) also found evidence that repeating a year's work does not assure students will overcome a deficiency in academic achievement. This study evaluated the effect of retention on the school performance of first graders. This was done by matching 30 pairs, each consisting of one retained first grade student and one continuously promoted second grade student. Students were matched on race, gender, age, mental age, socio-economic status, type of classroom assignment, and reading ability. Two years later students were tested using the Metropolitan Achievement Test. Findings revealed that held back students scored significantly lower in reading and mathematics than their continuously promoted peers. As a result, the investigators concluded that the objective of retention was not met. Researchers speculated that a feeling of success associated with continuous promotion may have been a motivating factor for increases in academic achievement.

An examination of standardized achievement tests revealed that the scores of held back students decreased during the years following retention. Abidin, Golladay and Howerton (1971) conducted a longitudinal study comparing retained students to students with low academic achievement who were promoted. This study consisted of 85 sixth grade students retained during first or second grade who were compared to 43 promoted

first grade students who performed below the 25th percentile on the Metropolitan Readiness Test. Findings indicated that by the end of sixth grade promoted students scored at or above grade level compared to retained students who scored below grade level. The investigators concluded that decreases in achievement and intelligence continued for held back students.

Godfrey (1972) conducted a study with 1,200 sixth and seventh grade students from 14 schools to determine if retention helped a student catch up academically. Findings based on reading ability and math scores found that promoted students were nearly two grade levels above those who were retained once, and about two and a half grade levels above those who were retained twice. These findings confirmed prior research concluding that retention does not help a child catch up academically. Godfrey recommended alternatives to grade retention including individualized instruction, peer tutoring, and summer school.

Ogden (1971) conducted an evaluation examining the academic and behavioral performance of 100 high school students who had been held back one or more times during elementary school. Held back students were compared to a randomly selected group of students who had been retained, as well as, socially promoted with various academic problems. Findings indicated that the academic problems continued in high school for half of the held back students. Ogden also found that held back students performed no better academically than at-risk students who were socially promoted.

Subsequent research supports Ogden's findings. Sandoval and Fitzgerald (1985) conducted a follow-up study with three groups of high school students. They compared

three groups of students to one another (a) 30 retained students, (b) 32 students who participated in an intervention program, and (c) 75 randomly selected continuously promoted students. Findings indicated that the held back students performed significantly lower academically than the other two groups. Also, the use of intervention programs was found to effectively address the needs of low-achieving students.

Shephard and Smith (1987) examined the impact of nonpromotion on the academic progress of kindergarteners. Forty held back and 40 continuously promoted first grade students were matched based on gender, age, initial school readiness, socio-economic status and language ability. They found no significant academic differences between the groups concluding that retention of kindergarteners was ineffective regardless of whether children had been retained due to developmental immaturity or insufficient academic achievement.

A study conducted by Owens and Ranick (1977) initiated much debate because it lent support to the belief that retention can be an effective educational intervention. In response to declining standardized achievement test scores, Owens and Ranick established The Greenville (Virginia) Program with one of the strictest promotion policies in the United States based on minimum competency education. The policy stated that "No students would be promoted until they showed, on achievement tests, the mastery of the skills for their grades" (p. 531). Students had to repeat a grade if they did not perform at grade level. Students were held back based on the premise that more time was needed in order to master particular skills. Retained students were then placed in special classes which enabled them to master most of the material required for their grade

and thus receive partial promotion or placement in a transitional grade. Results of this program have not only increased achievement tests and IQ scores but have also decreased retention and drop-out rates.

The results of the Greenville Program were questioned by Koons (1977) who cited some possible explanations for their findings; the first being that high achievement scores can be attributed to the Hawthorne effect which is due to changes in treatment rather than its quality. Other explanations include that teachers may have taught to the test and the strict promotion policy may have forced students to concentrate on passing the achievement tests in order to achieve high scores. Although other studies have found that having a strict promotion standard may improve overall scores, they also acknowledge that this is done at the expense of at-risk students. Koons noted that although some students may benefit from retention, a greater number are not helped and are actually harmed. Furthermore, Koons recommended that schools "should be made to fit the students, not the student fit the school" (p. 701).

Overall, the majority of previous research has found that retention and social promotion do not address the problems of low-achieving students. In particular, social promotion has produced many negative outcomes (Reiter, 1973). According to prior research the solutions may lie in alternatives to retention, such as tutoring, individualized instruction, summer school, after-school and at-home assistance, as well as changes in teaching methods and school structure. Reiter (1973) also noted that the answer lies in how students are treated at home and in school post retention or promotion.

Impact on Sociopsychological Adjustment

Although a large number of studies have been conducted on the academic achievement of retained students, studies have also explored the impacts of retention on peer relationships, behavior, self-concept, and school dropout rates. As early as 1941, Anfinson conducted a study comparing 58 held back and 58 continuously promoted junior high school students. Participants were matched based on school attendance, age, gender, IQ, and socio-economic status. Findings indicated that continuously promoted students scored significantly higher than those retained on social and personal adjustment.

Goodlad (1954) reached a similar conclusion in his study of the social and personal adjustment of 150 held back and continuously promoted students attending 11 elementary schools. Participants were matched on mental age, chronological age, and academic achievement. The California Test of Personality was used to discriminate between held back and continuously promoted students. The study found that continuously promoted students were significantly better adjusted socially and personally compared to retained students. Goodlad concluded that grade retention has detrimental effects on a students' social development.

Subsequent research has also revealed positive outcomes regarding retention. For example, Chansky (1964) compared scores from the California Test of Personality across a group of students held back in the first grade and a group of at-risk students who were socially promoted. Findings indicated that there were no significant differences between the groups on personal adjustment, nor was there a significant relationship between grade

retention and personal adjustment. Similarly, Briggs (1968) administered the California Test of Personality and Pupils Adjustment Inventory to 30 continuously promoted students and 29 fifth and sixth grade males who had been retained twice. The comparison of their scores indicated no significant differences between the groups.

Reinherz and Griffin (1970) conducted a study consisting of 57 retained males. Their goal was to determine the relationship between grade retention and immaturity. Findings indicated that retention was most beneficial for immature male students in the early grades. In regards to measuring the maturity levels of retained students, Chafe (1984) argued that it is impossible to know whether impacts are made in the classroom due to retention or the fact that the student is a year older when they repeat a grade.

Impact on Peer Relationships

Another important factor in a child's social development is peer relationships. Given that grade retention can create a negative peer reaction toward a retained student, Sandin (1944) decided to study the social and emotional adjustment of continuously promoted and retained students. Students were measured using not only sociometric rating scales but also classroom observations, interviews, and by reviewing student records. Findings indicated that retained students did not choose companions who were younger or smaller than themselves. Instead they chose to have relationships with students a grade above them who was their own age. Sandlin also found that as retained students entered upper grades they were discriminated against by their peers who no longer chose them to work on class projects. Goodlad (1954) confirmed these findings

noting that retained students faced significantly more rejection by classmates than continuously promoted students.

Similarly, Ashbury (1975) studied 118 students who were progressing slowly to determine the impact of peer rejection. Findings indicated that these students were chosen significantly less often by classmates than were continuously promoted students. In addition, struggling students scored lower on personality adjustment measures. Ashbury concluded that low achievement could contribute to personality maladjustment.

Plummer and Graziano (1987) examined the relationship between retention and social development by interviewing 219 retained and continuously promoted second and fifth grade students. The selection of partners for social activities and academic tasks was assessed using a rating scale. Findings indicated that retained students were discriminated against by those who were continuously promoted, especially in upper elementary grades. These results were moderated by grade level, gender, age, and the height of the retained student. Similarly, Morrison and Perry (1956) also found that discrimination against retained students was more evident in upper elementary grades.

Although the previously mentioned studies pertaining to peer relationships report varying degrees of discrimination against retained students, they also suggest that peer relationships strongly influence school adjustment. Furthermore, retention can cause feelings of rejection which may lead to negative social attitudes. Plummer, Lineberger, and Graziano (1986) suggested that experiencing frequent rejection or not receiving social approval is detrimental to the academic and social development of retained

students. Overall, the majority of the studies reviewed regarding peer relationships found that retention has negative effects on social and personal adjustment.

Impact on Conduct

Previous research has also focused on the effects of retention practices on student behavior. Jimerson, Carlson, Rotert, Egeland and Sroufe (1997) found that disruptive behavior plays a role in academic-related problems that are conducive with grade retention. Similarly, Reynolds (1992) found that student academic achievement and social behavior weigh into retention decisions. Caplan (1973) studied 50 continuously promoted and retained students that had the same demographic factors (e.g., age, gender, race, and grade level). Findings indicated that males have more learning problems than females. However, aggressive females earned significantly lower grades than their male counterparts because their teachers underestimated their ability level. Therefore, aggressive female students were more likely to be retained than girls who conformed to traditional gender norms. Caplan concluded that classroom behavior is a critical factor in the decision to promote or retain females. Similarly, Jackson (1975) found that students from minority groups were more often perceived by teachers and school administrators as having maladaptive behavior and/or learning problems. Briggs (1968) found similar results, concluding that retained students were more withdrawn and aggressive than their continuously promoted classmates.

Safer (1986) examined the relationship between retention and conduct and found differences between the consequences of retention in elementary versus middle school.

Findings indicated that retention led to negative behavior once a student entered high school. Safer noted a weakness in the studies ability to clearly understand the effects of retention on grade level. Therefore, he recommended further research be conducted to assess the relationship between grade level and personal adjustment. Overall, studies examining the impact of retention on student conduct found that retention may promote behavioral problems.

Impact on Self-Concept

Understanding the impact of grade retention on a student's self-concept has generated considerable debate. Previous research has found that retention in elementary school may lead to a negative perception of one's self. White and Howard (1973) studied 624 sixth grade students who were continuously promoted or previously retained one or more times. Findings indicated that continuously promoted students scored higher on the Tennessee Self-Concept test than retained students. Furthermore, students who were retained more than once had a lower self-concept than students who were retained only once. Numerous other studies have reached the same conclusion that retention in elementary school leads to a lower self-concept than their continuously promoted peers (Bedoian, 1954; Godfrey 1972; Johnson 1968; Morrison & Perry, 1956).

Although the aforementioned studies found lower self-concept among retained students than those who were continuously promoted, alternative findings have also been found. Hains (1981) examined 29 retained and 24 socially promoted randomly selected students in third, fourth and fifth grade to determine the effects of grade retention on self-

concept**.** Findings indicated that no significant difference existed between student self-concept scores. These results are in agreement with Ammons study from 1976 that used the Piers-Harris Children's Self Concept Scale to determine that after controlling for grade level, gender, race, IQ, and age there were no significant differences between the self-concept of retained and continuously promoted students. Ammons concluded that retention did not negatively impact the self-concept of students.

Plummer and Graziano (1987) found the opposite was true when it came to the effect of retention on self-concept. Their examination of 219 retained and continuously promoted students revealed that retention had positive effects on a student's self-concept. These findings were discovered using the Katz and Ziglar (1967) self-image disparity approach which found that continuously promoted students scored lower on self-concept measures than retained students. Katz and Ziglar speculated several possible explanations for these results including biases in previous research, positive effects of retention, and positive treatment by peers especially for retained students attending schools with a high ratio of student retention. The authors concluded that a high retention rate may contribute to increased social acceptance of retained students, thus improving their self-concept.

Although the above research addresses the effects of grade retention on self-concept, it does not establish a cause and effect relationship. Finlayson (1977) employed a repeated measures design to research the relationship between retention and self-concept during two academic years. Each year student self-concept was measured four times. During the second year, the study included 25 retained, 25 continuously

promoted, and 25 at-risk students. Findings indicated that following retention, self-concept scores steadily increased while the scores of at-risk and continuously promoted students remained stable. At the conclusion of the two year study, identical self-concept scores existed among retained and continuously promoted students. Therefore, the author concluded that retention did not negatively impact the self-concept of retained students.

Overall, the studies regarding the impact of retention on a student's self-concept, have found mixed results. Some studies indicated that retention is harmful, whereas others found it to enhance a student's self-concept, but still others found that no impact was made. However, it should be noted that measurements of self-concept are known to be unstable due to systematic response bias, response restriction, contextual effects and social desirability (Wylie, 1974).

Summary

Grade retention is a serious issue for all parties involved in retention decisions including parents and their students as well as school administrators and teachers. Students most likely to be held back meet certain demographic characteristics and/or family and school related factors. Prior research demonstrates that students who are held back have one or more of these factors present in their lives. Previous research has identified a number of predictors of grade retention that are categorized in this study by sociodemographic, family and school related factors, however, research to date has not examined all these predictors simultaneously, nor have classification trees previously been applied to predict grade retention. The purpose of this investigation, therefore, is to

use classification trees and logistic regression to identify subgroups of students more likely than average to be retained.

Table 1

Key References: Factors Related to Grade Retention

References	Gender	Race	Highly transient homes	Language proficiency	Parent education level	Parent participation	School mobility	Single parent home	Socio-economic status	Absenteeism	Academic achievement	Behavior	Remediation policy	Self-concept	Socialization
Abidin, Jr., Golladay, & Howerton (1971)	▲	▲								▲	▲				▲
Alexander, Entwisle, & Dauber (2003)	▲	▲	▲					▲	▲	▲	▲	▲			▲
Alexander, Entwisle, & Horsey (1997)	▲	▲							▲	▲		▲			
Brooks-Gunn, Guo, & Furstenberg (1993)						▲						▲			
Cadigan et al. (1988)										▲	▲				▲
Caplan (1973)	▲	▲													
Cosden, Zimmer, & Tuss (1993)	▲	▲													
Dauber, Alexander, & Entwisle (1993)	▲	▲			▲				▲		▲	▲			▲
Fernandez, Paulsen, & Hirano-Nikanishi (1989)		▲							▲						
Gottfredson, Fink, & Graham (1994)	▲										▲				
Grant (1997)			▲			▲		▲	▲						
Graue & DiPerna (2000)	▲	▲									▲				
Hayes (2005)				▲											
Holmes & Matthew's (1984)														▲	
Jackson (1975)	▲	▲													
Janosz et al. (1997)					▲				▲				▲		
Jimerson (1999)									▲	▲					
Jimerson (2001)										▲	▲	▲			▲
Jimerson, Egeland, & Teo (1999)													▲		
Jimerson & Kaufman (2003)	▲	▲													
McArthur & Bianchi (1993)	▲	▲	▲		▲						▲				
McCoy & Reynolds (1999)	▲		▲		▲	▲	▲		▲		▲				
Meisels & Liaw (1993)	▲	▲			▲				▲		▲				
Morris (1993)															▲
Morrison & On No (2007)				▲											
Plummer & Graziano (1987)														▲	
Powell (2005)			▲												
Reynolds (1992)	▲	▲			▲				▲		▲	▲			
Roderick (1994)		▲					▲						▲		
Rumberger (1995)		▲				▲	▲		▲	▲			▲		
Rumberger & Larson (1998)		▲					▲		▲	▲		▲			
Stroup & Robins (1972)					▲		▲			▲		▲			
Thomas & Knudsen (1965)	▲	▲													
Zill, Loomis, & West (1998)	▲	▲													

3. Method

This study employed two quantitative methodologies: classification trees and logistic regression. Classification trees and logistic regression were used to develop prediction rules that identify subgroups of students who are more likely to be retained. Data in a test set were used to evaluate the predictive accuracy of each of the models. This section describes each methodology as well as the sample, data source, instruments, and procedures.

Methodological Overview

Classification trees (Breiman, Friedman, Olshen & Stone, 1984) are known for providing a valuable tool for uncovering the hidden interactive structure of variables which are well established to be associated with grade retention. Classification trees also effectively solve common data analysis issues that emerge such as collinearity (Kitsantas, Moore & Sly, 2007) and missing data (Kitsantas, Moore & Sly, 2007; Sims et al., 2000; Sutton, 2005). Classification and Regression Tree (CART) software (Salford Systems, 2000) can create a diagnostic tool that serves as an early warning system (Harper, 1996; Kitsantas, Hollander & Li, 2006; Kitsantas, Moore & Sly, 2007; Marshall, 2001; Werneck, De Carvalho, Barroso, Cook & Walker, 1999) that is used for detecting the latent interactive structures including risk and protective factors. In the present study,

prediction rules that identify subgroups of students more likely to be retained were developed using CART.

CART is known to have several advantages over other classification methods such as logistic regression (Kitsantas, Hollander & Li, 2006). First, CART deals effectively with large data sets and issues of dimensionality (Brieman et al., 1984; Kitsantas, Moore & Sly, 2007). In addition, CART can effectively handle outliers, missing data and collinearities. Another common classification method is discriminant analysis which has a major weakness in that it assumes all predictor variables are normally distributed and are therefore continuous. However, CART makes no such distributional assumptions (i.e., normality) about the predictors and can handle both continuous and categorical predictors. One drawback to CART is its instability, which means that small changes in the data can result in different variables being selected. This in turn can change the structure of the tree, however the classification accuracy of the model can be very similar.

In the first step of tree construction the entire data set is split into various subsamples using binary splits based on a chosen splitting criterion, such as the measure of node purity called the Gini index (Breiman et al., 1984), with each split creating nodes of the tree. The parent node appears at the top of the tree, corresponding to the entire data set prior to the first binary split. An algorithm is used at each junction to search for the best binary split to produce nodes with the greatest decrease in node impurity. Binary splits occur until defined stopping rules for each child node are fulfilled, at which point a child node becomes a terminal node at the end of each branch. The terminal nodes of the

tree can be based on a different combination of variables and CART can estimate the classification accuracy for each terminal node. The basic structure of classification trees is illustrated in Figure 1.

It is important to note that the initially grown trees can be large, complex and difficult to interpret. The pruning process collapses tree branches of the initial tree. Cross-validation is used in the pruning process to measure the predictive accuracy of each tree in a sequence of pruned trees. Once the "right-sized" tree is determined by minimizing the cross-validation estimate of overall accuracy, the tree-building process is complete.

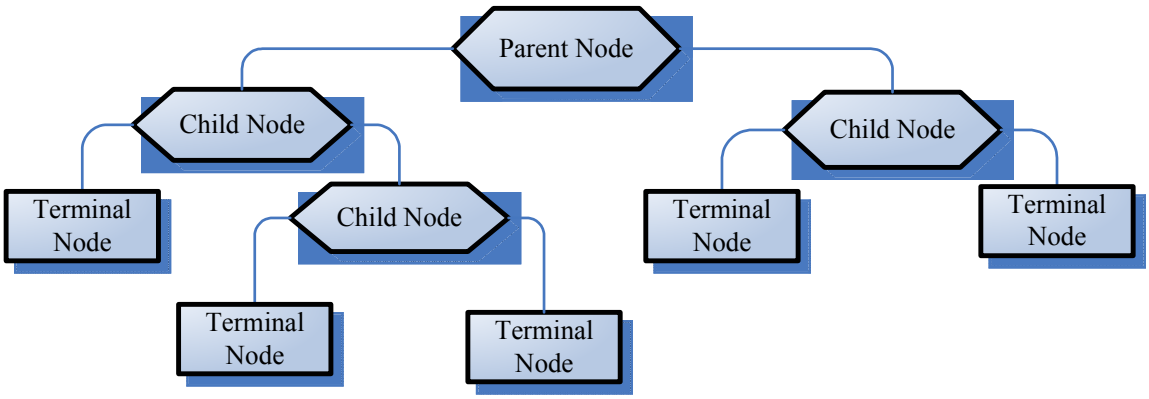

Figure 1. Sample Classification Tree

Classification trees are often used in conjunction with logistic regression. Multiple studies have found little difference between the performances of logistic regression models and tree-structured classifiers (Ennis, Hinton, Naylor, Revow, & Tibshirani, 1998; Kitsantas, Hollander, & Li, 2006; Kitsantas, Moore, & Sly, 2007; Long,

Griffith, Selker, & D'Agostino, 1993; Nelson, Bloch, Longstreth, & Shi, 1998; Sims et al., 2000; Werneck et al., 1999). Studies have also found that classification trees perform comparably to logistic regression, identifying similar risk factors (Kitsantas, Hollander, & Li, 2006; Sims, et al., 2000). Kitsantas, Hollander, and Li (2006) concluded that the use of one of these methods complements the other.

Logistic regression is most commonly used to model the relationship between a binary outcome variable and a set of predictors (Hosmer & Lemeshow, 1999). It is also used with nominal or ordinal outcome variables. Like classification tress, logistic regression can be used to classify data involving both categorical and continuous predictor variables. The goal of a logistic regression model is to estimate the probability of an event. This is accomplished by understanding the role of a vector of "input" variables (**x**) in explaining the "outcome" variable (Y) in the model, which can have two or more classes. Specifically, in binary logistic regression, letting p = P(Y=1), it is assumed that

$$\log(p / (1-p)) = f(\mathbf{x}),$$

where f(**x**) is a linear function of the predictors. The method of maximum likelihood is used to estimate the coefficients of f(**x**), and statistical tests are used to determine which variables should be included in the model. For example, the Wald Chi-Square statistic can be used to test for the statistical significance of each predictor, given the presence of the other predictors in the model.

With the backward elimination stepwise method, all of the variables are initially included in the model. Then the least significant variable is removed and the significance

of the other variables is reevaluated based on the new model. Once again the least significant variable is removed. This process is continued until only statistically significant variables remain in the model. Stepwise logistic regression using forward selection is similar. The process starts with no variables in the model and at each step the most significant variable is added. The process stops when none of the variables not yet added to the model are statistically significant given the presence of the variables already in the model.

Sample

This study is a secondary analysis of data from the National Educational Longitudinal Study of 1988 (NELS:88) which provides data from 1988 to 2000. Data from NELS were employed to provide a nationally representative sample. All students in the sample (*N*=10,140) were in eighth-grade at the start of data collection in the spring of 1988. Fifteen percent (1,570) of the students in this sample were previously retained in grade school. Characteristics of the sample are outlined in the next chapter.

Data Source

The NELS is a twelve-year project designed and funded by the National Center for Education Statistics (NCES). NELS consists of four main data collection components: student data collection, and questionnaires administered to parents, teachers, and school administrators. As illustrated in Figure 2, for this study the data were segmented in order to construct three classification models: *ever held back* (due to any

reason), held back due to a *parental request* and held back due to a *school request*. Each model also includes students who were not held back.

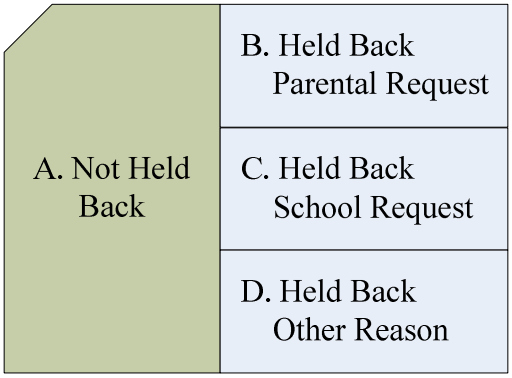

Figure 2. Data Segments used for the Various Models

The first classification model, ever held back, includes all students who were and were not held back (A+B+C+D). The second classification model, parental request, consists of those held back due to a parental request and students not held back (A+B). Similarly, the third model, school request, consists of students not held back and those students held back due to the request of the school (A+C). Students held back due to both parental and school requests were classified as being part of the school request data segment. This determination was made given that although a parental request may have been made, it is ultimately the decision of the school to retain a student.

Instruments

Instruments were developed for the NELS to collect data from students, parents, teachers, and school administrators. Questionnaires were administered during base year

with respondents who were then resurveyed through four follow-ups. During the base year of data collection (1988), trend data were collected about the critical transformations eighth grade students experience as they transition from middle or junior high school to high school. In 1990, the first set of follow-up questionnaires was administered, capturing data on students who dropped out before the end of 10th grade and monitoring the transition of students into secondary schooling. The second follow-up occurred in 1992, providing a culminating measurement of learning in the course of secondary school. This period of data collection also gathered information that will allow an investigation of the transitions from secondary school to the work force and postsecondary education. During the third follow-up in 1994, data were provided addressing the number of dropouts that have returned to school. The fourth follow-up was outside of the scope of this study given its focus on the current accomplishments of students in the workforce. Table 2 lists each wave of data collection used in this study.

Table 2

Data Collection Instruments by Wave, Year, and Respondent

Wave[a]	Year[b]	Respondent[c]
Base Year	1988	Student
Base Year	1988	Parent
First Follow-up	1990	Student
First Follow-up	1990	Teacher
First Follow-up	1990	School Administrator
Second Follow-up	1992	Student
Second Follow-up	1992	Parent
Third Follow-up	1994	Student

[a]Wave of data collection.
[b]Year of data collection.
[c]Type of data collection respondent.

Scoring

NELS questionnaire scoring is both normative and criterion-referenced. NELS scores include achievement quartiles and proficiency scores which are most fully described in appendix H (pp. H-31— H-38) of the *NELS:88 Second Follow-Up Student Component Data File User's Manual* (Ingels et al., 1994, NCES 94-374). The psychometric properties of the NELS scores are described in the *Psychometric Report for the NELS:88 Base Year Through Second Follow-Up* (Rock & Pollack, 1995, NCES 95-382). The psychometric report includes information about test reliability, construct validity and test specifications of the NELS content areas.

The reliabilities provided in the psychometric report for NELS:88 are based on weighted data. The reliabilities of internal consistency measures are based on coefficient Alpha. High coefficient Alpha reliabilities (eighties and above) suggest that the reliability tests are relatively unifactoral. To assess the internal consistency reliability of NELS:88, four content areas (reading, math, science and history) were tested. Test results demonstrated relatively high internal consistency reliability within each content area. It is important to note that higher reliabilities were found in reading (.84 overall coefficient alpha) and mathematics tests (.90 overall coefficient alpha).[2] Similarly, factor analytic results supported the discriminant validity of the four tested content areas.

[2] Science and history were not identified as predictors of student grade retention and were therefore not used in the present study.

Procedures

As previously mentioned, this study is a secondary analysis of data from the NELS. Data analysis started with recoding grade retention into a binary categorical variable. Children retained were coded 1, and those continuously promoted were coded 0. Next, the remaining variables were recoded. The following is a comprehensive list of the variables used in this study to analyze grade retention, including a breakdown of their characteristics and codes. The variables are grouped by grade retention, demographics, family background, and school related factors. In addition, the Appendix provides another look at the details pertaining to the nature of the variables (e.g., categorical), type of respondents, and wave of data collection.

- **Grade Retention**
    - ***Ever held back.*** Students that have ever been held back a grade in school (during kindergarten through eighth grade) identified themselves during the second follow-up. Retained students were coded as 1, and continuously promoted students were coded 0.
    - ***Parental request.*** Students held back during kindergarten through eighth grade due to a parental request were identified by parents during the base year of data collection. Retained students were coded as 1, and continuously promoted students were coded 0.
    - ***School request.*** Students held back during kindergarten through eighth grade due to a school request (and possibly a parental request as well)

were identified by parents during the base year of data collection. Retained students were coded 1, and continuously promoted students were coded 0.

- **Demographics**
  - ***Gender.*** Gender of student was self-identified during the third follow-up. Males were coded as 1, and females were coded 2.
  - ***Race/ethnic backgrounds.*** Race of student was self-identified during the third follow-up. Asian/Pacific Islanders were coded as 1, Hispanics were coded as 2, Blacks (not Hispanic) were coded as 3, Whites (not Hispanic) were coded as 4, and Native Americans were coded as 5.
- **Family Background**
  - ***Parent level of education.*** Mother and father separately reported whether they graduated college (coded 1), did not finish high school or earned a high school diploma (coded 0), during base year. Parents (or guardians) reported whether they graduated college (coded 1), did not finish high school or earned a high school diploma (coded 0), during the second follow-up survey.
  - ***Parent marital status***. Parents (or guardians) reported their marital status during the second follow-up. Parents who were single/never married were coded as 1, married were coded 2, divorced/separated were coded 3, widowed were coded 4, and living like married were coded 5.

- ***Socio-economic status.*** This family-level factor indicates low-income families coded as 1, and high-income families coded 2. Parents reported their total gross family income during the second follow-up survey.
- ***Highly transient homes.*** Students whose families moved within the past two years were coded as 1, whereas students who did not move were coded 2. Students reported this information during the first follow-up survey.
- ***School mobility***. Students reported the number of times they changed schools during the second follow-up. No school changes were coded as 0, and one or more school changes were coded 1.
- ***Language proficiency***. Students self-reported participation in an English language/language assistance program, if their native language was not English during base year. Students who participated in a language assistance program were coded as 1, and no participation was coded 2.
- ***Parent involvement.*** Teachers reported the level of parental involvement in their child's academic performance during the first follow-up. Parents with no influence on the academic performance of their child were coded as 0, somewhat or very involved parents were coded 1.

- **School Related**
  - ***Absenteeism.*** Students self-reported how often they cut or skipped classes during base year. Students absent from class never or almost never were

coded 0, less than once a week was coded 1, and at least once a week or daily was coded 2.

- ***Academic achievement.*** Achievement quartiles for standardized test scores, reading and mathematics were based on composite scores from the cognitive test battery administered during the first and second follow-up. Students performing below average in these subjects were coded as 1, whereas average or above scores were coded 2.
- ***Head start.*** Parents (or guardians) reported whether or not their child attended Head Start during base year. Students who attended the Head Start program were coded as 1. Those who did not participate in the program were coded 2.
- ***Suspensions.*** Students self-reported how often they were suspended or put on probation from school during the second follow-up survey. Students never suspended were coded as 0, one to two (1-2) were coded 1, three to nine (3-9) were coded 2, and ten or more were coded 3.
- ***School transfer.*** Students self-reported how many times they were transferred to another school for disciplinary reasons during the second follow-up survey. Students never transferred were coded as 0, one to two (1-2) were coded 1, three to nine (3-9) were coded 2, and ten or more were coded 3.
- ***Socialization with females.*** Students self-reported during the first follow-up if they made friends easily with girls. False, mostly false, and more

false than true responses were coded as 2. True, mostly true, and more true than false responses were coded 1.

- ***Socialization with males.*** Students self-reported during the first follow-up if they made friends easily with boys (coded 2) or not (coded 1).
- ***Social activity.*** Students self-reported during the first follow-up whether or not they think their peers saw them as being very or somewhat socially active (coded 1), or not at all (coded 0).
- ***Popularity with the opposite sex.*** Students self-reported during the first follow-up if they were very popular with the opposite sex (coded 1) or not (coded 2).
- ***Popularity among peers.*** Students self-reported during the first follow-up whether or not they think their peers saw them as them as being very or somewhat popular (coded 1), or not at all (coded 0).
- ***School promotes parental involvement.*** School administrators reported how much emphasis the school placed on promoting parental support and involvement during the first follow-up survey. Minor to no emphasis on promoting parental support was coded as 1, and major emphasis was coded 2.
- ***School encourages parental involvement in policy decisions.*** School administrators reported encouraging parents to be involved in policy decisions during the first follow-up survey. If this took place never,

seldom, or sometimes it was coded as 0. If it occurred usually or always it was coded 1.

- ***School remediation policy.*** School administrators reported on whether or not the school had a remediation policy during the first follow-up survey. Schools requiring students to maintain a minimum grade point average in order to participate in school activities such as organized team sports or theatrical productions were coded as 1. Schools without such requirements were coded 2.

Building and Testing the Models

For the purposes of this study the data segments available for each of the three models were broken into two parts, one to build the model and one to evaluate the model. A randomly selected test sample consisting of nine percent of the data (n=920) were used for the evaluation of the ever be held back model. Table 3 shows the breakdown of the data used to build and test each model.[3] The test sample was used to identify how many misclassification errors were present when the data were plugged into the fitted models. This leads to unbiased estimates of each models' predictive accuracy. The decision to select nine percent of the cases was made after estimating the predictive accuracy using test samples consisting of five, six, seven, eight, nine and ten percent of the data. Although having a larger test set means less data to build the model, this tradeoff is necessary because a test set is needed to provide an unbiased assessment of classification

[3] The data were rounded to the nearest ten as required by the U.S. Department of Education Institute of Education Sciences for restricted-use data.

accuracy. The most accurate predicting models were produced using nine percent of the cases.

Table 3

Data Used to Build and Test Each Model

Model	*Model Building Data* n_1[d]	*Test Data* n_2[e]
Ever Held Back[a]	9,220	920
Parental Request[b]	8,050	790
School Request[c]	8,510	850

[a]The ever held back model includes all participants in the sample.
[b]Only includes students held back due to a parental request or never held back.
[c]Only includes students held back due to a school request or never held back.
[d]Rounded number of students included to build each model.
[e]Rounded number of students included to test each model.

Model Building

Once all variables were recoded and the data were segmented to build and test each model, data analysis began by using CART to create classification trees. All variables were available for inclusion in the models. The cross-validation option was selected for tree construction, and the test sample was used to evaluate the predictive performance of each classifier.

The construction of classification trees was followed by logistic regression which was done using SPSS 17.0 (a comprehensive software program used for statistical data analysis that is a registered trademark of SPSS, Inc.). Predictors were entered into models using both backward elimination and forward selection stepwise logistic regression. First, all variables were included for possible selection in the models, as was

the case in the CART analysis; however, the iterative logistic regression methods terminated and were unable to find a final solution for five of the six runs. Because logistic regression cannot handle missing values when all of the variables are used, the sample size becomes too small (because all cases having any missing values are completely ignored when logistic regression is run using SPSS) and this can cause convergence to not occur.

Since SPSS's logistic regression routine ignores cases having any missing values, resulting in a large proportion of the data sets for the various models being not used, an alternative strategy is to ignore variables having a large proportion of missing values, since this will result in fewer cases being eliminated. Three different methods for reducing the number of variables were used. One method was to omit variables having more than 20 percent of the values missing or unknown. This resulted in not using three variables: language assistance, Head Start, and parent level of involvement. This filtering of the data still left a lot of cases not being usable, because they had one or more missing values for other variables, and so a further filtering of variables for which all variables with more than 10% missing values being ignored was also done. This criterion resulted in seven additional variables (marital status, school mobility, father and mother highest levels of education, school remediation policy, school promotion of parental involvement, and parental involvement in policy decisions) being excluded from the logistic regression modeling, but increased the number of usable cases (because fewer missing values remained). One other way of filtering the data was also considered: only variables used by CART were included (since these variables were found to be useful for

classification, and the other variables were not found to be useful by CART). So, altogether, for each of the three models (ever held back, parental request, and school request), four different filterings of the data (using all variables, using only variables with 20% or more missing values, using only variables with 10% or more missing values, and using only variables identified by CART as being useful for classification) were tried with both the forward selection and backwards elimination methods of stepwise logistic regression.

Model Evaluation

Each of CART's terminal nodes correspond to different portions of the student population. Some of these are groups of students more likely to be held back and some are groups of students less likely to be held back, with more likely and less likely being relative to the overall held back rate for the various subsets of the data being considered for the various models.[4] The model is judged on the sensibility of the nodes. To evaluate the sensibility of CART's terminal nodes, the CART model can be used to make predictions on the test set cases. Cases predicted to be more likely to be held back should contain a greater proportion of students who were actually held back than those cases which were predicted to be less likely to be held back. The accuracy of logistic regression models can be assessed by estimating the probability of being held back for each of the test set cases from the fitted logistic regression models. The estimated probabilities can be compared to the overall held back rate to identify whether a student

[4] The overall held back rates are 0.15, 0.03, and 0.08 for the ever held back, parental request, and school request models, respectively.

is more likely or less likely than average to be held back and the goodness of the logistic regression models can be judged by whether students in the data sets predicted to be more likely to be held back are in fact held back at an appreciably higher rate than the students not predicted to be more likely to be held back. The strongest predictors of grade retention will be identified by the parameter estimates (beta coefficients) and the smallest *p* values in each stepwise logistic regression model.

4. Results

This chapter presents the results of the study beginning with a description of the sample categorized by sociodemographic, family and school related factors. This is followed by a description of the classification tree results of each of the three models (ever held back, parental request and school request). Next the results of forward stepwise logistic regression are presented for each model which includes parameter estimates, as well as Wald test statistic values and *p* values for each predictor. Findings reported for both classification trees and logistic regression include the sensibility and accuracy of each model. These findings are followed by a comparison of the variables and cases used for each model, as well as a comparison of classification tree and stepwise logistic regression models. This comparison includes the percentage of students actually held back in each of the predicted groups and the variables identified across models.

Characteristics of the Sample

This section presents the risk and protective factors that were considered in the construction of classification trees and logistic regression models. The sample consisted of 10,140 participants.[5] Approximately 15.5% of the sample was classified as ever held

[5] The NELS data was collected at multiple intervals, using various survey instruments, resulting in different completion rates and missing values for non-responses. The data were also rounded to the nearest ten as required by the U.S. Department of Education Institute of Education Sciences for restricted-use data.

back. Table 4 presents the reasons students are ever held back. Of those held back, 17.2% were classified as held back due to a parental request. Students held back due to both parental and school requests or just school request, 50.3% altogether, were classified in the school request model. The remaining 32.5% of the held back students were held back for unspecified reasons.

Table 4

Reasons Students are Ever Held Back

Rationale	*n*[d]	%[e]
Parental Request[a]	270	17.2
School Request[b]	790	50.3
Unspecified[c]	510	32.5
Total	1570	100.0

[a]Students held back due to a parental request.
[b]Students held back due to a school request.
[c]Students held back for unspecified reasons.
[d]Rounded number of participants held back.
[e]Percentage of held back students in each category.

Sociodemographic Factors

The sociodemographic characteristics of the sample consist of the ethnic and gender distribution of participants. Table 5 presents the ethnic distribution of the sample. Native Americans were most likely to be held back. Although Native Americans only account for 1.2% of the overall sample, 25.0% were held back (See Table 1). This is almost twice the overall percentage (15.5%) for the entire sample. African Americans (24.7%) were the racial group second most likely to be held back, followed by Hispanics (21.3%), whites (13.7%), and Asians (8.9%).

Table 5

Ethnic Breakdown of the Total Sample

	Overall		*Held Back*	
Ethnicity	*n*[a]	%[b]	*n*[c]	%[d]
Asian/Pacific Islander	670	6.6	60	9.0
Hispanic	1220	12.0	260	21.3
Black (non-Hispanic)	850	8.4	210	24.7
White (non-Hispanic)	7220	71.1	990	13.7
American Indian	120	1.2	30	25.0
Unknown	60	0.6	20	33.3
Total	10140	100.0	1570	15.5

Note. From the NELS third follow-up, student questionnaire, 1994.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each ethnic group.

Table 6 presents the gender distribution of all participants, as well as specific information about those held back of each gender. Overall, there were fewer females (12.2%) than males (19.3%) among those held back, as well as those held back specifically due to a parental or school request. Likewise the male retention rate (19.3%) is higher than the overall retention rate for the sample (15.5%).

Table 6

Gender Statistics for the Total Sample

	Overall		*Held Back*	
Ethnicity	*n*[a]	%[b]	*n*[c]	%[d]
Male	4660	46.0	900	19.3
Female	5420	53.5	660	12.2
Unknown	60	0.6	10	16.7
Total	10140	100.0	1570	15.5

Note. From the NELS third follow-up, student questionnaire, 1994.
[a]The number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]The number of participants held back.
[d]The percentage of participants held back for each gender.

Family Background Factors

The tables below illustrate the family background descriptive statistics of the overall sample. Table 7 presents the percentage of students from highly transient homes, (defined to be students whose families moved within the past two years). The majority of students (82.9%) were not from highly transient homes. The percentage of retained students from highly transient homes (17.8%) was higher than the overall average (15.5%).

Table 7

Highly Transient Home among the Total Sample

	Overall		*Held Back*	
Highly Transient Home	*n*[a]	%[b]	*n*[c]	%[d]
Applies	1460	14.4	260	17.8
Does not apply	8410	82.9	1160	13.8
Unknown	270	2.7	150	55.5
Total	10140	100.0	1570	15.5

Note. From the NELS first follow-up, student questionnaire, 1990.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Table 8 presents the percentage of students who attended a language assistance program. The percentage of held back students who attended a language assistance program (19.4%) is higher than the overall average of ever held back students in the sample (15.5%). Similarly, the held back rate for students who did not attend a language assistance program is also higher than the overall average the overall average. It is also important to note the majority of students (80.9%) did not respond to this question.

Table 8

Language Assistance Provided to the Total Sample

	Overall		*Held Back*	
Language Assistance Program	*n*[a]	%[b]	*n*[c]	%[d]
Yes	310	3.0	60	19.4
No	1630	16.1	270	16.6
Unknown	8200	80.9	1240	15.1
Total	10140	100.0	1570	15.5

Note. From the NELS base year data collection, student questionnaire, 1988.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Tables 9, 10 and 11 address parental levels of education. The highest held back rates for parents' level of education were among students whose parents did not finish high school or earn a high school diploma.

Table 9

Parents' Highest Level of Education for the Total Sample

	Overall		*Held Back*	
Parents' Highest Level of Education	*n*[a]	%[b]	*n*[c]	%[d]
Dropout or H.S. diploma	6460	63.7	1190	18.4
College Graduate	2870	28.3	230	8.0
Unknown	810	8.0	150	18.5
Total	10140	100.0	1570	15.5

Note. From the NELS second follow-up, parent questionnaire, 1992.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Table 10

Mother's Level of Education for the Total Sample

	Overall		*Held Back*	
Mother's Level of Education	*n*[a]	%[b]	*n*[c]	%[d]
Dropout or H.S. diploma	6460	63.7	1040	16.1
College Graduate	2280	22.5	200	8.8
Unknown	1400	13.8	330	23.6
Total	10140	100.0	1570	15.5

Note. From the NELS base year data collection, parent questionnaire, 1988.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Table 11

Father's Level of Education for the Total Sample

	Overall		*Held Back*	
Father's Level of Education	*n*[a]	%[b]	*n*[c]	%[d]
Dropout or H.S. diploma	5720	56.4	950	16.6
College Graduate	2710	26.7	240	8.9
Unknown	1710	16.9	380	22.2
Total	10140	100.0	1570	15.5

Note. From the NELS base year data collection, parent questionnaire, 1988.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

The percentage of parental involvement reported by teachers is presented in Table 12. Although the majority of teachers (64.0%) did not respond to this question, most who responded reported that parents of their students are somewhat or very involved in the lives of their children. The percentage of held back students whose parents were uninvolved (25.0%) is more than one and a half times the overall average (15.5%). While only 27.7% of the parents were reported as being somewhat or very involved, that corresponds to 77.0% of the cases for which the level of involvement is known.

Table 12

Parents' Level of Involvement for the Total Sample

	Overall		*Held Back*	
Parents' Level of Involvement	*n*[a]	%[b]	*n*[c]	%[d]
Not involved	840	8.3	210	25.0
Somewhat/very involved	2810	27.7	330	11.7
Unknown	6490	64.0	1030	15.9
Total	10140	100.0	1570	15.5

Note. From the NELS first follow-up, teacher questionnaire, 1990.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Table 13 presents the data pertaining to school mobility. The majority of students reported that they did not change schools (68.8%). However the retention rates for school mobility were highest among students who changed schools one or more times in the past two years (15.7%).

Table 13

School Mobility for the Total Sample

	Overall		*Held Back*	
School Mobility	*n*[a]	%[b]	*n*[c]	%[d]
No school changes	6980	68.8	770	11.0
1 or more school changes	1340	13.2	210	15.7
Unknown	1820	18.0	590	32.4
Total	10140	100.0	1570	15.5

Note. From the NELS second follow-up, student questionnaire, 1992.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Table 14 presents the data pertaining to parents' marital status. The majority (68.4%) of students come from families whose parents are married, and the retention rates were lowest among this group. The highest retention rates were among students from homes with a widowed parent (22.2%), a single parent (20.0%), parents who are living like married (20.0%), or are divorced (19.0%).

Table 14

Marital Status of Parents from the Total Sample

	Overall		*Held Back*	
Marital Status of Parents	*n*[a]	%[b]	*n*[c]	%[d]
Single, never married	150	1.5	30	20.0
Married	6940	68.4	910	13.1
Divorced/separated	1050	10.4	200	19.0
Widowed	180	1.8	40	22.2
Living like married	100	1.0	20	20.0
Unknown	1710	16.9	370	21.6
Total	10140	100.0	1570	15.5

Note. From the NELS second follow-up, parent questionnaire, 1992.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Parents reported the socio-economic status of their families as low (53.5%) or high (45.8%), as presented in Table 15. The higher held back rate (22.4%) is among students from low income households.

Table 15

Socio-economic Status of the Total Sample

	Overall		*Held Back*	
Socio-economic Status	*n*[a]	%[b]	*n*[c]	%[d]
Low	4640	45.8	1040	22.4
High	5430	53.5	510	9.4
Unknown	70	0.7	20	28.6
Total	10140	100.0	1570	15.5

Note. From the NELS second follow-up, parent questionnaire, 1992.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

School Related Factors

The tables below illustrate the school descriptive statistics of the overall sample. Table 16 presents data pertaining to student absenteeism. The majority of students were never or almost never absent from class (89.9%). The percentage of held back students who were absent at least once a week or daily (30.8%) is nearly two times the overall average (15.5%). Similarly, the held back rate for students who missed class less than once a week (20.4%) is also higher than the overall average.

Table 16

Absenteeism for the Total Sample

	Overall		*Held Back*	
Absenteeism (skipped class)	*n*[a]	%[b]	*n*[c]	%[d]
Never or almost never	9120	89.9	1270	13.9
Less than once a week	540	5.3	110	20.4
At least once a week or daily	130	1.3	40	30.8
Unknown	350	3.5	150	42.9
Total	10140	100.0	1570	15.5

Note. From the NELS base year data collection, student questionnaire, 1988.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Table 17 presents the number of participants who attended the Head Start preschool program (8.4%). The highest held back rates were among students who attended Head Start (23.5%) which is one and a half times the overall average (15.5%).

Table 17

Head Start Attendees from the Total Sample

	Overall		*Held Back*	
Attended Head Start	*n*[a]	%[b]	*n*[c]	%[d]
Yes	850	8.4	200	23.5
No	5900	58.2	780	13.2
Unknown	3390	33.4	590	17.4
Total	10140	100.0	1570	15.5

Note. From the NELS base year data collection, parent questionnaire, 1988.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Cognitive test batteries were administered to students during the first and second follow-up. Table 18 indicates that nearly half of the sample (42.9%) has below average standardized test scores, and these students were held back at the higher rate (28.0% versus 6.0%).

Table 18

Standardized Test Scores for the Total Sample

	Overall		*Held Back*	
Standardized Test Score	*n*[a]	%[b]	*n*[c]	%[d]
Below average	4350	42.9	1220	28.0
Average or above	5790	57.1	350	6.0
Total	10140	100.0	1570	15.5

Note. From the NELS first and second follow-up, cognitive test battery, 1990 and 1992.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Similarly, nearly half of student reading ability scores (43.6%) are below average, as indicated in Table 19, and these students were held back at the higher rate (26.2% versus 7.2%). Table 20 indicates that nearly half of the sample (42.7%) have below average scores in mathematics and these students were held back at the higher rate (27.9% versus 6.2%).

Table 19

Reading Ability of the Total Sample

	Overall		*Held Back*	
Reading Ability	*n*[a]	%[b]	*n*[c]	%[d]
Below average	4420	43.6	1160	26.2
Average or above	5700	56.2	410	7.2
Unknown	20	0.2	0	0.0
Total	10140	100.0	1570	15.5

Note. From the NELS first and second follow-up, cognitive test battery, 1990 and 1992.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Table 20

Math Scores for the Total Sample

	Overall		*Held Back*	
Math Score	*n*[a]	%[b]	*n*[c]	%[d]
Below average	4330	42.7	1210	27.9
Average or above	5810	57.3	360	6.2
Total	10140	100.0	1570	15.5

Note. From the NELS first and second follow-up, cognitive test battery, 1990 and 1992.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Table 21 presents data regarding student suspensions. The majority of students were never suspended (87.2%), whereas a little more than four percent were known to have been suspended from school at some point during their academic career. The percentage of held back students generally increases according to the number of suspensions, with the highest rate (50.0%) being among those suspended 10 or more times, and the lowest rate (12.1%) being for those who were never suspended.

Table 21

Suspensions for the Total Sample

	Overall		*Held Back*	
Suspended	*n*[a]	%[b]	*n*[c]	%[d]
Never	8840	87.2	1070	12.1
1-2 Times	340	3.3	80	23.5
3-9 Times	60	0.6	10	16.7
10 or More Times	20	0.2	10	50.0
Unknown	880	8.7	400	45.5
Total	10140	100.0	1570	15.5

Note. From the NELS second follow-up, student questionnaire, 1992.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Table 22 presents data regarding student transfers for disciplinary reasons. Less than one percent of students were known to have been transferred to another school during their academic career. Students transferred one or two times were more than one and a half times more likely to be held back than the overall average (25.0% versus 15.5%).

Table 22

Participants Transferred for Disciplinary Reasons

	Overall		*Held Back*	
Transferred for Disciplinary Reasons	*n*[a]	%[b]	*n*[c]	%[d]
Never	9200	90.7	1160	12.6
1-2 Times	40	0.4	10	25.0
3-9 Times	10	0.1	0	0.0
10 or More Times	10	0.1	0	0.0
Unknown	880	8.7	400	45.5
Total	10140	100.0	1570	15.5

Note. From the NELS second follow-up, student questionnaire, 1992.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

The majority of schools (69.4%) have a policy requiring students to meet a minimum grade point average in order to participate in school activities such as organized team sports or theatrical productions, as shown in Table 23. Interestingly both the "yes" and "no" retention rates (14.2% and 11.9%) are appreciably less than the retention rate for the unknown category.

Table 23

Remediation Policies of Student's Schools

	Overall		*Held Back*	
School Remediation Policy	*n*[a]	%[b]	*n*[c]	%[d]
Yes	7040	69.4	1000	14.2
No	1180	11.6	140	11.9
Unknown	1920	19.0	430	22.4
Total	10140	100.0	1570	15.5

Note. From the NELS first follow-up, school administrator questionnaire, 1990.
[a]Rounded number of participants whose schools have retention policies.
[b]The percentage of participants whose schools have retention policies.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Students' social skills were self-assessed during the first follow-up. The majority of students (83.8%) reported making friends easily with girls, as indicated in Table 24. However the retention rate (18.8%) is higher among students who did not make friends easily with girls compared to students who reported making friends easily with girls.

Table 24

Makes Friends Easily with Girls

	Overall		*Held Back*	
Makes Friends Easily with Girls	*n*[a]	%[b]	*n*[c]	%[d]
False	1170	11.5	220	18.8
True	8500	83.8	1140	13.4
Unknown	470	4.6	210	44.7
Total	10140	100.0	1570	15.5

Note. From the NELS first follow-up, student questionnaire, 1990.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

The majority of students (85.3%) reported making friends easily with boys, as indicated in Table 25. However the retention rate (17.5%) is higher among students who did not make friends easily with boys compared to students who reported making friends easily with boys.

Table 25

Makes Friends Easily with Boys

	Overall		*Held Back*	
Makes Friends Easily with Boys	*n*[a]	%[b]	*n*[c]	%[d]
False	970	9.6	170	17.5
True	8650	85.3	1180	13.6
Unknown	520	5.1	220	42.3
Total	10140	100.0	1570	15.5

Note. From the NELS first follow-up, student questionnaire, 1990.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

The majority of students (74.1%) did not indicate a problem with their popularity with the opposite sex, however a fifth of students (20.8%) did, as indicated in Table 26. Students who are not very popular with the opposite sex are held back at a higher rate (17.5%) than students who are popular with the opposite sex.

Table 26

Student Popularity with Opposite Sex

	Overall		*Held Back*	
Not Very Popular with Opposite Sex	*n*[a]	%[b]	*n*[c]	%[d]
False	7510	74.1	990	13.2
True	2110	20.8	370	17.5
Unknown	520	5.1	210	40.4
Total	10140	100.0	1570	15.5

Note. From the NELS first follow-up, student questionnaire, 1990.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Student levels of popularity were also self-assessed during the first follow-up. The majority of students (78.1%) reported being perceived by their peers as very or somewhat popular, as indicated in Table 27. Students who are not perceived by their peers as being popular are held back at a higher rate (18.3%) than students who are perceived by their peers as being popular.

Table 27

Peers Think of Student as Being Popular

	Overall		*Held Back*	
Peers Think of Student as Being Popular	*n*[a]	%[b]	*n*[c]	%[d]
Not at all	1640	16.2	300	18.3
Very/Somewhat	7920	78.1	1040	13.1
Unknown	580	5.7	230	39.7
Total	10140	100.0	1570	15.5

Note. From the NELS first follow-up, student questionnaire, 1990.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

The majority of students (78.8%) reported being perceived by their peers as very or somewhat socially active, as indicated in Table 28. Students who are not perceived by their peers as socially active are held back at a higher rate (17.9%) than students who are perceived by their peers as socially active.

Table 28

Peers Think of Student as Socially Active

	Overall		*Held Back*	
Peers Think of Student as Socially Active	*n*[a]	%[b]	*n*[c]	%[d]
Not at all	1560	15.4	280	17.9
Very/Somewhat	7990	78.8	1050	13.1
Unknown	590	5.8	240	40.7
Total	10140	100.0	1570	15.5

Note. From the NELS first follow-up, student questionnaire, 1990.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

There are two types of parental involvement promoted at the school level. The first, is the level of emphasis schools place on parental involvement in the lives of their children. Results indicate that the majority of students (63.5%) attend schools that promote parental involvement, as indicated in Table 29. Students who attend schools that placed no emphasis or minor emphasis on parental involvement are held back at a higher rate (16.0%) compared to students who attend schools that placed major emphasis on parental involvement.

Table 29

School Promotes Parental Involvement

	Overall		*Held Back*	
School Promotes Parental Involvement	*n*[a]	%[b]	*n*[c]	%[d]
No/Minor Emphasis	1940	19.1	310	16.0
Major Emphasis	6440	63.5	860	13.4
Unknown	1760	17.4	400	22.7
Total	10140	100.0	1570	15.5

Note. From the NELS first follow-up, school administrator questionnaire, 1990.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

The second type of parental involvement presented in Table 30 features school encouragement of parental involvement in policy decisions. In the study, nearly fifty percent of students attend schools that never, seldom, or sometimes encourage parental involvement in policy decisions. However, the percentage of held back students was nearly the same based on the level of school encouragement of parental involvement in policy decisions.

Table 30

School Encourages Parental Involvement in Policy Decisions

School Encourages Parental Involvement in Policy Decisions	*Overall*		*Held Back*	
	n[a]	%[b]	*n*[c]	%[d]
Never/Seldom/Sometimes	4970	49.0	700	14.1
Usually/Always	3350	33.0	460	13.7
Unknown	1820	18.0	410	22.5
Total	10140	100.0	1570	15.5

Note. From the NELS first follow-up, school administrator questionnaire, 1990.
[a]Rounded number of participants in the overall study.
[b]The percentage of participants in the overall study.
[c]Rounded number of participants held back.
[d]The percentage of participants held back in each category.

Classification Tree Models

In this study, all students in the dataset were included in the classification tree analysis. Below results are presented for each model (ever held back, school request and parental request).

Ever Held Back Model

The ever held back model (Figures 3 and 4) included all students (*n*=10,140) who were either continuously promoted (*n*=8570) or held back (*n*=1570) during their academic career. This model includes all groups illustrated in Figure 2. This model was built using 91% of the data[6] and consists of 24 terminal nodes or subgroups. In Figures 3 and 4, each node contains the number of cases, and the associated percentages for each class. For instance, in the ever held back model, the root node (the node that contains the entire sample) contains 7810 (84.70%) students who were never held back, indicated by class 1, and 1410 (15.3%) held back students, indicated by class 2.

Of the 24 terminal nodes (subgroups) for the ever held back classification tree model, twelve of them are associated with being more likely to be held back. The primary split was found on academic performance and more specifically on the variable that indicates standardized test performance. In Figure 3, male students with low standardized test score, low math score, and low socio-economic status were grouped in a terminal node consisting of students greater than two and a half times more likely than average to be held back (42.4% versus the overall rate of 15.3%). This terminal node is listed in Table 31, where it can be seen that it has the second highest retention rate and is the fourth largest terminal node overall. Females with the same characteristics (low standardized test score, low math score, and low socio-economic status) were classified using three additional variables. For example, the females in this part of the tree having low reading ability were classified as being more likely to be held back, with their 29.7%

[6] The other 9% of the data were set aside to test the goodness of the model.

retention rate being almost twice the overall average. Their terminal node, labeled 7 in Figure 3 and in Table 31, is the third largest of the 24 terminal nodes. In sharp contrast, in another part of the tree, female students with high standardized test scores were directly classified as less likely to be held back, as illustrated in Figure 4. These students, in the node labeled 17 in Figure 4 and Table 31, have a retention rate of only 3.9 percent, which is only a little over one quarter the overall retention rate.

The highest held back rates in the ever held back model stem from the left side of the tree. As illustrated in Figure 3, female students with low scores on standardized tests, low math scores, low socio-economic status, with high reading ability but had changed schools one or more times were most likely to be held back (50.0%). Next male students with low standardized test scores, low math scores, low socio-economic status were from the second most likely group to be held back (42.4%). The third highest held back rate (40.0%) was among students who were popular with the opposite sex, had low standardized test scores, high math scores, no parental involvement, and had not changed schools in the past two years.

It can be noted from Table 31 that while the second highest retention rate occurs with a large node, the 1st, 3rd, and 4th highest rates occur with a small number of cases. Combined, these three nodes contain only 100 cases, which is less than one percent of the data used to build the tree.

Table 31

Estimated Retention Rates: Ever Held Back Model

Terminal Node	*n*[a]	*% Held Back*[b]
1	20	50.0
2	990	42.4
3	50	40.0
4	30	33.3
5	60	33.3
6	40	31.7
7	1110	29.7
8	70	28.6
9	540	25.9
10	120	25.0
11	190	21.0
12	60	16.7
13	200	15.0
14	290	10.3
15	2240	7.6
16	140	7.1
17	2840	3.9
18	40	0.0
19	20	0.0
20	20	0.0
21	30	0.0
22	90	0.0
23	20	0.0
24	10	0.0

Note. The overall retention rate for the cases used to build the ever held back model is 15.3%.
[a]The number of participants for each terminal node of the ever held back model.
[b]The percentage of participants held back in each terminal node of the ever held back model.

To assess the sensibility of the tree for the ever held back model, the randomly selected test sample consisting of nine percent of the data was used. The model identified (with its terminal nodes) segments of the population corresponding to higher-risk groups of students, and Table 32 shows that students in these groups collectively proved to be almost twice as likely to be held back (31.4%) compared to the overall retention rate of the test sample (17.4%).

Table 32

Classification Accuracy: Ever Held Back Model

Actual Retention Status	*n*[a]	*Predicted to be Less Likely*[b]	*Predicted to be More Likely*[c]	*Percent Correct*[d]
Continuously Promoted	760 (82.6%)	520 (91.2%)	240 (68.6%)	68.9%
Ever Held Back	160 **(17.4%)**	50 **(8.8%)**	110 **(31.4%)**	69.6%
Total	920 (100%)	570 (100%)	350 (100%)	69.0%

[a]Rounded number of participants who were or were not held back.
[b]Rounded number of participants predicted to be less likely to be held back.
[c]Rounded number of participants predicted to be more likely to be held back.
[d]The percentages of continuously promoted students predicted to be less likely to be held back, and of held back students predicted to be more likely to be held back.

Subsets of the students who are less likely to be held back are also identified. The test sample students predicted to be at lower-risk of being held back had a retention rate (8.8%) that is half of the overall average, as shown in Table 32. So, overall, the model is quite successful in separating students into higher-risk and lower-risk groups, with those students deemed to be of high risk being retained at a rate which is more than three and a half times higher than the retention rate for the low-risk group (31.4% versus 8.8%). Table 32 also assesses the predictive accuracy of the model in another way. It shows that

68.9% of the students who were continuously promoted were classified as being low risk, and 69.6% of the students who were ever held back were classified as being high risk.

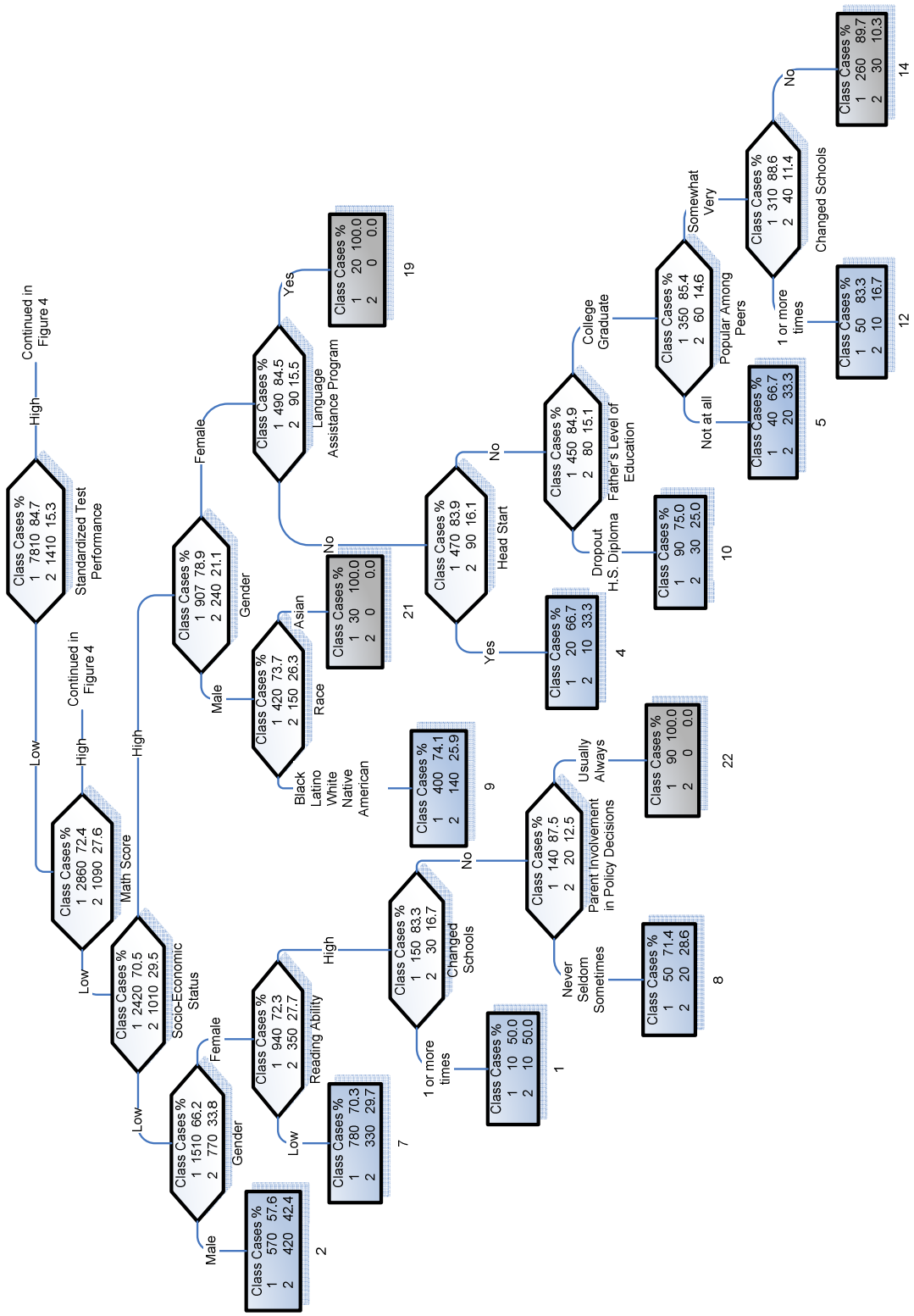

Figure 3. First Part of Ever Held Back Classification Tree (class 1 in gray = less likely to be held back, class 2 in blue = more likely to be held back).

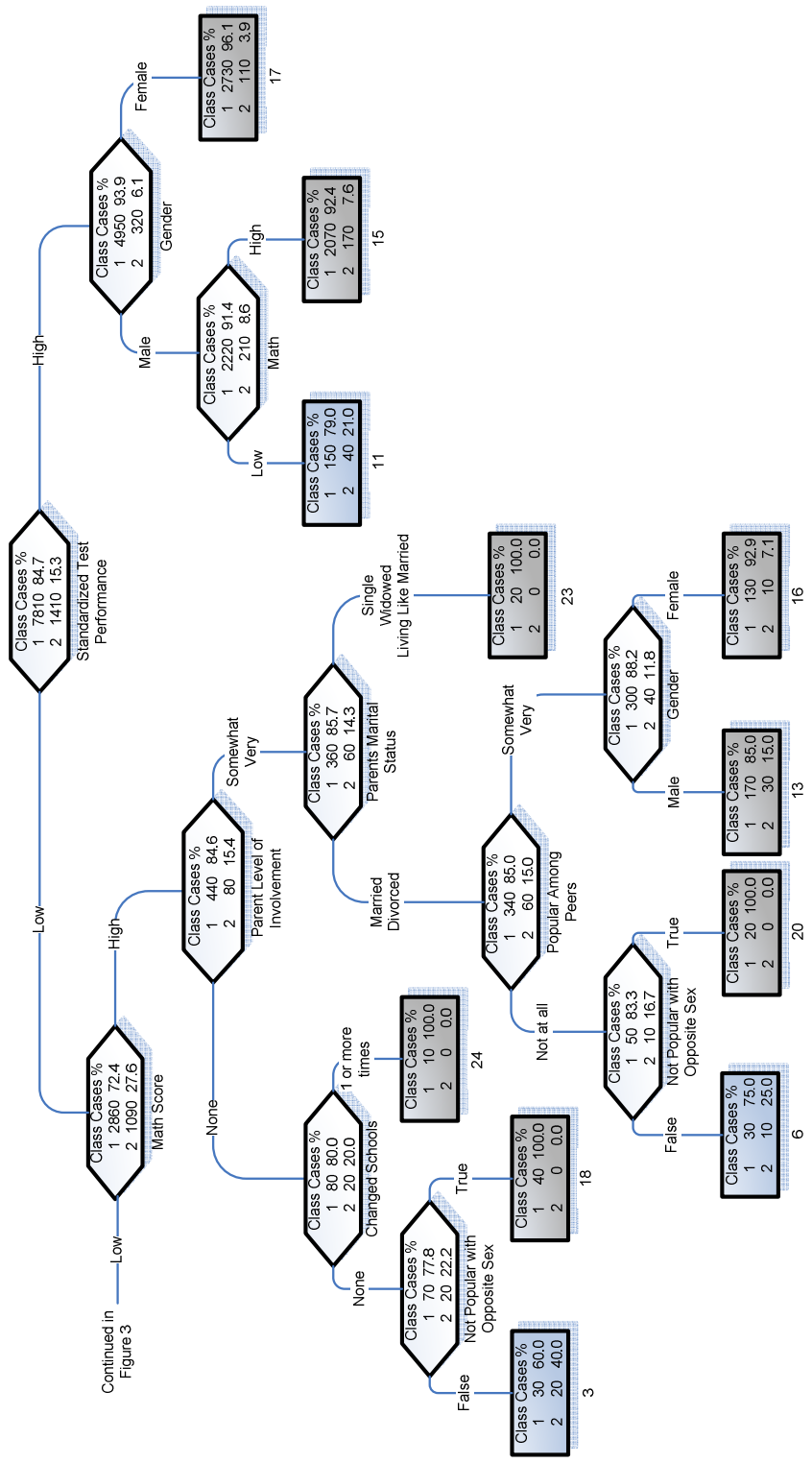

Figure 4. Second Part of Ever Held Back Classification Tree (class 1 in gray = less likely to be held back, class 2 in blue = more likely to be held back).

Parental Request Model

The parental request model (Figure 5) is based on students held back due to the request of their parents (*n*=240) along with students never held back, that is groups A (students not held back) and B (students held back due to a parental request) illustrated in Figure 2. In this model students less likely to be held back due to parental request are indicated by class 1 and students more likely to be held back due to parental request are indicated by class 2. The classification tree for parental request divides the data into 9 subgroups of which 4 correspond to students more likely to be held back (Figure 5).

This tree is smaller in comparison to the ever held back model and the classifier initially split on math score. Students with a low math score and low reading ability are found to be more than twice as likely to be held back (6.3% compared to 3.0% overall), as indicated in Table 33. Also, students with a low math score and high reading ability that changed schools one or more times during the past two years are also classified as more likely to be held back (9.1%), more than three times the overall average. The largest terminal node in this model directly classifies female students with high math scores as less likely to be held back, with a retention rate of only 0.8 percent.

Table 33

Estimated Retention Rates: Parental Request Model

Terminal Node	*n*[a]	*% Held Back*[b]
1	110	9.1
2	120	8.3
3	240	8.3
4	2210	6.3
5	350	2.9
6	2020	2.0
7	2650	0.8
8	230	0.0
9	120	0.0

Note. Overall, 3.0% of the students were held back in the parental request model.
[a]The number of participants for each terminal node of the parental request model.
[b]The percentage of participants held back in each terminal node of the parental request model.

The highest held back rates in the parent request model stem from both sides of the tree. Students most likely to be held back due to a parental request had low math scores, high reading ability but changed schools one or more times (9.1%). The next highest held back rate was among males with low math scores, high reading ability, had not changed schools, but attended schools that placed minimal emphasis on parental involvement in policy decisions (8.3%). The same held back rate (8.3%) was found among male students on the right side who did not attend Head Start and had high math scores but low standardized test scores.

To assess the sensibility of the tree for the parental request model, the randomly selected test sample consisting of nine percent of the data was used. The model identified (with its terminal nodes) segments of the population corresponding to higher-risk groups of students, and Table 34 shows these groups collectively proved to be almost twice as

likely to be held back (7.7%) compared to the overall retention rate of the test sample (3.8%).

Table 34

Classification Accuracy: Parental Request Model

Actual Retention Status	*n*[a]	*Predicted to be Less Likely*[b]	*Predicted to be More Likely*[c]	*Percent Correct*[d]
Continuously Promoted	760 (96.2%)	520 (98.1%)	240 (92.3%)	68.4%
Ever Held Back	30 **(3.8%)**	10 **(1.9%)**	20 **(7.7%)**	60.0%
Total	790 (100%)	530 (100%)	260 (100%)	68.1%

[a]The actual number of participants who were or were not held back.
[b]The number of participants predicted to be less likely to be held back.
[c]The number of participants predicted to be more likely to be held back.
[d]The percentages of continuously promoted students predicted to be less likely to be held back, and of held back students predicted to be more likely to be held back.

Subsets of the students who are less likely to be held back are also identified. The test sample students predicted to be at lower-risk of being held back had a retention rate (1.9%) less than the overall average, as shown in Table 34. So, overall, the model is quite successful in separating students into higher-risk and lower-risk groups, with those students deemed to be of high risk being retained at a rate which is four times higher than the retention rate for the low-risk group (7.7% versus 1.9%). Table 34 also assesses the predictive accuracy of the model in another way. It shows that 68.4% of the students who were continuously promoted were classified as being low risk, and 60.0% of the students who were ever held back were classified as being high risk.

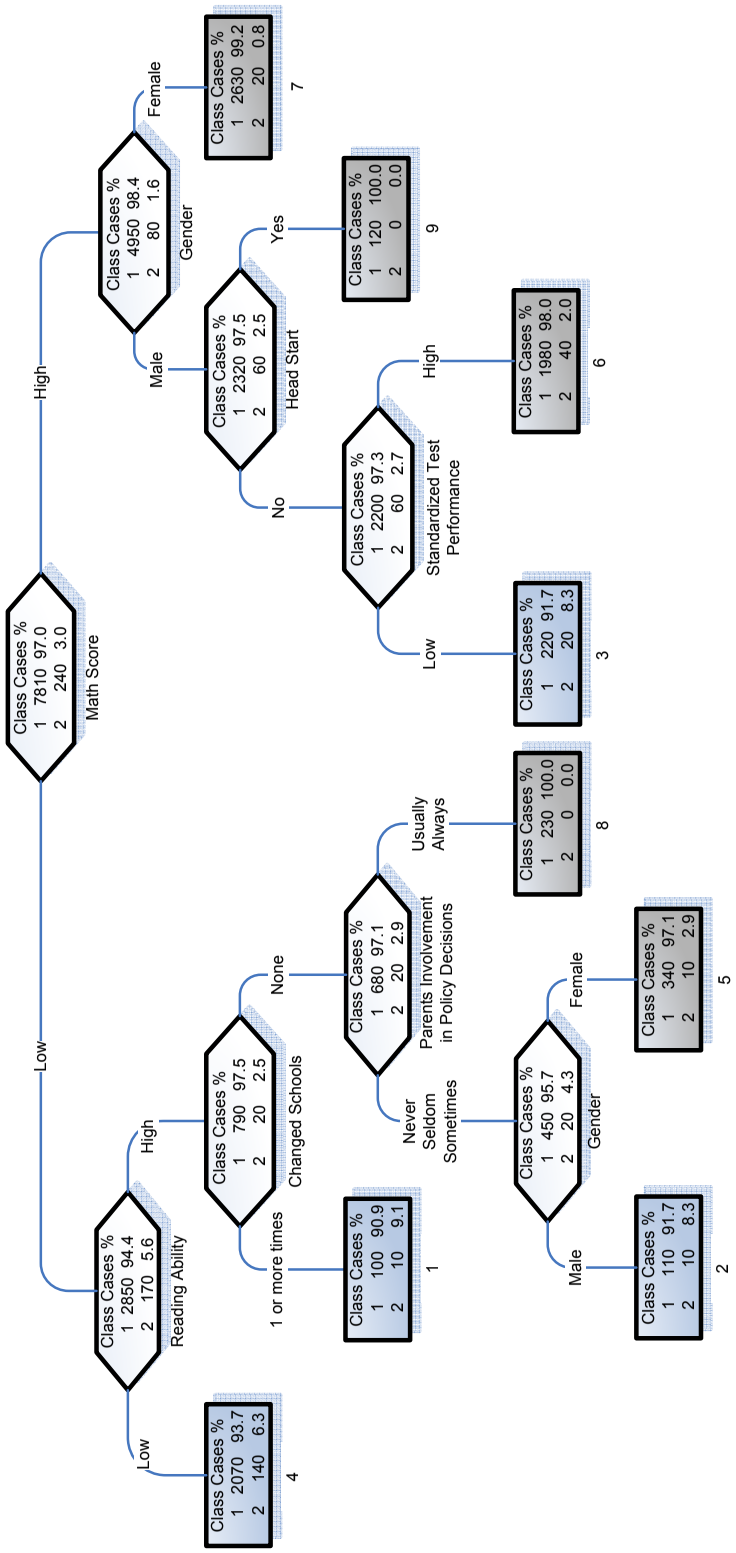

Figure 5. Parental Request Classification Tree (class 1 in gray = less likely to be held back, class 2 in blue = more likely to be held back).

School Request Model

The school request model (Figures 6 and 7) is based on students held back due to the request of the school (*n*= 790) along with students never held back, that is groups A (students not held back) and C (students held back due to a school request) illustrated in Figure 2. As previously mentioned, this model also contains students who were held back due both parental and school requests because it is ultimately the decision of the school to retain a student. In this model, less likely to be held back students are indicated by class 1 and more likely to be held back students are indicated by class 2.

The classification tree for school request divides the data into 22 subgroups, of which 11 correspond to students more likely to be held back (Table 35, Figures 6 and 7). This tree is larger in comparison to the parental request model. The classifier initially split on standardized test performance. As illustrated in Figure 7, the vast majority of students in this model were directly classified as less likely to be held back if they had high scores on standardized tests and in mathematics. This terminal node consists of over fifty percent of the cases, and only 2.2 percent of them were held back, which is about one fourth of the overall rate of 8.3 percent. The second largest terminal node identified Black, Latino, White, and Native American females, with low standardized test scores, low socio-economic status and low scores in reading and math to be more than two times as likely than average to be held back (20.0% versus 8.3%), as illustrated in Figure 6.

The highest held back rates in the school request model stem from the left side of the tree. As illustrated in Figure 6, male students with low standardized test scores, low socio-economic status, and low math scores were more than three times more likely than

the overall average to be held back (29.1% versus 8.3%). Female students who were not popular among their peers and also had low scores on standardized tests, low scores in reading ability, and high socio-economic status were also more likely to be held back (25.0%), as illustrated in Figure 7. The third highest held back rate (20.0%) was among Black, Latino, White and Native American females from low socio-economic status with low scores on standardized tests, and low scores in mathematics and reading.

Table 35

Estimated Retention Rates: School Request Model

Terminal Node	*n*[a]	*% Held Back*[b]
1	790	29.1
2	80	25.0
3	950	20.0
4	350	17.1
5	120	16.7
6	70	14.3
7	70	14.3
8	80	12.5
9	100	10.0
10	270	7.4
11	390	5.1
12	4610	2.2
13	50	2.0
14	80	0.0
15	90	0.0
16	100	0.0
17	30	0.0
18	70	0.0
19	60	0.0
20	90	0.0
21	40	0.0
22	30	0.0

Note. The overall retention rate for the cases used to build the school request model is 8.3%.
[a]The number of participants for each terminal node of the school request model.
[b]The percentage of participants held back in each terminal node of the school request model.

To estimate the probability that participants in various subsets of the population would be held back due to a school request, a randomly selected test sample consisting of nine percent of the data was used. The test data were used to determine if students categorized by the model as being more likely than average to be held back actually had a higher than average held back rate. This unbiased assessment of classification accuracy is depicted in Table 36. The model identified (with its terminal nodes) segments of the population corresponding to high-risk groups of students, who collectively proved to be more than twice as likely to be held back (21.4%) compared to the overall retention rate of the test sample (9.5%).

Table 36

Classification Accuracy: School Request Model

Actual Retention Status	*n*[a]	*Predicted to be Less Likely*[b]	*Predicted to be More Likely*[c]	*Percent Correct*[d]
Continuously Promoted	760 (90.5%)	540 (94.7%)	220 (78.6%)	71.0%
Ever Held Back	80 **(9.5%)**	30 **(5.3%)**	60 **(21.4%)**	69.9%
Total	840 (100%)	570 (100%)	280 (100%)	70.9%

[a]The actual number of participants who were or were not held back.
[b]The number of participants predicted to be less likely to be held back.
[c]The number of participants predicted to be more likely to be held back.
[d]The percentages of continuously promoted students predicted to be less likely to be held back, and of held back students predicted to be more likely to be held back.

Subsets of the students who are less likely to be held back are also identified. The test sample students predicted to be at lower-risk of being held back had a retention rate (5.3%) nearly half the overall average. So, overall, the model is quite successful in separating students into higher-risk and lower-risk groups, with those students deemed to be of high risk being retained at a rate which is more than four times higher than the

retention rate for the low-risk group (21.4% versus 5.3%). Table 36 also shows that 71.0% of the students who were continuously promoted were classified as being low risk, and 69.9% of the students who were ever held back were classified as being high risk.

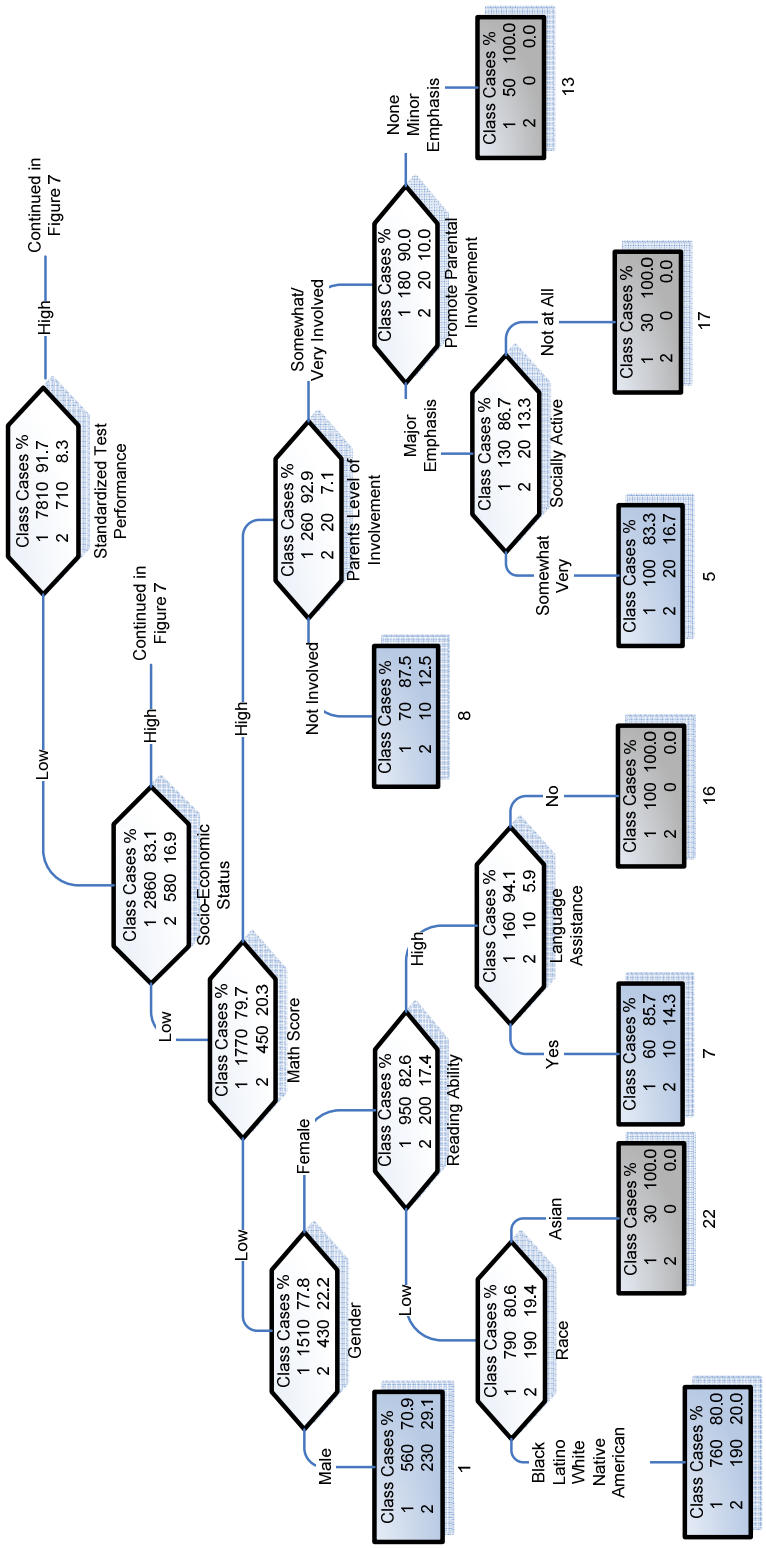

Figure 6. First Part of School Request Classification Tree (class 1 in gray = less likely to be held back, class 2 in blue = more likely to be held back).

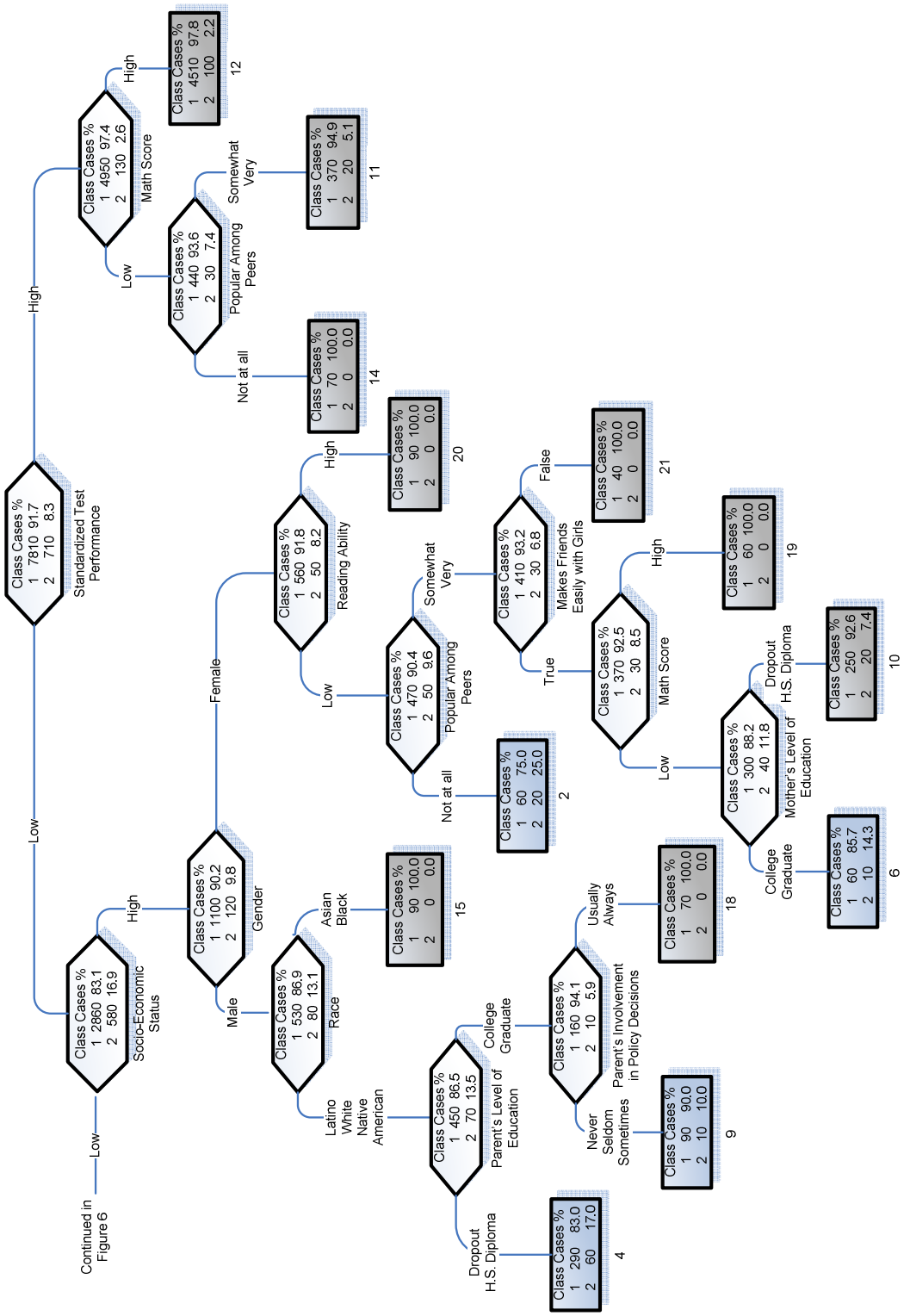

Figure 7. Second Part of School Request Classification Tree (class 1 in gray = less likely to be held back, class 2 in blue = more likely to be held back).

Forward Stepwise Logistic Regression Models

To determine the most accurate logistic regression results, forward and backward stepwise logistic regression was run for each model (ever held back, parental request and school request). As indicated in Table 37, multiple combinations of variables were considered across each model. In addition to trying all variables (which results in a large number of cases being eliminated due to missing values), various subsets of the variables were considered by eliminating variables with 10% and 20% or more missing values. A fourth subset using only variables from classification tree models was also considered. Table 37 shows that only three of the eight variations resulted in convergence for all three models. Based on the results of these three variations, it was decided to use models based on eliminating variables with 20% or more missing values (which are language assistance, Head Start, and parent level of involvement).

Table 37

Percentage Correct for Each Set of Variables and Stepwise Direction Across Models

	Less likely than average			*More likely than average*			*Overall*		
	Ever	*Parent*	*School*	*Ever*	*Parent*	*School*	*Ever*	*Parent*	*School*
	%	%	%	%	%	%	%	%	%
LR All F[a]	62.71	---	---	73.29	---	---	64.57	---	---
LR All B[b]	---	---	---	---	---	---	---	---	---
LR Tree F[c]	75.49	68.47	---	67.00	63.16	---	74.38	68.32	---
LR Tree B[d]	75.49	68.47	---	67.00	63.16	---	74.38	68.32	---
LR ≤20% F[e]	82.48	64.48	75.34	56.16	68.18	71.70	79.39	64.62	75.07
LR ≤20% B[f]	77.32	---	77.55	59.72	---	66.67	75.25	---	76.74
LR ≤10% F[g]	76.62	71.96	75.24	57.81	52.17	66.67	73.80	71.35	74.47
LR ≤10% B[h]	76.35	73.94	75.19	57.63	55.00	68.75	73.67	72.93	74.62

Note. LR = Logistic Regression, Tree = Classification Trees, Ever = Ever Held Back, Parent = Parental Request, School = School Request, F=Forward, B = Backward, --- = did not converge.

[a]Forward stepwise logistic regression with all variables.

[b]Backward stepwise logistic regression with all variables.

[c]Forward stepwise logistic regression using variables identified in the classification trees.

[d]Backward stepwise logistic regression using variables identified in the classification trees.

[e]Forward stepwise logistic regression eliminating variables with 20% or more missing values.

[f]Backward stepwise logistic regression eliminating variables with 20% or more missing values.

[g]Forward stepwise logistic regression eliminating variables with 10% or more missing values.

[h]Backward stepwise logistic regression eliminating variables with 10% or more missing values.

Ever Held Back Model

The overall majority of participants (3730/4110 = 90.7%) were not held back.[7] Forward stepwise logistic regression analysis was employed to estimate the probability that a student with a given set of characteristics would be held back. The predictor variables appearing in the final model include gender, standardized test performance, math score, socio-economic status, absenteeism, school mobility, popular among peers, and school promotion of parental involvement. A test of the final selected model versus a model with the intercept only was statistically significant, χ^2 (8, N = 4110) = 309.45, $p <$ .0005.

To assess the accuracy of the ever held back model, the randomly selected test sample consisting of nine percent of the data was used. The test data were used to determine if students categorized by the model as being more likely than average to be held back actually had a higher than average held back rate. This unbiased assessment of classification accuracy is shown in Table 38. The logistic regression model identified a high-risk segment of the population that is two and a half times more likely than average to be held back (28.6%), compared to the overall retention rate[8] (11.3%).

A subset of students who are less likely to be held back is also identified. The test sample students predicted to be at lower-risk of being held back had a retention rate (6.3%) less than 60 percent of the overall average. So, overall, the ever held back logistic regression model is quite successful in separating students into higher-risk and lower-risk

[7] The reduced number of cases is accounted for by the 5110 cases removed from the data set due to missing values and the 9% of the data set aside to test the goodness of the model.

[8] The overall retention rates for the logistic regression models differs from the classification trees because only a subset of the original data can be used for the logistic regression models.

groups, with those students deemed to be of high risk being retained at a rate which is more than four and a half times higher than the retention rate for the low-risk group (28.6% versus 6.3%). Table 38 also shows, that 82.5% of the students who were continuously promoted were classified as being low risk, and 56.2% of the students who were ever held back were classified as being high risk.

Table 38

Classification Accuracy: Ever Held Back Logistic Regression Model

Actual Retention Status	*n*[a]	*Predicted to be Less Likely*[b]	*Predicted to be More Likely*[c]	*Percent Correct*[d]
Continuously Promoted	550 (88.7%)	450 (93.7%)	100 (71.4%)	82.5%
Ever Held Back	70 **(11.3%)**	30 **(6.3%)**	40 **(28.6%)**	56.2%
Total	620 (100%)	480 (100%)	140 (100%)	79.4%

[a]Rounded number of participants who were or were not held back.
[b]Rounded number of participants predicted to be less likely to be held back.
[c]Rounded number of participants predicted to be more likely to be held back.
[d]The percentages of continuously promoted students predicted to be less likely to be held back, and of held back students predicted to be more likely to be held back.

Table 39 shows the logistic regression coefficient, Wald test statistic value, and *p* value for each of the predictors. Employing a 0.05[9] criterion of statistical significance, gender, standardized test performance, math score, socio-economic status, absenteeism, school mobility, popular among peers, and school promotion of parental involvement are significant. The negative beta associated with gender indicates males are more likely than females to be held back. (As indicated in Chapter 3, males are coded 1 and females 2.) In Table 39, the negative coefficient for gender (-0.689) means that females are less likely to be held back than males. Similarly, the positive coefficient associated with

[9] All of the *p* values in forward stepwise logistic regression are less than 0.05 given that only statistically significant variables are entered into the model.

school mobility (0.457) indicates students who changed schools more than once are more likely to be held back, given that no change was coded 0 and one more school changes was coded 1. Overall the coefficients in Table 39 indicate that males with low scores on standardized tests and in mathematics, low socio-economic status, high absenteeism and high school mobility were more likely to be held back. Furthermore, students who were not popular among their peers and whose school placed no emphasis on parental support were also more likely to be held back.

Table 39

Coefficients and Significance of Variables used in the Ever Held Back Model

Variable	β	Wald χ^2	*p*
Gender	-0.689	35.81	.000
Standardized Test Performance	-0.963	30.62	.000
Math Score	-0.762	19.69	.000
Socio-Economic Status	-0.308	6.76	.009
Absenteeism	0.416	7.15	.008
School Mobility	0.457	9.69	.002
Popular among Peers	-0.345	5.32	.021
School Promotion of Parental Involvement	-0.366	8.67	.003

Parental Request Model

The overall majority of participants (3730/3830 = 97.4%) were not held back. Forward stepwise logistic regression analysis was employed to estimate the probability that a participant would be held back. The predictor variables for the parental request model include gender, math score, school promotion of parental involvement, and mothers highest level of education. A test of the final selected model versus a model with the intercept only was statistically significant, χ^2 (4, N = 3830) = 63.91, p < .0005.

To assess the accuracy of the parental request model, the randomly selected test sample consisting of nine percent of the data was used to determine if students categorized by the model as being more likely than average to be held back actually had a higher than average held back rate. This unbiased assessment of classification accuracy is shown in Table 40. The logistic regression model identifies high-risk groups of students who collectively proved to be more than two and a half times as likely to be held back (9.1%) compared to the overall retention rate of the test sample (3.4%).

Table 40

Classification Accuracy: Parental Request Logistic Regression Model

Actual Retention Status	*n*[a]	*Predicted to be Less Likely*[b]	*Predicted to be More Likely*[c]	*Percent Correct*[d]
Continuously Promoted	560 (96.6%)	360 (97.3%)	200 (90.9%)	64.5%
Ever Held Back	20 **(3.4%)**	10 **(2.7%)**	20 **(9.1%)**	68.2%
Total	580 (100%)	370 (100%)	220 (100%)	64.6%

[a]Rounded number of participants who were or were not held back.
[b]Rounded number of participants predicted to be less likely to be held back.
[c]Rounded number of participants predicted to be more likely to be held back.
[d]The percentages of continuously promoted students predicted to be less likely to be held back, and of held back students predicted to be more likely to be held back.

A subset of students who are less likely to be held back is also identified. The test sample students predicted to be at lower-risk of being held back had a retention rate (2.7%) about 80 percent the overall average. So, overall, the logistic regression model is quite successful in separating students into higher-risk and lower-risk groups, with those students deemed to be of high risk being retained at a rate which is more than three times higher than the retention rate for the low-risk group (9.1% versus 2.7%). Table 40 also shows that 64.5% of the students who were continuously promoted were classified as

being low risk, and 68.2% of the students who were ever held back were classified as being high risk.

Table 41 shows the logistic regression coefficient, Wald test statistic value, and *p* value for each of the predictors. Employing a 0.05 criterion of statistical significance, gender, math score, school promotion of parental involvement, and mother's highest level of education are significant. The coefficients in Table 41 indicate that males with low math scores from schools that placed no emphasis on parental involvement are more likely to be held back. It was also found that students whose mother's earned a college degree were more likely to be held back. One possible explanation for this is that mothers with a college degree placed more emphasis on their child's levels of achievement, perhaps not wanting their child to advance before they were ready.

Table 41

Coefficients and Significance of Variables used in the Parental Request Model

Predictor	β	Wald χ^2	*p*
Gender	-0.644	9.11	.003
Math score	-1.535	47.22	.000
School promotion of parental involvement	-0.497	4.95	.026
Mother's highest level of education	0.528	5.25	.022

School Request Model

The overall majority of participants (3730/3930 = 94.9%) were not held back. Forward stepwise logistic regression analysis was employed to estimate the probability that a participant with a given set of characteristics would be held back. The predictor variables for the school request model include gender, standardized test performance,

math score, socio-economic status, and school mobility. A test of the final selected model versus a model with the intercept only was statistically significant, χ^2 (5, N = 3930) = 248.34, $p < .0005$.

To assess the accuracy of the school request model, the randomly selected test sample consisting of nine percent of the data was used to determine if students categorized by the model as being more likely than average to be held back actually had a higher than average held back rate. This unbiased assessment of classification accuracy is shown in Table 42. The logistic regression model identifies a high-risk group of students who collectively proved to be more than two and a half times as likely to be held back (20.0%) compared to the overall retention rate of the test sample (7.1%).

Table 42

Classification Accuracy: School Request Logistic Regression Model

Actual Retention Status	*n*[a]	*Predicted to be Less Likely*[b]	*Predicted to be More Likely*[c]	*Percent Correct*[d]
Continuously Promoted	660 (92.9%)	500 (98.0%)	160 (80.0%)	75.3%
Ever Held Back	50 **(7.1%)**	10 **(2.0%)**	40 **(20.0%)**	71.7%
Total	710 (100%)	510 (100%)	200 (100%)	75.1%

[a]Rounded number of participants who were or were not held back.
[b]Rounded number of participants predicted to be less likely to be held back.
[c]Rounded number of participants predicted to be more likely to be held back.
[d]The percentages of continuously promoted students predicted to be less likely to be held back, and of held back students predicted to be more likely to be held back.

A subset of students who are less likely to be held back is also identified. The test sample students predicted to be at lower-risk of being held back had a retention rate (2.0%) less than one-third the overall average. So, overall, the logistic regression model is quite successful in separating students into higher-risk and lower-risk groups, with

those students deemed to be of high risk being retained at a rate which is more than ten times higher than the retention rate for the low-risk group (20.0% versus 2.0%). Table 42 also shows that 75.3% of the students who were continuously promoted were classified as being low risk, and 71.7% of the students who were ever held back were classified as being high risk.

Table 43 shows the logistic regression coefficient, Wald test statistic value, and *p* value for each of the predictors. Employing a 0.05 criterion of statistical significance, gender, standardized test performance, math score, socio-economic status, and school mobility had significant effects. The coefficients in Table 43 indicate that males with low scores on standardized tests and in mathematics, low socio-economic status and high school mobility were more likely to be held back.

Table 43

Coefficients and Significance of Variables used in the School Request Model

Predictor	β	Wald χ^2	*p*
Gender	-0.809	27.05	.000
Standardized test performance	-1.321	28.14	.000
Math score	-0.876	13.63	.000
Socio-economic status	-0.602	13.85	.000
School mobility	0.483	6.32	.012

Comparison of Tree and Logistic Regression Models

This section presents the sociodemographic, family and school related factors identified across each model in the construction of classification trees and logistic

regression models. In addition, strengths and weaknesses of the two classification methods will be addressed.

An important thing to notice is that classification trees are able to make predictions on more cases since the tree classifiers can deal with missing values (because they use surrogate variables). As indicated in Table 44 the number of cases included in each logistic regression model total is less than half (44.6 to 47.5 percent) the number of cases used to build each classification tree.

Table 44

Data Used to Build Each Model

	Classification Tree	*Logistic Regression*	
Model	n_1[a]	n_2[b]	%[c]
Ever Held Back	9,220	4,110	44.6
Parental Request	8,050	3,830	47.5
School Request	8,510	3,930	46.2

[a]Number of students used to build each classification tree model.
[b]Number of students used to build each logistic regression model.
[c]The percentage of classification tree data used to build each logistic regression model.

The overall held back rate differed across methods, as shown in Table 45. This difference is due to the different samples used for the classification tree and logistic regression models. This difference makes a comparison of accuracy measures across each model problematic.

Table 45

Overall Retention Rate for each Predicted Group Across Each Model Using Classification Trees and Logistic Regression

	Ever Held Back	*Parent Request*	*School Request*
Tree[a]	15.3%	3.0%	8.3%
LR ≤ 20% F[b]	11.3%	3.4%	7.1%

Note. Tree = Classification Trees, LR = Logistic Regression, F= Forward.
[a]The overall retention rate for classification trees.
[b]The overall retention rate for forward stepwise logistic regression eliminating variables with 20% or more missing values.

Table 46 presents the percentage of students actually held back in each of the predicted groups. To assess the sensibility of the predicted groups for each model, the randomly selected test sample consisting of nine percent of the data was used. The models were successful in separating students into higher-risk and lower-risk groups with those students deemed to be of high-risk being retained at a rate three to ten times higher than the retention rate for the low-risk group.

Table 46

Percentage Held Back for the Predicted Groups Across Each Model Using Classification Trees and Logistic Regression

	Ever Held Back		*Parent Request*		*School Request*	
	Predicted Less Likely	*Predicted More Likely*	*Predicted Less Likely*	*Predicted More Likely*	*Predicted Less Likely*	*Predicted More Likely*
Tree[a]	8.8%	31.4%	1.9%	7.7%	5.3%	21.4%
LR ≤ 20% F[b]	6.3%	28.6%	2.7%	9.1%	2.0%	20.0%

Note. The tabulated percentages are the percentages of students in the test sample actually held back in each of the predicted groups. Tree = Classification Trees, LR = Logistic Regression, F= Forward.
[a]Classification trees.
[b]Forward stepwise logistic regression eliminating variables with 20% or more missing values.

As shown in Table 46, retention rates in classification trees are more than three and a half times higher (31.4% versus 8.8%) for the ever held back model, four times higher for the parental request model (7.7% versus 1.9%) and school request model (21.4% versus 5.3%). Similarly, the retention rates in stepwise logistic regression are four and a half times higher (28.6% versus 6.3%) for the ever held back model, more than three times higher for the parental request model (9.1% versus 2.7%), and ten times higher for the school request model (20.0% versus 2.0%). Overall, stepwise logistic regression models were more successful at identifying meaningfully different groups of

students at high-risk for grade retention compared to classification trees, but it should be kept in mind that the data sets used are different for the two methods, and that due to not being able to deal with missing values the logistic regression models can only be applied to a subset of the population.

Table 47 gives the percentages of held back students who were predicted to be more likely to be held back, and also the percentages of continuously promoted students who were predicted to be less likely to be held back. The largest discrepancies between classification trees and logistic regression are found in the ever held back model. For instance, with the ever held back model, logistic regression only classified 56.2% of the students actually held back as being more likely to be retained, whereas the classification tree identified 69.6% of these students as being at a higher risk of being held back. But the tree classifier's tendency to correctly identify high risk students' results in too many students who were not held back being classified as more likely than average to be held back. For continuously promoted students, logistic regression does better for the ever held back model. Logistic regression identified 82.5% of the students who were never held back as being of lower risk of being held back, while the tree-structured classifier only predicted 68.9% of these continuously promoted students as being of low risk.

Table 47

Classification Tendencies of Each Model using Classification Trees and Logistic Regression

	Less likely than average			*More likely than average*			*Overall*		
	Ever	*Parent*	*School*	*Ever*	*Parent*	*School*	*Ever*	*Parent*	*School*
	%	%	%	%	%	%	%	%	%
Tree[a]	68.9	68.8	71.0	69.6	60.0	69.8	69.0	68.1	70.9
LR ≤20% F[b]	82.5	64.5	75.3	56.2	68.2	71.7	79.4	64.6	75.1

Note. The tabulated percentages include the percentages of held back students predicted to be more likely to be held back and continuously promoted students predicted to be less likely to be held back. Tree = Classification Trees, LR = Logistic Regression, Ever = Ever Held Back, Parent = Parental Request, School = School Request.

[a]The percentage correct of participants more or less likely to be held back using classification trees.

[b]The percentage correct running forward stepwise logistic regression eliminating variables with 20% or more missing values.

Table 48 illustrates the presence of sociodemographic variables across each model. Gender is the only sociodemographic variable identified within each model and method of analysis.

Table 48

Sociodemographic Variables Across Models

	Ever		*Parent*		*School*	
Variable	*Tree*	*LR*	*Tree*	*LR*	*Tree*	*LR*
Gender	▲	▲	▲	▲	▲	▲
Ethnicity	▲	--	--	--	▲	--

Note. Tree = Classification Trees, LR = Logistic Regression, ▲ = Variable present in data analysis results.

Table 49 illustrates the presence of family background variables across each model. School mobility and socio-economic status are the family related variables that were identified the most during data analysis.

Table 49

Family Background Variables Across Models

	Ever		*Parent*		*School*	
Variable	*Tree*	*LR*	*Tree*	*LR*	*Tree*	*LR*
Highly transient home	--	--	--	--	--	--
Language assistance program	▲	--	--	--	▲	--
Parents highest level of education	--	--	--	--	▲	--
Mother's Level of Education	--	--	--	▲	▲	--
Father's Level of Education	▲	--	--	--	--	--
Parents level of involvement	▲	--	--	--	▲	--
School Mobility	▲	▲	▲	--	--	▲
Marital Status	▲	--	--	--	--	--
Socio-economic Status	▲	▲	--	--	▲	▲

Note. Tree = Classification Trees, LR = Logistic Regression, ▲ = Variable present in data analysis results.

Table 50 illustrates the presence of school related variables across each model. Math score is the only school related variable identified within each model and method of analysis.

Table 50

School Related Variables Across Models

	Ever		*Parent*		*School*	
Variable	*Tree*	*LR*	*Tree*	*LR*	*Tree*	*LR*
Absenteeism (skipped class)	--	▲	--	--	--	--
Attended Head Start	▲	--	▲	--	--	--
Standardized Test Score	▲	▲	▲	--	▲	▲
Reading Ability	▲	--	▲	--	▲	
Math Score	▲	▲	▲	▲	▲	▲
Suspended	--	--	--	--	--	--
Transferred for Disciplinary Reasons	--	--	--	--	--	--
School Remediation Policy	--	--	--	--	--	--
Makes Friend Easily with Girls	--	--	--	--	--	--
Makes Friend Easily with Boys	--	--	--	--	--	--
Not Very Popular with Opposite Sex	▲	--	--	--	--	--
Peers Think of Student as being Popular	▲	▲	--	--	▲	--
Peers Think of Student as Socially Active	--	--	--	--	▲	--
School Promotes Parental Involvement	--	▲	--	▲	▲	--
School Encourages Parent Involvement in Policy Decisions	▲	--	▲	--	▲	--

Note. Tree = Classification Trees, LR = Logistic Regression, ▲ = Variable present in data analysis results.

5. Discussion

This study contributes to the existing research on grade retention through a specific focus on the predictive factors of students who are more likely than average to be held back. The information gained from this study also sheds light on the benefits of classification trees when used in conjunction with stepwise logistic regression. Key findings are presented in this chapter according to sociodemographic, family background, and school related factors. The methods used to identify these findings are also discussed. Finally, educational implications, limitations of this study, and recommendations for future research are discussed.

Sociodemographic Factors

Gender and ethnicity are the demographic characteristics assessed in this study. Findings indicate gender is a predictor of students more likely than average to be retained. This is true across each of the three models (ever held back, parental request, and school request) for classification trees and stepwise logistic regression. This aligns with the work of Dauber, Alexander and Entwisle (1993) who also found gender to be a predictor of grade retention. The findings from this study align with literature that found males to be more likely to be held back than females (NCES, 2006). Classification tree analysis identified gender differences based on various interactions between variables.

For instance, in Figure 3, females with low standardized test scores, math scores, reading ability and socio-economic status were classified as more likely than average to be held back. Similarly, in Figure 3, males with low standardized test scores, math scores and socio-economic status were also classified as more likely than average to be held back.

The ethnic distribution of the sample only appeared in two classification tree models; ever held back and school request. As with gender, these findings were also based on interactions between variables. Based on previous literature African American males are most likely to be retained (Abidin, Golladay & Howerton, 1971; McArthur & Bianchi, 1993; Zill, Loomis & West, 1998). In this study, Asians were less likely to be held back than Blacks, Whites, Hispanics, and American Indians in some segments of the population in the ever held back and school related classification tree models. However, ethnicity does not appear in any of the three logistic regression models.

Family Background

This study found that predictive factors of grade retention that pertain to family background are parent's marital status, parent education level, parent academic support, socio-economic status, school mobility, and language proficiency. Findings regarding parent level of education align with previous research which found that low parent education increases the likelihood of grade retention (Alexander, Entwisle, & Dauber, 2003; McCoy & Reynolds, 1999). On the basis of the findings in Figure 7, students whose parents did not finish high school or earned a high school diploma were more likely than average to be held back if they had low standardized test scores, high socio-economic status and were Latino, White or Native American males. Similarly, in Figure

6, students whose parents were not involved[10] in their academic performance were more likely to be retained if they also had low standardized test scores, low socio-economic status and high scores in mathematics. This also aligns with previous research which identified low parental involvement as a risk factor of student retention (McCoy & Reynolds, 1999). Mother's level of education appears in the parental request logistic regression model and the school request classification tree model. In Figure7, the opposite result is found for mother's level of education. Students were more likely to be held back if their mother earned a college degree (14.3%). However, this finding is limited to a small number of cases compared to the overall size of the data set.

As illustrated in Figure 3, female students who changed schools one or more times were also more likely than average to be held back if they had low standardized test scores, math scores, socio-economic status, and high reading ability. This too supports previous literature (McCoy & Reynolds, 1999; Roderick, 1994; Rumberger,1995; Rumberger & Larson, 1998; Stroup & Robins, 1972). Socio-economic status also appears to be a predictor of grade retention for students in both the ever held back and school request classification tree and stepwise logistic regression models. This also aligns with previous literature which found students from low income householders were more likely to be held back (Reynolds, 1992). However, within these same models mixed results for participation in a language assistance program were found.[11] This can

[10] Parental involvement was eliminated from logistic regression models because this variable had 20% or more missing values.

[11] Language assistance was eliminated from logistic regression models because this variable had 20% or more missing values.

be attributed to the small number of cases regarding language assistance.[12] This study only identified a small segment of the population impacted by parent's marital status, as illustrated in Figure 4. One family background factor failed to support prior research which found highly transient homes to be a predictor of grade retention (Alexander, Entwisle, & Dauber, 2003; Grant, 1997; McArthur & Bianchi, 1993; McCoy & Reynolds, 1999; Powell, 2005; Stringer, 1960). One explanation for this is that the variables were coded 1 if the families moved within the past two years, and 2 if they did not move. Perhaps recoding the variable 1, if the family moved one or fewer times in the past two years, and 2 if they moved two or more times would have produced stronger findings. It is also problematic that whether or not a student was retained was not measured at the same time as home transiency.

School Related Factors

Findings from this study identified multiple predictive factors of grade retention that are school related including: academic achievement, Head Start participation, absenteeism, socialization, popularity, and school encouragement of parental involvement. Standardized test performance appeared at the top of the ever held back and school request classification tree models, whereas math score was the primary split for the parental request classification tree model. Reading ability also appeared in each classification tree model indicating that students with low scores as more likely to be held back. Similarly, the parameter estimates (beta coefficients) for stepwise logistic

[12] Small numbers of cases that appear in classification trees can be due to the hierarchical nature of the trees. CART analysis does not uniformly focus on the entire sample (some variables only pertain to a subset of the population) unlike global logistic regression models.

regression models identified standardized test performance as the strongest predictor of grade retention in the ever held back and school request models, whereas math score made the largest difference in the probability of a student being held back in the parental request logistic regression model. Therefore, academic achievement variables are the strongest predictors of students more likely to be held back in both classification tree and stepwise logistic regression models. These findings align with previous research which identified low academic achievement as a risk factor of student retention.

Findings from this study revealed that school encouragement of parental involvement in policy decisions varied by model. In the ever held back classification tree model, female students whose schools never, seldom, or sometimes encouraged parental involvement were more likely to be held back if they had low standardized test scores, low math scores, low socio-economic status, high reading ability and did not change schools in the past two years. Similarly, in the school request classification tree model male students attending schools that never, seldom or sometimes emphasize parental involvement in policy decisions were more likely than average to be held back. These students also had low standardized test scores, high socio-economic status, college educated parents, and were Latino, White or Native American. The parental request classification tree model found that students attending schools that usually or always emphasize the importance of parental involvement in policy decisions were less likely to be held back if they had low math scores, high reading ability and had not changed schools in the past two years. These findings align with Reynolds (1992) study which identified low parental involvement as a risk factor of grade retention. Different results emerged in the school request classification tree model with regards to school promotion

of parental involvement, which found that students in some segments of the population were less likely to be held back if the school placed minor to no emphasis on parental involvement. It can be speculated that this disparity is attributed to the positive influence parental involvement has in a student's life regardless of whether or not a school emphasizes the importance of parental involvement.

In line with existing literature, participation in the Head Start preschool program[13] had mixed results. In the parental request classification tree model male students who attended Head Start were less likely than average to be held back if they had high math scores.[14] This aligns with Zill, Loomis and West's study (1998), while the ever held back tree model agreed with the work of Reynolds (1992), finding that students who attended Head Start were more likely than average to be held back. Reynolds (1992) included similar variables in his study including gender, ethnicity, socio-economic status, parental involvement, school mobility, reading achievement and math achievement; however, only Head Start was found to be a statistically significant predictor of grade retention.

On the basis of the findings from this study, students are more likely to be held back if they are not at all popular;[15] however, students who are somewhat or very socially active are also more likely to be retained.[16] However, these findings are limited to a small number of cases compared to the overall size of the data set. Aligned with previous research, the ever held back logistic regression model found that absenteeism was a

[13] Head Start was eliminated from logistic regression models because this variable had 20% or more missing values.
[14] Head Start pertains to a small subset of the population.
[15] Findings regarding student popularity among peers are shown in the ever held back models for classification trees and logistic regression as well as the school request classification tree.
[16] Findings regarding student social activity appear in the school request classification tree.

predictor of student retention. Four school related factors failed to support prior research which found the ability to make friends (Abidin, Jr., Golladay, & Howerton, 1971; Alexander, Entwisle, & Dauber, 2003; Cadigan, Entwisle, Alexander, & Pallas, 1988; Dauber, Alexander, & Entwisle, 1993; Jimerson, 2001), school remediation policy (Janosz, LeBlanc, Boulerice, & Tremblay, 1997; Jimerson, Egeland, & Teo, 1999; Roderick, 1994; Rumberger, 1995), school suspensions and school transfers (McCoy & Reynolds, 1999; Roderick, 1994; Rumberger, 1995; Rumberger & Larson, 1998; Stroup & Robins, 1972) to be predictors of grade retention.

Comparison of Classification Trees and Stepwise Logistic Regression

The use of classification trees and stepwise logistic regression identified segments of the population more likely than average to be held back and segments less likely than average to be held back. Both classification trees and logistic regression models revealed similar risk factors of grade retention; however, classification trees also suggested the interactive nature of the predictors.

In this study, the sensibility of logistic regression and tree models for the predicted groups were assessed using the randomly selected test sample consisting of nine percent of the data. Findings reported regarding the predictive accuracy of the models were similar for both methodologies. For example, in each model students who were predicted to be more likely to be held back contained a larger proportion of students who were actually held back than those cases who were predicted to be less likely to be held back. Classification trees identified segments of the population corresponding to higher-risk groups of students in the terminal nodes, whereas the accuracy of logistic

regression models was assessed by estimating the probability of being held back for each of the test set cases from the fitted logistic regression models. When compared to the overall held back rate, the estimated probabilities in both classification trees and logistic regression models were able to correctly identify students more likely to be held back.

Overall, classification tree models possess certain advantages over stepwise logistic regression models. Classification trees can be significantly easier to interpret by individuals who do not have advanced training in statistical methodology. Classification trees provided a graphic display of the results which supports their interpretation. Also, the graphical display of tree models can assist in visualizing the interactive nature of predictors as well as the hierarchical importance of each variable in predicting the modeled response by attaching meaning to terminal node (subgroup) combinations. Furthermore, classification trees automatically deal with missing values, leaving more data to build the models, and allowing predictions to be made on a larger proportion of cases. As indicated in Table 44, for each of the three models the number of cases available for logistic regression modeling was less than half the number of cases used for classification tree construction.

Educational Implications

Teachers and parents need to take a number of factors into account when they consider making a request for a student to be retained. Teachers can use the characteristics of students who are more likely than average to be retained to identify at-risk students in need of an intervention. Interventions based on variables that negatively impact student academic performance such as low reading ability and low math scores

can help reduce the risk of student grade retention. For example, Gatti (2004) conducted a study to provide empirical evidence of the effectiveness of several classroom activities designed to help students with low scores in reading and mathematics. One hundred seventy-seven students in grades 2, 3, and 4 were placed in the eight-week intervention based on teacher nominations of students they believed would be referred for grade retention within the next month. Students placed in the intervention received a series of assessments three times a week in 20-25 minute sessions. Assessments included teacher nominations for retention which were used to increase the accuracy of referrals to special education services.

Limitations

There are possible limitations to the current study caused by the use of existing data. The variables used to build each of the three models (ever held back, parental request, and school request) were selected based on grade retention factors that were identified in previous studies. Although the database used in this study collected data on each of the previously identified factors of grade retention, data were collected at three different intervals. The first interval was in 1988 during the base year of the study when students were in eighth grade. Subsequent follow-up data collection was conducted every two years when students were in tenth and twelfth grades in 1990 and 1992, respectively. For instance, some variables, such as high rates of home transiency, were collected later in a student's academic career. Data on students from homes with high rates of transiency were collected during the first follow-up in 1990 which asked students to report whether or not their families had moved within the past two years. This

presents a problem when predicting early year retention given that students reported moving to a new home between tenth and twelfth grades.

A similar issue exists regarding the predictability of student academic achievement. For example, standardized test scores were based on the composite scores from the cognitive test battery administered to students during tenth and twelfth grades; however, grade retention occurred during kindergarten through eighth grade. Therefore, other variables pertaining to student academic achievement are needed in order to predict student achievement during the time grade retention occurred.

Another issue pertains to variables having high proportions of unknown values (e.g., for 80.9% of the participants it's unknown whether or not they attended a language assistance program, and for 64.0% of the participants the parents' level of involvement is unknown). A large number of missing values makes it challenging to use classification in order to identify retained students since these cases are ignored in logistic regression and surrogate variables are used in CART analysis. It can be noted that 24 of the 26 predictor variables have missing values. For 22 of these 24 variables, the retention rate is higher than average for the participants in the unknown category. The retention rate for the "unknowns" exceeds 30% for 11 of the 24 predictors, exceeds 40% for three of them, and exceeds 50% for one of them, suggesting that unknown variable values are commonly associated with subsets of the participants having extremely high retention rates. The fact that higher rates of unknown values are both common and often associated with rather high retention makes it extremely difficult to accurately predict students most likely to be retained.

Recommendations for Future Research

Recommendations for future research relate to the aforementioned limitations of the study and have implications for educational research methodologists. In reference to the limitations of the study, data used in future research pertaining to the predictors of grade retention need to include data collected at the time retention occurred. Findings from this study can also benefit the methodological practices used in educational research. Current practices employed to evaluate the predictive nature of variables in educational research primarily use stepwise logistic regression; however, as indicated in this study, employing classification trees in conjunction with stepwise logistic regression will enable educational research methodologists to explore the predictive nature of variables that were otherwise limited, due to the inability of logistic regression to handle missing values. Given the ability of classification trees to effectively handle missing data and identify interactions between variables, the next reasonable step in facilitating the use of classification trees in educational research is for methodologists to apply this technique to studies that were previously limited by the use of only logistic regression. Future research in this area should also strive to have a dataset with few missing values which will likely generate appreciably more accurate results than were found in the present study. It is also recommended that future research evaluating the impact of grade retention be conducted using predictor variables measured at the sixth grade level (rather than eighth grade and beyond) that are coded using actual numerical scores instead of quartiles for academic achievement. Adding interaction terms to each logistic regression model is also recommended.

Summary

Prior to this study, research looking at all previously identified characteristics of retained students had not been conducted to my knowledge, nor had a study using classification trees to predict grade retention been done. The current study not only identified the predictive factors of grade retention but also demonstrated the effectiveness of using classification trees in conjunction with stepwise logistic regression in educational research. Both classification tress and stepwise logistic regression provided an estimate of the probability that a student with a given set of characteristics would be held back. Overall, these analyses suggest that classification trees can be used to develop relatively efficient models to identify students more likely than average to be held back. Classification tree models also provide a detailed visual presentation of multiple variable interactions, consequently making it easier for educators to identify students more likely to be held back. In order to take advantage of the complementary nature of these two methods, it is recommended that both approaches be considered when developing predictive models.

Appendix:

Key Variables by Data Collection Sources

The key variables used in the study are listed below including the type of respondents, wave of data collection, and the nature of each variable after being recoded.

Model	*Nature of the Variable*	*Respondent*	*Wave of Data Collection*
Ever held back	Binary	Student	Second Follow-up (1992)
Parental request	Binary	Parent	Base Year (1988)
School request	Binary	Parent	Base Year (1988)

Sociodemographic Variables	*Nature of the Variable*	*Respondent*	*Wave of Data Collection*
Gender	Binary	Student	Third Follow-up (1994)
Race/Ethnicity	Categorical	Student	Third Follow-up (1994)

Family Background	*Nature of the Variable*	*Respondent*	*Wave of Data Collection*
Highly transient home	Ordinal	Student	First Follow-up (1990)
Language assistance program	Binary	Student	Base Year (1988)
Parents' highest level of education	Ordinal	Parent	Second Follow-up (1992)
Mother's level of education	Ordinal	Parent	Base Year (1988)
Father's level of education	Ordinal	Parent	Base Year (1988)
Parents' level of involvement	Ordinal	Teacher	First Follow-up (1990)
School mobility	Ordinal	Student	Second Follow-up (1992)
Marital status of parents	Categorical	Parent	Second Follow-up (1992)
Socio-economic status	Ordinal	Parent	Second Follow-up (1992)

School Related Factors	*Nature of the Variable*	*Respondent*	*Wave of Data Collection*
Absenteeism	Ordinal	Student	Base Year (1988)
Attended Head Start	Binary	Parent	Base Year (1988)
Standardized test score	Ordinal	Cognitive test battery	First and Second Follow-up (1990 & 1992)
Reading ability	Ordinal	Cognitive test battery	First and Second Follow-up (1990 & 1992)
Math score	Ordinal	Cognitive test battery	First and Second Follow-up (1990 & 1992)
Suspended	Ordinal	Student	Second Follow-up (1992)
Transferred for disciplinary reasons	Ordinal	Student	Second Follow-up (1992)
School retention policy	Binary	School Administrator	First Follow-up (1990)
Peers think of student as being popular	Categorical	Student	First Follow-up (1990)
Peers think of student as socially active	Categorical	Student	First Follow-up (1990)
Makes friends easily with girls	Categorical	Student	First Follow-up (1990)
Makes friends easily with boys	Categorical	Student	First Follow-up (1990)
Makes friends easily with the opposite sex	Categorical	Student	First Follow-up (1990)
Not very popular with the opposite sex	Categorical	Student	First Follow-up (1990)
School promoted parental involvement	Categorical	School Administrator	First Follow-up (1990)
School encourages parent involvement in policy decisions	Categorical	School Administrator	First Follow-up (1990)

References

References

Abidin, R. R., Jr., Golladay, W. M., & Howerton, A. L. (1971). Elementary school retention an unjustifiable, discriminatory and noxious education policy. *Journal of School Psychology, 9*, 410-417. doi:10.1016/0022-4405(71)90032-X

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

Alexander, K. L., Entwisle, D. R., & Horsey, C. S. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of Education, 20,* 87-107. doi:10.2307/2673158

Ammons. J. (1976). A study of the effects of nonpromotion and promotion as related to achievement and self-concept of elementary students. *Dissertation Abstracts International, 36,* 5011A.

Anfinson, R. D. (1941). School progress and pupil adjustment. *The Elementary School Journal, 41*, 1941, 507-514.

Ashbury, C. (1975). Maturity factors related to discrepant achievement of white and black first graders. *Journal of Negro Education, 44*, 493-501. doi:10.2307/ 2966634

Bedoian, V. (1954). Social acceptability and social rejection of the underage, at-age, and overage pupils in the sixth grade. *The Journal of Education Research, 47*, 513-520.

Bocks, W. (1977). Nonpromotion: A year to grow? *Educational Leadership, 34,* 379-383.

Briggs, L. (1968). The impact of failure on elementary school pupils. *Dissertation Abstracts International, 27,* 2719A.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees.* Boca Raton, Chapman and Hall.

Brooks-Gunn, J., Guo, G., & Furstenburg, F. F. (1993). Who drops out and who continues beyond high school? A 20-year follow-up of Black urban youth.

Journal of Research on Adolescence, 3, 271-294. doi:10.1207/s15327795jra 0303_4

Buratti, S., Benedetti, S. Scampicchio, M., & Pangerod, E.C. (2004). Characterization and classification of Italian Barbera wines by using an electronic nose and an amperometric electronic tongue. *Analytica Chimica Acta, 525*, 133-139. doi:10.1016/j.aca.2004.07.062

Cadigan, D., Entwisle, D. R., Alexander, K. L., & Pallas, A. M. (1988). First-grade retention among low achieving students: A search for significant predictors. *Merrill-Palmer Quarterly, 34*, 71-88.

Caplan, P. J. (1973). The role of classroom conduct in the promotion and retention of elementary school children. *The Journal of Experimental Education, 41*, 8-9.

Chafe, W. L. (1984). Integration and involvement in speaking, writing and oral literature. In D. Tanen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 35-53. Norwood, NJ: Ablex.

Chansky, N. M. (1964). Asking the right questions about teacher preparation: Contributions of research on teaching thinking. *Educational Researcher, 17*(2), 5-12.

Coffield, W. (1954). A longitudinal study of the effects of nonpromotion on educational achievement in the elementary school. *Dissertation Abstracts International, 42,* 2383A.

Cosden, M., Zimmer, J., & Tuss, P. (1993). The impact of age, sex, and ethnicity on kindergarten entry and retention decisions. *Educational Evaluation and Policy Analysis, 15*, 209-222.

Dauber, S. L., Alexander, K. L., & Entwisle, D. R. (1993). Characteristics of retainees and early precursors of retention in grade: Who is held back? *Merrill-Palmer Quarterly, 39*, 326-343.

Dobbs, V., & Neville, D. (1967). The effects of nonpromotion on achievement of groups matched from retained first graders, and promoted second graders. *The Journal of Educational Research, 60*, 427-474.

Ennis, M., Hinton, G., Naylor, D., Revow, M., & Tibshirani, R. (1998). A comparison of statistical learning methods on the GUSTO database. *Statistics in Medicine, 17,* 2501-2508. doi:10.1002/(SICI)1097-0258(19981115)17:21<2501::AID-SIM938 >3.0.CO;2-M

Fernandez, R. M., Paulsen, R., & Hirano-Nakanishi, M. (1989). Dropping out among Hispanic youth. *Social Science Research, 18,* 21-52. doi:10.1016/0049-089X(89) 90002-1

Finlayson, H. L. (1977). Nonpromotion and self-concept development. *Phi Delta Kappan, 59*(3), 205-206.

Finlayson, H. (1975). *The effect of nonpromotion upon the self-concept of pupils in primary grades*. Unpublished doctoral dissertation, Temple University, Philadelphia. (ERIC Document Reproduction Service No. ED155566)

Gatti, S. L. (2004). Identifying students at risk for academic failure: The application of a prereferral screening model including responsiveness to intervention. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 65*, 2087.

Godfrey, E. (1972). The tragedy of failure. *Educational Digest, 37*(5)*,* 34-35.

Goodlad, J. I. (1954). Some effects of promotion and nonpromotion upon the social and personal adjustment of children. *The Journal of Experimental Education, 22*, 301-328.

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

Grant, J. (1997). *Retention and its prevention: Making informed decisions about individual children.* Rosemont, NJ: Modern Learning Press.

Graue, M. E., & DiPerna, J. (2000). Redshirting and early retention: Who gets the "gift of time" and what are its outcomes? *American Educational Research Journal, 37*, 509-534.

Hains, A. A. (1981). The effect of retention on the self-concept of elementary students in grades three through five as compared to the self-concept of elementary students who have been socially promoted. *Dissertation Abstracts International, 42,* 5077.

Harper, P. (1996). A review and comparison of classification algorithms for medical decision making. *Health Policy, 71,* 315-331. doi:10.1016/j.healthpol.2004.05.002

Hayes, C. L. (2005). Predictors of grade retention and special education placement: Implications for improving educational outcomes in the state of Florida. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 66*, 1636.

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. 1633). London: The Farmer Press, 16-33.

Holmes, C. T., & Matthews, K. M. (1984). The effects of nonpromotion on elementary and junior high school pupils: A meta-analysis. *Review of Educational Research, 54,* 225-236.

Hosmer, D. W., Jr., & Lemeshow, S. (1999). *Applied survival analysis: Regression modeling of time to event data.* New York: Wiley.

Ingels, S.J., Dowd, K.L., Baldridge, J.D., Stipe, J.L., Bartot, V.H., & Frankel, M.R. (1994). *NELS:88 Second Follow-Up: Student Component Data File User's Manual* (NCES 93-374). Washington, DC: National Center for Education Statistics.

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

Janosz, M., LeBlanc, M., Boulerice, B., & Tremblay, R. E. (1997). Disentangling the weight of school dropout predictors: A test on two longitudinal samples. *Journal of Youth and Adolescence, 26*, 733-762. doi:10.1023/A:1022300826371

Jimerson, S. R., (1999). On the failure of failure: Examining the association between early grade retention and education and employment outcomes during late adolescence. *Journal of School Psychology,* 37, 243-272. doi:10.1016/S0022-4405(99)00005-9

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

Jimerson, S. R., Carlson, E., Rotert, M., Egeland, B., & Sroufe, L. A. (1997). A prospective, longitudinal study of the correlates and consequences of early grade retention. *Journal of School Psychology, 35,* 3-25. doi:10.1016/S0022-4405(96) 00033-7

Jimerson, S. R., Egeland, B., & Teo, A. (1999). Achievement across time: A longitudinal study of deflections, considering early school and family factors. *Journal of Educational Psychology, 91*, 116-126. doi:10.1037/0022-0663.91.1.116

Jimerson, S. R., & Kauffman, A. M. (2003). Reading, writing and retention: A primer on grade retention research shows that neither grade retention nor social promotion improves educational success. Familiarity with this research is essential when seeking intervention strategies. *The Reading Teacher, 56*, 622-635.

Johnson, A. A. (1968). A study of the relationship between nonpromotion and the male student's self-concept of academic ability and his perceived parental, friends', and teachers' evaluations of his academic ability. *Dissertation Abstracts, 29*(02A), 0409.

Katz, P. A., & Ziglar, E. (1967). Self-image disparity: A developmental approach. *Journal of Personality and Social Psychology, 5,* 186-195. doi:10.1037/h00 24190

Kitsantas, P. K., Moore, T. W., & Sly, D. F. (2007). Using classification trees to profile adolescent smoking behaviors. *Addictive Behaviors, 32,* 9–23. doi:10.1016/j. addbeh.2006.03.014

Kitsantas, P., Hollander, M., & Li, L. (2006). Using classification trees to assess low birth weight outcomes. *Artificial Intelligence in Medicine, 38,* 275-289. doi:10.1016/j.artmed.2006.03.008

Koons, C. L. (1977). Nonpromotion: A dead road. *Phi Delta Kappan, 58*, 701-702.

Kuhnert, P. M., Do, K. A., & McClure, R. (2000). Combining non-parametric models with logistic regression: An application to motor vehicle injury data. *Computational Statistics & Data Analysis, 34,* 371-386. doi: 10.1016/S0167-9473(99)00099-7

Lee, V. E., Brooks-Gunn, J., & Schnur, E., (1988). Does Head Start work? A 1-year follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. *Developmental Psychology, 24*, 210-222. doi:10.1037/0012-1649.24.2.210

Livingston, H. (1959). Key to the dropout problem: The elementary school. *Elementary School Journal, 58*, 267-270. doi:10.1086/459729

Long, W. J., Griffith, J. L., Selker, H. P., & D'Agostino, R. B. (1993). A comparison of logistic regression to decision tree induction in a medical domain. *Computers and Biomedical Research, 26,* 74-97. doi:10.1006/cbmr.1993.1005

Marshall, R. J. (2001). The use of classification and regression trees in clinical epidemiology. *Journal of Clinical Epidemiology, 54,* 603-609. doi:10.1016/ S0895-4356(00)00344-9

McArthur, E. K., & Bianchi, S. M. (1993). *Characteristics of children who are 'behind' in school.* Alexandria, VA. American Statistical Association, 1993 Proceedings of the Social Statistics Section.

McCoy, A. R., & Reynolds, A. I. (1999). Grade retention and school performance: An extended investigation. *Journal of School Psychology, 37*, 273-298. doi:10.1016/ S0022-4405(99)00012-6

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

Morris, D. R. (1993). Patterns of aggregate grade retention rates. *American Educational Research Journal, 30*, 497-514.

Morrison, I. E., & Perry, I. F. (1956). Acceptance of overage children by classmates. *The Elementary School Journal, 56*, 217-220. doi:10.1086/459447

Morrison, K., & On No, A. I. (2007). Does repeating a year improve performance? The case of teaching English. *Educational Studies, 33*, 353-371. doi:10.1080/03055690701423333

National Center for Educational Statistics. (2006). *The condition of education: Grade retention.* U.S. Department of Education: Institute of Education Sciences.

Nelson, L. M., Bloch, D. A., Longstreth, W. T., & Shi, H. (1998). Recursive partitioning for identification of disease risk subgroups: A case-control study of subarachnoid hemorrhage. *Journal of Clinical Epidemiology, 51*, 199-209. doi:10.1016/S0895-4356(97)00268-0

Ogden, K. (1971). An evaluation of nonpromotion as a method of improving academic performance. *Dissertation Abstracts International, 32,* 795A.

Owens, S., & Ranick, D. (1977). The Greenville program: A commonsense approach to basics. *Phi Delta Kappan, 58,* 531-533.

Plummer, D., & Graziano, W. (1987). Impact of grade retention on the social development of elementary school children. *Developmental Psychology, 23*, 267-275. doi:10.1037/0012-1649.23.2.267

Plummer, D., Lineberger, M., & Graziano, W. (1986). Academic and social consequences of grade retention: A convergent analysis. In L. G. Katz (Ed.). *Current topics in early childhood education, 6,* 224-252. Norwood, NJ: Ablex Publishing Corporation.

Powell, P. J. (2005). The effects of grade retention: Life histories of adults who were retained as children. *Dissertation Abstracts International, 66*(09A), 3213 (UMI No. 3189068).

Raymond, E. G., Tafari, N., Troendle, J. F., & Clemens, J. D. (1994). Development of a practical screening tool to identify preterm, low birth weight neonates in Ethiopia. *Lancet, 344,* 524-527**.** doi:10.1016/S0140-6736(94)91905-4

Reinherz, H., & Griffin, C. (1970). The second time around: Achievement and progress of boys who repeated one of the first grades. *School Counselor, 17,* 213-218.

Reiter, R. (1973). The promotion/retention dilemma, what research tells us. Report No. 7416, Philadelphia: Philadelphia School District Office of Research and Evaluation, (ERIC Document Reproduction Service No. ED 099412).

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

Rock, D.A., & Pollack, J.M. (1995). *Psychometric Report for the NELS:88 Base Year Through Second Follow-Up* (NCES 95-382). Washington, DC: National Center for Education Statistics.

Roderick, M. (1994). Grade retention and school dropout: Investigating the association. *American Educational Research Journal, 31*, 729-759.

Rose, J., Medway, F., Cantrell, V., & Marus, S. (1983). A fresh look at the retention–promotion controversy. *Journal of School Psychology, 21,* 201-211. doi:10.1016 /0022-4405(83)90015-8

Rumberger, R. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal, 32*, 583-625.

Rumberger, R., & Larson, K. (1998). Student mobility and the increased risk of high school dropout. *American Journal of Education, 107*, 1-35.

Rummel, N. P. (2007). Teacher self-efficacy beliefs and their impact on recommendations for student retention at grades K-2. *Dissertation Abstracts International, 68*(04A), 1261 (UMI No. 3257328). doi:10.1086/444201

Safer, D. J. (1986). Nonpromotion correlates and outcomes at different grade levels. *Journal of Learning Disabilities, 19*, 500-503. doi:10.1177/0022219486019 00812

Salford Systems (2000). CART Software Release 4.0 edition. San Diego, CA7 Salford Systems.

Sandin, A. A. (1944). Social and emotional adjustment of regularly promoted and non-promoted pupils. New York: Bureau of publications, Teachers College, Columbia University.

Sandoval, J., & Fitzgerald, P. (1985). A high school follow-up of children who were nonpromoted or attended a junior first grade. *Psychology in the Schools, 22,* 164-170. doi:10.1002/1520-6807(198504)22:2<164::AID-PITS2310220210>3.0.CO; 2-V

Scruggs, E., & Mastropieri, M. A. (2002). On babies and bathwater: addressing the problems on identification of learning disabilities. *Learning Disability Quarterly, 25*(3), 155-168). doi:10.2307/1511299

Shephard, L. A., & Smith, M. L. (1987). Effect of kindergarten retention at the end of first grade. *Psychology in the Schools, 24*, 346-357. doi:10.1002/15206807(1987 10)24:4<346::AID-PITS2310240409>3.0.CO;2-7

Sims, C. J., Meyn, L., Caruana, R., Rao, R.B., Mitchell, T., & Krohn, M. (2000). Predicting cesarean delivery with decision tree models. *American Journal of Obstetrics and Gynecology, 183,* 1198-1206. doi:10.1067/mob.2000.108891

Staskowski, M., & Rivera, E. A. (2005). Speech—language pathologists' involvement in responsiveness to intervention activities: A complement to curriculum-relevant practice. *Topics in Language Disorders, 25*, 132-147.

Stringer, L. A. (1960). Report on a retentions program. *Elementary School Journal, 60,* 370-375. doi:10.1086/459830

Stroup, A., & Robins, L. (1972). Elementary school predictors of high school dropout among black males. *Sociology of Education, 45*, 212-222. doi:10.2307/2112008

Sutton, C. D. (2005). Classification and regression trees, bagging, and boosting. In C.R. Rao, E. J. Wegman, & J. L. Solka (Eds.), *Handbook of Statistics, Volume 24, Data Mining and Data Visualization* (pp.303-329). Elsevier B. V., The Netherlands.

Thomas, S., & Knudsen, D. D. (1965). The relationship between nonpromotion and the dropout problem. *Theory into Practice, 4*, 90-94. doi:10.1080/00405846509541952

Toschke, A. M., Beyerlein, A., & von Kries, R. (2005). Children at high risk for overweight: a classification and regression trees analysis approach. *Obesity Research, 13*, 1270-1274. doi:10.1038/oby.2005.151

Vaughn, S., & Fuchs, L. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice, 18*, 137-146. doi:10.1111/1540-5826.00070

Werneck, G. L., De Carvalho, D. M., Barroso, D. E., Cook, E. F., & Walker, A. M. (1999). Classification trees and logistic regression applied to prognostic studies: a comparison using meningococcal disease as an example. *Journal of Tropical Pediatrics, 45,* 248-251. doi:10.1093/tropej/45.4.248

White, K., & Howard, J. L. (1973). Failure to be promoted and self-concept among elementary school children. *Elementary School Guidance and Counseling, 7,* 182-187.

Wylie, R. C. (1974). The measured academic achievement of two groups of first grade students matched along five variables when one group has been retained. *Dissertation Abstracts International, 41,* 3418A.

Zhang, H., & Bracken, M. B. (1995). Tree-based risk factor analysis of preterm delivery and small-for-gestational-age birth. *American Journal of Epidemiology, 141,* 70-78.

Zhang, H., & Bracken, M. B. (1996). Tree-based, two-stage risk factor analysis for spontaneous abortion. *American Journal of Epidemiology, 144*, 989-996.

Zhang, H., & Singer B. (1999). *Recursive partitioning in the health sciences.* New York: Springer-Verlag.

Zill, N., Loomis, L. S., & West, J. (1998). The elementary school performance adjustment of children who enter kindergarten late or repeat kindergarten. Findings from national surveys. NCES 98-97. National Center for Education Statistical Analysis Report. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Curriculum Vitae

Deanna Kelley-Winstead graduated from West Potomac High School, Alexandria, Virginia, in 1998. She received her Associate of Arts from Northern Virginia Community College in 2001 and Bachelor of Arts from George Mason University, Fairfax, Virginia, in 2003 while employed as a Youth Counselor in Alexandria, Virginia from 1999 to 2004. She received her Master of Education from George Mason University in 2005 while employed as a Workforce Development Skills Instructor in Alexandria, Virginia and Research Associate in Gaithersburg, Maryland. She has also held multiple graduate research assistant positions at George Mason University between 2004 and 2008. During her doctoral studies she has been employed as a Research Analyst and Consultant in Northern Virginia.