ABSTRACT

LYNCH, LIAN. Repeated Test Taking on the SAT: The Effects of Ethnicity, Gender, Financial Aid, and High School Location. (Under the direction of Kevin P. Brady, Ph.D. & Shevaun D. Neupert, Ph.D.)

The purpose of this study was to determine whether there was a significant relationship between repeated test taking on the SAT and several demographic measures. These measures included ethnicity (American Indian/Alaskan Native, Asian/Asian American/Pacific Islander, Black/African American, Mexican/Mexican American, Puerto Rican, Other Hispanic/Latino/Latin American, White, Other, and Unreported); gender (male, female, unreported); need for financial aid (yes, no, unknown, unreported); or high school location (large city, medium-sized city, small city or town, suburban, rural, and unreported) for North Carolina State University applicants between 2003 and 2008. An additional goal was to investigate the implications the new SAT reporting option, Score Choice, could have on future applicants' scores. The study examined SAT scores submitted to North Carolina State University between 2003 and 2008 for its applicants, totaling 151,901 individual SAT scores.

Overall, this study found there was a significant statistical relationship between repeatedly taking the SAT and ethnicity, gender, need for financial aid, and high school location. Furthermore, it found that the new SAT reporting policy, Score Choice, could indeed impact North Carolina State University undergraduate applicants' scores. On average, there was a within-person change in student SAT scores over time when looking at the North Carolina State University applicant pool between 2003 and 2008. Additionally, there was not a disadvantage to taking the SAT multiple times, as on average scores did not decrease across the selected population. This indicated that the more times students took the SAT, the greater

the potential to increase their score. The findings also indicated that there were differences in whether students' scores increased with each additional attempt of the SAT based on ethnic background, gender, and need for financial aid. The SAT score that students would report through Score Choice was found to be related to ethnicity, gender, financial aid, and location of a student's high school.

A paired samples *t* test revealed a statistically reliable difference between the score a student reports using Score Choice (the highest combination of math and verbal scores during one sitting) and the combination of an applicant's highest scores from each portion of the SAT to obtain the highest possible final SAT score ("superscore"). Therefore, it appears that the practice of taking the highest math and highest verbal score across all sittings of the SAT is the more beneficial to a student than the score submitted by Score Choice.

Repeated Test Taking on the SAT: The Effects of Ethnicity, Gender, Financial Aid, and High School Location

by
Lian Lynch

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Educational Research and Policy Analysis

Raleigh, North Carolina

2012

APPROVED BY:

_____	_____
Kevin P. Brady, Ph.D. Committee Co-Chair	Shevaun D. Neupert, Ph.D. Committee Co-Chair
_____	_____
Robert C. Serow, Ph.D.	Louis D. Hunt, Ph.D.

BIOGRAPHY

Lian Lynch (nee Oxenham) is originally from Leeds, England, and is the youngest child of Olwyn and William Oxenham. When Lian was 12 years old, her family immigrated to North Carolina, where she has lived ever since. After graduating from East Carolina University with a Bachelors of Science in Business Administration (2002) and MBA (2003), she worked in the department of Registration and Records at North Carolina State University for five years. In 2009, she started working in the Graduate School at North Carolina State University. Lian married Michael Lynch of Chapel Hill, North Carolina, in 2006. They currently reside in Holly Springs, North Carolina, with their daughter Emerson Kate.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

"Everyone may agree that testing can be a wedge, but some see the wedge forcing open the gates of opportunity while others see it as the doorstop keeping the gates tightly shut"(Office of Technology Assessment, 1992, p. 8).

Background of Study

The college and university admissions process in the United States has undergone numerous changes over the years. There is an existing quandary regarding the criteria today's college and university admissions officers use to predict applicants' academic success. The SAT Reasoning Test (which, for the purposes of this study, will be referred to simply as the SAT) has been the predictive method of choice. The SAT is produced and administered by the College Board, a not-for-profit association founded in 1900. College and university admissions offices use the SAT, along with additional factors such as high school grade point average (GPA), advanced placement (AP) scores, and extracurricular activities as predictors of an applicants' potential future academic success (Rigol, 2003). The combination of these criteria is often used to determine whether the student will be accepted to a particular college or university.

According to the Western Interstate Commission for Higher Education (2008), the total number of students graduating from high school, both public and private, exceeded 3.34 million during the 2007-2008 school year (p. 2). Of the students who graduated, approximately 1.5 million took the SAT (College Board, 2008). Prior to 2010, the College Board's policy was to send colleges and universities all scores on the SAT, regardless of

whether a student requested previous SAT scores be sent or not (Graves, 2008). Due to the fact that colleges and universities would often have multiple SAT scores for one individual student, there is a widespread university admissions practice of combining applicants' highest scores from each portion of the SAT to obtain the highest possible combined SAT score. However, a new policy named Score Choice was recently implemented in 2010 giving students the ability to mask their previous SAT test scores (Hatch, 2008; Kaplan, Henig & Philips, 2008; Miners, 2008). Score Choice permits students to send colleges and universities only their highest SAT scores as opposed to the previous practice of sending all of their SAT scores (Hoover, 2009; Miners, 2008). Rimer and Lewin (2008) state that "under Score Choice, students can choose their best overall SAT sitting to send to colleges and universities, but they will not be able to mix and match scores from different sitting" (p. 14).

The primary focus of this study is to determine whether there is a significant relationship between repeated test taking on the SAT and the ethnicity, gender, need for financial aid, or high school location of North Carolina State University applicants. The existing, peer-reviewed literature discusses the use of the SAT during the admissions process. However, with the introduction of Score Choice, postsecondary institutions will no longer receive as much SAT student score information as they previously did. The implication of such a practice will certainly change the way college admissions offices use this information.

Statement of Problem

The highest average student SAT score is often a pivotal element in determining whether a student is granted admission to a particular college or university. Students can take the SAT as often as they choose, contingent upon the frequency of the test offering,

which results in a wide range of SAT scores varying greatly from student to student. The cost of taking the SAT was $49 in 2010 for each administration of the exam and students with low socioeconomic status could obtain up to two fee waivers (Bunin, 2009). According to Vigdor (2003), the majority of students nationwide take the SAT at least twice.

Because many college admissions offices look at SAT scores as a primary factor in gauging student academic success, it is imperative to determine whether taking the SAT multiple times is advantaging a certain group of students while simultaneously disadvantaging other students. With the introduction of the Score Choice option by the College Board, a college or university's knowledge of the number of iterations of the test a student has taken will be eliminated. Therefore, the amount of data collected by the university to determine admissions is being reduced. According to Teare (2009):

> When I think about Score Choice, I remember the day that a great public university invited a group of counselors to a "fishbowl" admissions-committee meeting. We sat around the perimeter of the room while admissions officers presented and evaluated candidates. I felt as if I were watching pros play poker with GPA's, SAT scores, and other criteria submitted by applicants. Now, with Score Choice, which cards get into the committee's hands will supposedly be up to the candidates themselves. (p. A27)

In a time when colleges and universities across the country are attempting to determine whether they will comply with the Score Choice option, or choose to require that students report all scores, it is essential that colleges and universities deliberate with a full understanding of the impact of their actions on a wide range of students. This study will investigate the implications of implementing this new policy. It will also determine whether

there is a relationship between repeated test taking of the SAT and ethnicity, gender, need for financial aid, or high school location consequently determining whether this new policy will advantage or disadvantage certain groups of students applying to colleges and universities.

Purpose of Study

The purpose of this study is to determine whether there is a significant relationship between repeated test taking of the SAT and ethnicity (American Indian/Alaskan Native, Asian/Asian American/Pacific Islander, Black/African American, Mexican/Mexican American, Puerto Rican, Other Hispanic/Latino/Latin American, White, Other, and Unreported); gender (male, female, unreported); need for financial aid (yes, no, unknown, unreported); or high school location (large city, medium-sized city, small city or town, suburban, rural, and unreported) for North Carolina State University applicants between 2003 and 2008. The study also considers the implications of the new SAT reporting option, Score Choice, and its impact on future applicants' scores.

Significance of Study

The potential significance of this study on policy is paramount, since many colleges and universities are currently reevaluating the use of the SAT in light of the introduction of Score Choice. Score Choice has caused many colleges and universities to reflect upon their own admissions practices, while simultaneously determining whether they should still require students to submit all their SAT scores. The results from this study will provide information which should allow North Carolina State University and other similar institutions to make more informed policy decisions regarding the use of the Score Choice option.

North Carolina State University is a land-grant university and as such has an official mission and duty to serve the people of North Carolina, as the university was originally created by and receives significant funding from the state. Desmond and Goldman (2008) state that "land-grants exist to serve, protect and advance the public's values, ideals and interests" (p. 18). It is vital to comprehend the impact of Score Choice in order to assure that the purpose of the university is being fulfilled.

The literature on SAT scores tends to focus on whether the test is a reliable predictor of academic success at the college and university level. Score Choice has been discussed at length in peer-reviewed, scholarly journals; however, the focus of the discussions lacks empirical evidence and support. This gap in the existing literature is a primary focus in this study. For example, previous studies indicate that postsecondary degrees earned and income are positively correlated (Bowen & Bok, 1999). Therefore, the long-term effect of the criteria used by college admissions offices is much more significant than it initially appears.

Definition of Terms

1. **High stakes testing** - tests on which "high-stakes decisions with important consequences for individual students" are based (Heubert & Hauser, 1999, p.1).
2. **SAT** – previously called the Scholastic Aptitude Test, renamed in 1993 to the SAT I: Reasoning Test, renamed again in 2004 to the SAT Reasoning Test. However, for purposes of this study will be referred to as its most common name the SAT.
3. **SAT II** - subject tests that focus on one specific area. These tests aren't required by many universities but are often encouraged as supplemental information.

4. **AP tests** - are advanced placement tests focused on a particular subject. These tests are often used by universities to place students into the appropriate level of a course, or from a student's perspective to 'place out' of certain courses in university.
5. **Score Choice** - a new reporting option implemented by the College Board for the incoming university/college class of 2010, which gives students the ability to pick and choose the SAT scores they would like to report (Hatch, 2008; Hoover, 2009; Kaplan, et al., 2008; Miners, 2008).
6. **Math Score** – one of the three sections (critical reading, mathematics, and writing) on the SAT. The score a student can obtain on this portion of the test ranges from 200-800 points.
7. **Verbal Score** – will refer to verbal scores prior to March 2005 and critical reading scores post March 2005. The score a student can obtain on this portion of the test ranges from 200-800 points.

Research Questions

The purpose of this study is to determine whether there is a significant relationship between repeated test taking on the SAT and ethnicity, gender, need for financial aid, or high school location for North Carolina State University applicants. This study also investigates the implications of the new SAT reporting option, Score Choice, could have on future applicants' scores. The study focuses on the following five related research questions:

(1) On average, is there change in student SAT scores associated with repeated testing?

(2) Is there a point at which scores decrease with each additional attempt of the SAT?

(3) Does the relationship between repeated testing and SAT scores depend on:

(a) ethnicity;

(b) gender;

(c) financial aid; and/or,

(d) location of a student's high school?

(4) Does the point at which students' SAT scores decrease with each additional attempt of the SAT depend on ethnicity, gender, financial aid, location of a student's high school?

(5) Is a student's SAT score reported through Score Choice statistically related to ethnicity, gender, financial aid, or location of a student's high school?

Limitations of the Study

Since the data in this study is specific to North Carolina State University, the findings may not be generally applied to other colleges and universities. According to the North Carolina State University website:

> With more than 34,000 students and nearly 8,000 faculty and staff, North Carolina State University is a comprehensive university known for its leadership in education and research, and globally recognized for its science, technology, engineering and mathematics leadership. North Carolina State students, faculty and staff are focused. As one of the leading land-grant institutions in the nation, North Carolina State is committed to playing an active and vital role in improving the quality of life for the citizens of North Carolina, the nation and the world. (2011)

The data used in this study are limited to students who applied to North Carolina State University. It is worth noting that some of the applicants used in this study may have initially sent North Carolina State University their SAT scores but may have failed to send

additional scores if they decided against attending North Carolina State University. Also, applicants' whose initial SAT scores were high may have decided not to take the SAT additional times. Additionally, this study looks at SAT scores over time, but the amount of time is not specified. Therefore, two people could have taken the SAT three times. One may have taken it three times within a year, while the other may have taken it three times over three years.

The variable used in this study as a proxy for socioeconomic status, "financial aid need," has limitations. The question asked on the Student Descriptive Questionnaire (SDQ) was "Do you plan to apply for financial aid at any college?" It does not specify whether the financial aid is need-based or merit-based. North Carolina State University collects three variables on applicants that could be used for proxies of socioeconomic status: expected family contribution (EFC), parents' education level, and financial aid need. The first variable, expected family contribution, is derived from calculations from data provided on the Free Application for Federal Student Aid (FAFSA) and is collected by the Office of Scholarships and Financial Aid. According to the U.S. Department Of Education (2012), EFC "is a number used by your school to calculate the amount of federal student aid you are eligible to receive." The second variable, parents' education level, is collected on the North Carolina State University undergraduate admissions application; however, it is not a required question. The third variable is collected on the SAT survey. Of the applicants utilized in this study, only 23% had corresponding EFC data available, and just 33.54% reported their parents' education level on the North Carolina State University undergraduate admissions application. Due to the low response rates, it was determined that these would be inappropriate variables

to use in the MLM models. Therefore, the "need for financial aid variable" collected on the SDQ was used instead.

Assumptions

This study included the following assumptions: (a) The students in the study responded accurately on the self-reported survey questions which were submitted along with their SAT scores (the self-reported items are gender, ethnicity, whether they require financial aid, and the type of high school they attended); and (b) The population analyzed is reflective of future applicants of North Carolina State University.

Organization of Study

Chapter One frames the basis of this study, presenting a brief background, which includes the research problem, purpose, significance, definition of terms, key research questions and limitations. Chapter Two reviews literature and research associated with the SAT and Score Choice. Furthermore, Chapter Two delves into the discussions regarding university admissions practices and educational opportunity issues associated with the implementation of Score Choice. Chapter Three presents the methodology that will be used to conduct the study, a combination of multilevel modeling and regression. Chapter Four presents the results from the study. Finally, Chapter Five presents a discussion associated with the results and the conclusion.

Chapter 2: Literature Review

Introduction

The purpose of this study is to determine whether there is a significant relationship between repeated test taking on the SAT and ethnicity, gender, need for financial aid, or high school for North Carolina State University undergraduate applicants. A secondary purpose of this study is to investigate the implications of the new SAT reporting option, Score Choice, and its impact on future applicants' scores. The university admissions process is a topic frequently studied by educational researchers. This research is usually deliberated within the framework of student success or educational access. The body of research indicates that there is much dispute regarding whether a correlation exists between student test scores and success (Bracey, 2001; Sackett, Borneman, & Connelly, 2008).

This literature review will discuss the impact of high stakes testing on the college admissions process. For the purposes of this study, high stakes testing will be defined as tests on which "high-stakes decisions with important consequences for individual students" are based (Heubert et al., 1999, p. 1). The word "consequence" is prevalent in the educational literature regarding high stakes testing. Some of the positive consequences include promotion to the next grade level and graduation; there are also negative consequences that may occur, such as repeating a grade level and failure to graduate, thus being prevented access to higher educational institutions (Heubert et al., 1999; Nichols & Berliner, 2007).

A brief history of the SAT and the implementation of Score Choice will be explored, followed by a discussion of college and university admissions processes. Because this study

is investigating Score Choice and the data that will be analyzed is from a land-grant institution, the history of land-grant universities will be examined. Furthermore, as this study utilizes a multilevel modeling method, the dependent (SAT scores) and independent variables (ethnicity, gender, socioeconomic status, and high school attended) will be discussed.

High Stakes Testing

Where did high stakes testing begin? This phenomenon could be tracked as far as the launch of a Russian space-orbiting satellite, Sputnik, in 1957 (Amrein & Berliner, 2010). At this time, formal assessments were not prevalent in U.S. schools; however, the launch of this satellite by a rival country caused a demand from Americans for their educational system to be reevaluated: "Sputnik became an instant metaphor for the poor quality of U.S. schools" (Ravitch, 2000, p. 361). The fact that a rival country had essentially just defeated the United States in the pursuit of being the first to arrive in space instigated an examination into America's educational system and where it was failing (Amrein & Berliner, 2002; Amrein & Berliner, 2010; Ravitch, 2000). This uproar prompted President Dwight D. Eisenhower to sign the National Defense Education Act (NDEA) in 1958, an act that increased funding for math, science and foreign language programs and was designed to help America compete in the so-called "space race" (Ravitch, 2000; U.S. Department of Education, 2009). In discussing the impact of the NDEA, in his 1961 State of the Union address, Eisenhower stated:

> The National Defense Education Act of 1958 is already a milestone in the history of American education. It provides broad opportunities for the intellectual development of all children by strengthening courses of study in science, mathematics, and foreign

> languages, by developing new graduate programs to train additional teachers, and by providing loans for young people who need financial help to go to college. (Eisenhower, 1961)

Prior to this point, there was minimal governmental involvement in assessing student achievement. However, the demand for analysis of the educational system continued to increase, resulting in the establishment of the National Assessment of Educational Progress (NAEP) in the 1960s.

In 1965, the Elementary and Secondary Education Act (ESEA) was passed by Congress in a hope to improve education on a wider scale, not solely in the math and science sector (Frederiksen, 1994; Koretz, 2008; Thomas & Brady, 2005). Title I was an integral component of the ESEA and its purpose was to increase the academic achievement of underprivileged youths (Borman, Stringfield & Slavin, 2001). This initiative was part of Lyndon Johnson's "War on Poverty" initiative in which he established the Title I compensatory educational program in order to allocate federal funds to disadvantaged children:

> If a child was disadvantaged, the federal money would follow that child to whatever school he or she attended, public or private. However, there would have to be a "public trustee" administering the funds for all these children, and that trustee almost always would be the local public school district. (Borman, et al., 2001, pp. 3-4)

The implementation of this compensatory educational program marked the first time a formal evaluation program was required of primary and secondary education by the federal government, thus resulting in the establishment of the Title I Evaluation and Reporting

System (TIERS) in 1974 (Koretz, 2008, p. 55). Standardized testing was the method used by TIERS to accomplish this formal evaluation (Koretz, 2008):

> Neither NAEP nor TIERS imposed consequences on students or teachers based on test results. Nonetheless, in retrospect it seems that these two federal programs marked the onset of a sea change in educational testing in the United States. They signified the beginning of a fundamental shift in the goals of testing, from diagnosis and local evaluation to large-scale monitoring of performance and, ultimately, to test-based accountability. (Koretz, 2008, p. 55)

Minimum competency tests were introduced beginning in the 1970s (Boger & Orfield, 2005; Heubert et al., 1999; Koretz, 2008; Hamilton, Stecher, & Klein, 2002). These minimum competency tests were a direct response to implications that students were being awarded high school diplomas without actually obtaining an appropriate education (Heubert et al., 1999). These tests were administered as a mandatory requirement for graduation. They were also used to determine which students were permitted to move to the next grade level (Hamilton et al., 2002).

The *High Stakes Testing and High School Completion* report conducted for the National Board on Educational Testing and Public Policy by Clarke, Haney, and Madaus (2000) investigated the success of minimum competency tests. They used evidence from five studies to conclude that there was a negative correlation between graduation rates and high stakes testing programs. In 1986, the first study found that five of the ten states that had the lowest dropout rates did not utilize minimum competency tests; four of the other states used minimum competency tests for remediation purposes; and only one state used them for

accountability. Basically, none of the states used these tests for graduation or to advance a student to the next grade level (Clarke et al., 2000), whereas nine of the ten states with the highest dropout rates did use minimum competency tests to make decisions about graduation or grade promotion.

The second study, using data from the 1988 and 1990 National Educational Longitudinal Surveys, established that students who were in schools with a higher number of students from a lower socioeconomic status and schools that used minimum competency tests had higher dropout rates (4 to 6 % higher) between 8th and 10th grades compared to their wealthier counterparts (Clarke et al., 2000).

The third study found that students who generally achieved "moderately good grades (in the range of 1.5 to 2.5 on a 4-point scale)" but performed badly on the Florida high school graduation test were more likely to leave school (Clarke et al., 2000). The study also indicated that whether the student was a minority or non-minority did not impact the findings.

The fourth study indicated that there was a significant difference between minority and non-minority students when looking at the Texas Assessment of Academic Skills (TAAS) exit test; it found that after controlling for socio-economic status, academic track, language program participation, and school quality, Black and Hispanic students were three times more likely than non-Black and non-Hispanic students to drop out of school (Clarke et al., 2000). This study indicated that the use of the TAAS exit exam as a requirement to receive a diploma actually caused approximately 40,000 sophomores to drop out of school in 1993.

The final study gave evidence to indicate that retaining a student in a grade level instead of promoting them to the next grade level for the purpose of preparing them for a minimum competency exam does more harm than good. It affects a student's "sense of academic worth" and the results showed that this action had a greater negative effect on Black students than other ethnicities (Clarke et al., 2000). Furthermore, it increased the likelihood that a student would be held back in another grade. Although this study and others question the effectiveness of minimum competency exams, "in 1984, 40 states were actively involved in some aspect of minimum competency testing" (Winfield, 1990, p. 157).

The attitude that American schools were failing continued from the launch of Sputnik, through the implementation of Title I, through the minimum competency test movement, and was the basis for the formation of The National Commission on Excellence in Education by President Reagan. The National Commission on Excellence in Education was a group appointed to research the American education system and to provide the President with recommendations (Borek, 2008). This group published the 1983 report *A Nation at Risk: The Imperative for Educational Reform*, which gave the following warning: "Our Nation is at risk." It indicated that American students were not adequately educated in order to compete in the global market:

> The educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future as a Nation and a people. What was unimaginable a generation ago has begun to occur—others are matching and surpassing our educational attainments. If an unfriendly foreign power had attempted to impose on America the mediocre educational performance that exists today, we

> might well have viewed it as an act of war. (National Commission on Excellence In Education, 1983)

These words were disturbing to many and *A Nation at Risk* gained a great deal of attention. It addressed inequities in the educational system, stating that**:**

> All, regardless of race or class or economic status, are entitled to a fair chance and to the tools for developing their individual powers of mind and spirit to the utmost. This promise means that all children by virtue of their own efforts, competently guided, can hope to attain the mature and informed judgment needed to secure gainful employment, and to manage their own lives, thereby serving not only their own interests but also the progress of society itself. (National Commission on Excellence In Education, 1983)

The commission recommended five areas of improvement: content, expectations, time, teaching, as well as leadership and fiscal support. This publication again focused attention on educational assessment, resulting in many states appointing committees to research areas of potential improvement (Ravitch, 2000).

High stakes testing did not end in the twentieth century; in fact, in January 2002, the No Child Left Behind Act (NCLB) was signed into law by President Bush and has imposed educational high stakes testing at an even higher level. The purpose of this law is to ensure students are acquiring an appropriate education; however, there are high stakes associated with it as schools that fail to meet the standard could lose funding.

There are opposing views surrounding high stakes testing. Some believe that this type of assessment is beneficial and enhances learning, while others believe it results in negative

effects. A meta-analysis conducted by Sackett et al. (2008) established that standardized high stakes tests predict both professional and academic results. They found:

> (a) that tests of developed abilities are generally valid for their intended uses in predicting a wide variety of aspects of short-term and long-term academic and job performance; (b) that validity is not an artifact of SES; (c) that coaching is not a major determinant of test performance; (d) that tests do not generally exhibit bias by underpredicting the performance of minority group members; and, (e) that test taking motivational mechanisms are not major determinants of test performance in these high-stakes settings. (Sackett et al., p. 225)

However, there are also critics who believe that high stakes tests have a negative impact on student learning and potentially decrease student motivation (Amrein et al., 2002; Amrein et al., 2003). Heubert et al. (1999) made the following suggestions regarding high stakes tests in their book *High Stakes: Testing for Tracking, Promotion, and Graduation:*

1. Use the correct test for the situation. Tests are only useful when they are being used appropriately;
2. Tests are not perfect. They are not faultless measures of a student's knowledge, as circumstances may alter a student's performance (such as health on the day of the test);
3. Educational decisions should not be based purely on one test score; and,
4. Test scores cannot justify a bad decision. (p.275)

The suggestions made by Heubert et al. (1999) could assist high stakes test supporters and critics to accomplish their goals regarding student education. Critics tend to focus on

how tests are used inappropriately; therefore, the first suggestion by Heubert et al. (1999) would address this critique. Likewise, it would allow the supporters to continue to conduct their tests. Basically, if tests are used appropriately and are used for an appropriate purpose, they can be successful measurements.

The fact is that high stakes tests are prevalent throughout the American education system and do not appear to be diminishing. Such tests are not exclusively administered to K-12 students; these tests are also rampant in the university admissions process, thus affecting who gains admission to higher educational institutions. According to the College Entrance Examination Board (1999), over 85% of colleges and universities in the U.S. require an entrance examination. The vast majority of these schools require the SAT or ACT. Marchant (2004) explain that "[t]he SAT is not an achievement test, but an aptitude test designed to predict college achievement; however, because of its influence on college admissions decisions, it is also considered a high-stakes test" (p. 2).

A couple of phenomena have directly affected the use and intent of high stakes college entrance exams. Originally, many of these exams were formed with the intention of giving those other than the aristocracy access to a post-secondary education (Epstein, 2009). However, university rankings such as those reported in *U.S. News & World Report* (USNWR) have made universities view each other as direct competitors and "in the dominant ranking system, the SAT is at center stage" (Epstein, 2009, p. 10). The weight placed on such rankings is great because these published results shape university applicant pools, which in turn affect the make-up of universities, which influences the university

reputation, thus having a direct impact on alumni contributions (Webster, 2001). According to Webster (2001):

> An examination of the ranking criteria used by USNWR suggests that multicollinearity is pervasive. Multicollinearity in this instance refers to the degree to which changes in the value of one or more of the ranking criteria are related to, and are affected by, changes in one or more of the other ranking criteria. It could be argued, for example, that an institution's academic reputation is influenced by knowledge of the SAT scores of admitted students. Retention rates, enrollments, and alumni contributions are likely to be affected by academic reputation, which, in turn, would influence an institution's financial resources, per-student expenditures, faculty/student ratios, faculty compensations, etc. (p. 236)

The competition caused by USNWR rankings has increased the value placed on entrance exams such as the SAT, consequently transforming a simple entrance exam into a high stakes test.

University rankings have been altering the landscape of universities since their introduction. Universities are so driven by these publications that many practices have been altered in order to increase overall rankings. One such practice is that of merit scholarships. Scholarships were once provided on a needs basis, thus giving those who would otherwise not have the financial resources to attend college an equal opportunity. More recently, merit-based scholarships have been increasingly utilized in order to draw stellar students to attend a particular institution, regardless of the students' economic need (Thacker, 2005). The more merit-based scholarships are distributed, the more likely there is to be an increase in the

universities' overall scores, thus increasing their ranking (Thacker, 2005). However, this practice of increasing rankings is directly affecting those who used to receive need-based financial aid, as this funding is being reallocated, thus decreasing equity. Therefore, it is apparent that the SAT is not just a high stakes test from an individual student's perspective, but also from a higher education institutional perspective, as much is to be lost if a school's rankings decline.

The Role of the SAT

As previously stated, more than 85% of U.S. colleges and universities require some type of standardized admission test (College Entrance Examination Board, 1999). One of the most popular entrance exams is the SAT, which was created in 1926. This test is administered by The College Board, a non-profit organization, and is composed of three sections: critical reading, mathematics, and writing, with each section scored between 200 and 800 points. This high stakes test is often a pivotal point in the university admissions process. According to the National Association for College Admission Counseling's (NACAC) 2003-2004 annual report "The State of College Admission," the top factors in university admission decisions were as follows: grades in college prep courses, standardized admission tests, overall grade point average, and class rank (Hawkins, 2004, p. 5).

Controversy has surrounded the SAT from its origins to current use, as it has always had equally avid support and opposition. The SAT was originally named the Scholastic Aptitude Test, a name that was changed in 1993 to the SAT I: Reasoning Test. Carl C. Brigham created the SAT in 1926 based on Army IQ tests, following the publication of his book *A Study of American Intelligence*. Brigham dedicated a significant portion of the book

to the topic of ethnicity and how it related to intelligence. He also created what he referred to as the "race hypothesis," concluding that the integration of races would cause American education to decline (Brigham, 1923). Lemann (2000), who disagreed with the implications made in Brigham's book, acknowledges that Brigham was following the eugenicist views that were popular during that period of time. Although Brigham recanted his views and disowned his book years after its publication, there is still a stigma attached with the SAT today due to Brigham's involvement in its creation.

The SAT was originally believed to measure a person's intelligence, or innate ability, thus the name "Aptitude Test." However, when the test became more extensively used and the data clearly indicated students could increase their scores over time, the College Board changed the name of the test to a "Reasoning Test". The test was re-normed in April 1995 "which re-established a mean score of 500. To compare the 1996 scores with those of previous years, the College Board also re-centered the average scores of tests taken since 1972"(Geraghty & Guernsey, 1996). This re-centering caused some controversy, as some believed it was a tactic used to increase overall test scores. Others claimed "the worth of the test - especially its longevity as the primary gauge of college-bound seniors - is all but lost" (Innerst, 1996). Many believe the SAT contains bias (Crouse & Trusheim, 1988; Fleming & Garcia, 1998; Rooney & Schaeffer, 1998), while others believe it is a fair and valid test (College Board, 2011).

Due to controversy surrounding university admissions standardized tests, certain universities, such as the University of Washington, are beginning to take a more holistic approach to their application process. This particular school decided to look at other factors

surrounding the applicant, such as family income, the courses the student took in high school, clubs they participated in, among other factors (Frey, 2006). In February 2001, Richard C. Atkinson, president of the University of California system, encouraged his nine-campus system (containing upward of 170,000 students) to become the first public university system to drop the SAT requirement for their applicants (Gose & Selingo, 2001; Soares, 2007). More recently, Wake Forest University in North Carolina eliminated the SAT requirement; the president of university stated "this step away from standardized tests will help us and other institutions of higher education move closer to the goals of greater educational quality and opportunity" (Hatch, 2008).

Regardless of the controversy surrounding the SAT, it is still widely used by universities to predict a student's future success in university. According to its website, The College Board is currently "composed of more than 5,600 schools, colleges, universities and other educational organizations" (2011). The question being investigated in this paper is not whether the SAT should be used during the university application process, but whether there is a significant relationship between repeated test taking of the SAT and ethnicity, gender, need for financial aid, or high school location. A secondary goal is to investigate the implications the new SAT reporting option, Score Choice, could have on future applicants' scores.

Students have the option of taking the SAT as many times as they wish. There is a $49 fee each time it is taken and the test is offered seven times a year in the U.S. Students who meet the financial eligibility guideline have the option of receiving two SAT fee waivers which are initiated by the student's high school counselor. The test has been adjusted over

time, with the most recent change being the implementation of Score Choice. In 2005 the test was altered to add a third section—writing. Prior to March 2005, the SAT was composed of two sections: verbal and math. Currently there are three sections: critical reading, mathematics, and writing; the verbal scores prior to March 2005 correspond to critical reading scores post March 2005. Likewise, the math scores on the old test are comparable to the mathematics scores on the current test (College Board, 2009d).

Score Choice

Score Choice, a new reporting option implemented by the College Board for the incoming university/college class of 2010, gives students the ability to pick and choose the SAT scores they would like to report (Hatch, 2008; Hoover, 2009; Kaplan et al., 2008; Miners, 2008). According to the College Board's website (2009), this option was implemented in order to reduce stress from students taking the SAT. The College Board (2009) states the following regarding Score Choice:

1. Colleges and universities will only receive the scores that students send them. Scores will not be released for admission purposes without a student's specific consent.
2. Students who have not reported scores to any institutions will receive special e-mail reminders when traditional score-submission deadlines approach.
3. Score Choice is optional, and if a student does not actively decide to use it, all of his or her scores will be sent automatically at the time of ordering a score report.
4. Scores from an entire SAT test will be sent - scores of individual sections from different test dates cannot be selected independently for sending.

5. Students can send any or all scores to a college on a single report - it will not cost more to send one, multiple or all test scores.
6. Students are responsible for complying with the admissions requirements of the colleges, universities and scholarship programs to which they apply.

The sixth point is perhaps the most challenging, as colleges and universities do not have a mechanism for checking whether a student has sent the college or university all of his/her scores or just select scores. Furthermore, the College Board has stated that they have no intention of getting into the business of monitoring this action, as it is the responsibility of the student to comply with university policies (College Board, 2009b). In the past, universities would receive all scores a student had previously received on the SAT when the student submitted his/her current scores. Score Choice has changed this practice because, at no additional cost, students now have the opportunity to choose which scores they would like to report. The College Board cannot release students' scores without their permission, so even if they know a school requires all scores, a student could potentially choose to only send one. The implication of such a practice will certainly change the way college admissions offices use this information. In fact, some universities have already stated that they want students to send all SAT scores with their admissions application, thus asking students to not adhere to the Score Choice option. For example, Stanford Universities Undergraduate Admissions website (2011) states that "(a) applicants must self-report and submit *all* SAT scores and *all* ACT scores, and (b) applicants may not use the College Board's Score Choice feature or 'hide' any scores with either testing agency." Many colleges and universities use

the practice of combining a students' highest score on each section of the SAT to calculate the highest possible score (Rimer et al., 2008). This practice is referred to as "superscoring":

> Superscoring is done by some colleges for students who take the same admissions test more than once. Colleges look at the results of each test a student takes. If a student takes a test twice or more, the college will look at the subscores from each section (math, reading, etc) on every test date submitted, and take the highest for each. Taking a test multiple times gives kids a chance to raise a score in a particular area. (Straus, 2010)

If a student uses the Score Choice option they will forfeit the chance to combine scores, as they will only have reported one attempt of the test; thus, students may actually be putting themselves at a disadvantage.

Burke (1997) found that increased test scores are possible based purely on retesting, without coaching, as the experience of taking a test more than once can improve scores. This point was supported by the College Board's Vice President of Research, when he attributed the decrease in SAT scores during 2006 partially to the fact that there was a decrease in the number of students who took the test two or more times that year (Camara, 2006). According to the College Board (2008):

1. 55% of juniors taking the test improved their scores as seniors.
2. 35% had score drops.
3. 10% had no change.
4. The higher a student's scores as a junior, the more likely that student's subsequent scores will drop.

5. The lower the initial scores, the more likely the scores will go up.
6. On average, juniors repeating the SAT as seniors improved their combined critical reading, mathematics, and writing scores by approximately 40 points.
7. About 1 in 25 gained 100 or more points on critical reading or mathematics, and about 1 in 90 lost 100 or more points.

The information regarding changes in scores over time will not be evident to admissions offices if a student chooses to use the Score Choice option. Students' previous scores would be masked, so a university could potentially compare a student's score on their fourth attempt on the SAT to another student's score on their first attempt of the SAT and would never know that this was occurring.

This policy change is not the first of its kind. The ACT, another popular university admissions exam, has a similar procedure. The ACT only sends universities scores a student has requested to be sent, much the same as Score Choice. The College Board previously permitted a score choice option on the SAT II beginning in 1993; however, they dropped this policy in 2002, stating that students who could afford to take multiple iterations of the tests had an unfair advantage over their poorer counterparts (Bowman, 2002; Rimer et al., 2008). Many believe that this situation will be repeated on the SAT with Score Choice advantaging the wealthy (Hoover, 2009; USA Today, 2009). Others believe it is a positive change (Bunin, 2009; College Board, 2009b). The results of this study will address these opposing views.

Postsecondary Admissions and Land Grant Universities

The college and university admissions process varies greatly. However, the criteria for land grant universities differ from other types of higher educational institutions. The

mission of a land grant university is not to just attract the best and the brightest students, but also to meet the needs of the citizens within the state. During the period when land grant universities were established, more than half the workforce was in farming, whereas now less than 1% of the U.S. population claim farming as their profession (Brannon, R. Morgan Dean, & J. Morgan Dean, 2002; U.S. Environmental Protection Agency, 2009). Because of this shift from a more agricultural workforce, the mission of these institutions has morphed over the years to meet the changing needs of the people within the state. When discussing the mission of Cornell as a land grant university, Brannon et al. (2002) stated that the "reality of the land grant university is change. The core and soul of the land grant university is the mutual informing of research and outreach education so that our emerging scholarship is relevant to the needs and problems that our citizens face" (p. 1).

In President Barack Obama's speech at the National Academy of Sciences Annual Meeting, April 27, 2009, President Obama stated:

> A few months after a devastating defeat at Fredericksburg, before Gettysburg would be won, before Richmond would fall, before the fate of the Union would be at all certain, President Abraham Lincoln signed into law an act creating the National Academy of Sciences -- in the midst of civil war.
>
> Lincoln refused to accept that our nation's sole purpose was mere survival. He created this academy, founded the land grant colleges, and began the work of the transcontinental railroad, believing that we must add -- and I quote – 'the fuel of interest to the fire of genius in the discovery... of new and useful things.

The rise of the U.S. public land-grant postsecondary institution had an equity-based focus, as the original purpose of these institutions was to widen the access Americans had to higher education. The Morrill Acts of 1862 and 1890, named for the Vermont Senator Justin Morrill, gave 30,000 acres of public land to each senator and representative in Congress (Greenleaf, 1934). According to the National Association of State Universities and Land-Grant Colleges (NASULGC) (1995), the intention was for this land to be used to create institutions that would "teach agriculture, military tactics, and the mechanical arts as well as classical studies so that members of the working classes could obtain a liberal, practical education" (p. 3). It was a way to provide applicable education to citizens of the state in which the land-grant institution was established.

Currently, there are two land-grant universities in North Carolina: North Carolina State University and North Carolina A&T State University. North Carolina State University was part of the original land-grant universities established with the first Morrill Act of 1862. North Carolina A&T State University was established in 1890 when the second Morrill Act permitted land-grant institutions to admit students of all ethnicities. The land-grant universities at that time had the option of permitting minority students into their university, or a separate institution could be established specifically for minority students. Most southern states, North Carolina included, opted to create a separate institution, thus North Carolina A&T was established (NASULGC, 1995).

The Hatch Act of 1887 funded agricultural experiment stations at each land-grant institution as a response to the need for agricultural research, and the Smith-Lever Act of 1914 gave funding for cooperative extension services. However, the land-grant universities

of 1890 were not privy to this funding. Instead, additional funding was given to these institutions through other means such as the Evans-Allen program (NASULGC, 1995).

Table 2.1

Land-Grant Timeline

Year	Act	Description
1862	Morrill Act	Granted 30,000 acres of land to each senator and representative
1887	Hatch Act	Experiment stations created
1890	Morrill Act	Additional funding given to the original land-grant institutions. 17 predominantly Black institutions created in southern states
1914	Smith-Lever Act	The Cooperative Extension System was created
1994	Equity in Education Land-Grant Status Act	29 Native American tribal colleges and universities were given land-grant status. Provided funding for these institutions to conduct non-formal/outreach activities to assist Native American people and provide services to their communities

McDowell (2001) joked at the end of his book, *Land-grant Universities and Extension into the 21st Century*, that when he stated he was writing a book discussing land-grant universities into the future, his friend asked, "Do they have a future?" (p.191). He goes on to talk about the need for land-grant universities to again become the people's universities, arguing that land-grants need to become more engaged with the community. North Carolina State University has developed an office of Extension, Engagement and

Economic Development to accomplish just that: “The mission of North Carolina State University Extension, Engagement and Economic Development is to partner the resources of the university and communities to produce mutual benefits” (North Carolina State University, 2009). The programs developed by this office are done so with the mission of creating a mutually beneficial relationship between the residents of North Carolina and North Carolina State University. Furthermore, the development of the Centennial Campus at N.C. State has increased the mutually beneficial relationship between the university and the people of North Carolina. According to the N.C. State Centennial Campus website (http://centennial.ncsu.edu), “North Carolina State University’s Centennial Campus is an extraordinary success story – a research park and campus providing its corporate, governmental and non-profit partners unusually close proximity to world-class research and a highly educated workforce – all in an amenity-rich environment” (North Carolina State University, 2009). Although North Carolina State University has made efforts to increase relationships between the residents of North Carolina and the university, high stakes tests are still factored into the admissions requirements.

According to *Returning to Our Roots: Executive Summaries of the Reports of the Kellogg Commission on the Future of State and Land-Grant Universities*, “Educational opportunity in America is still far from equal” (p. 7). The report further discusses how the purpose of land-grant institutions was to “open opportunity and broaden access to higher education” and how “today, this historic commitment must encompass the different educational needs of many different kinds of students coming from different and ever-more diverse backgrounds” (*Kellogg Commission,* 2001, p. 7).

The Relationship between Score Choice and North Carolina State University

North Carolina State University is the largest university in the state of North Carolina. It is a publicly supported and highly selective institution. North Carolina State University also offers admission to an Agricultural Institute, which gives students the opportunity to participate in a two-year technical program in order to obtain an associate of applied science degree in an agricultural field.

Another item that should be taken into account when assessing North Carolina State University is the current economic climate. The current recession has required some students who would have previously applied to private institutions to apply to selective public universities, thus potentially altering the applicant pool (Foderaro, 2009). This study will examine the implications the new SAT reporting option, Score Choice, could have on future North Carolina State University applicants. Currently, the North Carolina State University admissions office requires applicants submit an SAT score or an ACT score for admissions consideration. It will be interesting to see whether the Score Choice option could potentially change the dimension of the incoming freshman class.

Student Demographic Variables

In recent years, there has been much discussion regarding diversifying the student body in higher educational institutions. Justice Sandra Day O'Connor was the "swing vote" regarding the ruling for the Supreme Court case *Grutter v. Bollinger* in 2003, where she acknowledged "that race-conscious admissions policies are constitutional because they serve the compelling state interest of promoting diversity and its associated educational benefits" (Schmidt, 2010). Bowen et al. (1998) found that graduating students "regularly stress that

much of what they gained from their educational experience came from what they learned from their fellow students" (p. 24). The importance of diversity is evident; however, American universities are far from evenly diverse.

The report released by the Department of Education (2004) entitled *Achieving Diversity: Race-neutral Alternatives in American Education* discussed two overarching ways in which American educational institutions are attempting to increase the diversity of their student body. The first is through "developmental approaches;" the report refers to developmental approaches as those that "enrich the pipeline of applicants prepared to succeed in any academic setting" (Department of Education, 2004, p. 10). The second is through "admissions approaches," which include items such as socioeconomic preferences, lottery systems, or comprehensive reviews (Department of Education, 2004, p. 67). Admission approaches are defined in the report as "approaches that focus on the admissions process, attempting to diversify the range of students who are admitted into educational institutions" (Department of Education, 2004, p. 67). The focus of this research paper is on admission approaches. Previous research has indicated that ethnicity, gender, socioeconomic status, and the high school a student attended are correlated to the likelihood of that student being accepted into certain universities.

Ethnicity. In 1954, the primary goal of the Supreme Court ruling in *Brown vs. Board of Education* was to provide equal access to educational opportunities for all students. However, 55 years later, the role of ethnicity in regards to college admission is still a significant topic of debate (Carter, 1992; Wightman, 1997; Williams, 2007)**.** Antonio (2003) stated "racial and ethnic diversity influences an entire campus, including its culture, its

values, and eventually its ethos" (p. 14). However, Blacks and Hispanics are still underrepresented in higher educational institutions (Carnevale & Rose, 2003). According to Carnevale et al. (2003), in 1995 the freshman classes of highly selective universities, as defined by Barron's selectivity measure, were composed of approximately 6% African Americans and Hispanics (p. 10). This gap is also evident at other higher educational institutions, not just highly selective ones:

> White students are more likely than average to graduate from high school and enroll in college, while Black and Hispanic students are less likely than average to do the same. The gap is most pronounced at four-year colleges, where Blacks and Hispanics constitute only 17 percent of the undergraduate population, even though together they constitute 31 percent of the national college-age population. (Hawkins, 2004, p. 4)

Herrnstein and Murray (1994) analyzed SAT data obtained from two campuses within the University of California system and the University of Virginia and discovered that the "gap between minorities and Whites among freshmen at state universities may be larger than at the elite private schools" (p. 453). Further analysis indicated additional ethnicity gaps. They found that in 1993, more than 7000 White students earned a score greater than 700 on the verbal portion of the SAT, where as a mere 210 Black students and 234 Latinos obtained this same score (Herrnstein & Murray, 1994, p. 454). The use of SATs in the admissions process has implications on the mission of land-grant institutions, given that there is apparent inequity between ethnicities. Furthermore, the adoption of Score Choice could potentially expand this gap.

Some universities have implemented affirmative action in the hopes of increasing diversity. The 1978 Supreme Court trial *Regents of the University of California v. Bakke* determined that affirmative action was permitted in university admissions processes as long as universities looked at other factors in addition to a student's race (Cornell University Law School, 2009a; University of Missouri-Kansas City School of Law, 2009). However, the 1996 Hopwood decision, which was a ruling made at the Fifth Circuit Court of Appeals (encompassing Alabama, Louisiana, and Texas), contradicted that of *California v. Bakke*, determining that race could not be used as an explicit factor in decisions regarding admissions to public universities in these states (University of Texas at Austin, 2008). Similarly, California's Proposition 209 (1996), Washington's Initiative 200 (1998), Michigan's Proposal 2 (2006) and Nebraska's Initiative 424, voter-approved anti-affirmative action laws, prevent the states from discriminating or giving racial preferences (West-Faulcon, 2009). With these laws implemented California, Washington and Michigan, university systems made the decision not to use race as a factor in the admissions process (West-Faulcon, 2009). The decision to put into practice a race-free admission process has caused the number of minority students attending college in these universities to decrease (West-Faulcon, 2009). Furthermore, this practice at the University of California Berkeley caused a significant decrease in minority student representation, thus prompting a group of minority students to sue the university in 1998 (*Rio v. Regents of the University of California*).

In 2000, Florida implemented the One Florida initiative, which also banned the use of ethnicity and gender during university admissions decisions. Part of One Florida, Talented 20

admissions program, guaranteed the top 20% of graduating high school classes in Florida admission to the state's public universities (Blair, 2000; Black Issues in Higher Education, 2002; Rapp, 2000). The thought process behind this controversial initiative was to encourage student achievement and to increase minority college admissions. However, the feedback received was far from positive, as many believed that in fact it would accomplish quite the opposite by decreasing equity and minority admissions (Rapp, 2000). One Florida was challenged by the NAACP, but "an appeals court has ruled the NAACP does not have standing to challenge rules eliminating racial and gender preferences in university admissions" (*Black Issues in Higher Education,* 2002, p. 10).

In 2003, Gratz and Hamacher filed suit against the University of Michigan claiming they were discriminated against during the undergraduate admissions process. Their basis was that the University of Michigan used a point system to establish who would and would not be admitted. The point system awarded more points to minorities, due to their ethnicity. As Gratz and Hamacher were both Caucasians they did not receive the additional points and therefore believed that they had been unfairly discriminated against on the basis of race (Cornell University Law School, 2009b). A similar suit, *Grutter v. Bollinger et al*, was filed within the same time frame against the Michigan University Law School,. The court ruling in both these cases was similar to that of the *Regents of the University of California v. Bakke* case; they both concluded that affirmative action was tolerable if ethnicity was not the only factor considered and the primary result was to diversify the student body (McGaghie & Kreiter, 2005). According to Edwards (2004), "*Brown* sought to foster equality through

integration by prohibiting forced segregation on the basis of race; *Grutter* aims to foster equality by permitting forced racial integration to achieve diversity" (p. 947).

It is evident that differing decisions have resulted from court cases regarding the role of ethnicity in university admissions processes. It is important that we investigate whether there is a significant relationship between repeated test taking on the SAT and ethnicity in order to see whether the introduction of Score Choice could increase diversity or decrease diversity within higher educational institutions. Some universities are not requesting ethnicity be disclosed at the time of the admission application; if they also do require the SAT, it is vital that they fully understand the impact the new Score Choice policy may have in order to make informed decisions.

Gender. When discussing gender and education, Title IX of the Education Amendments of 1972 is paramount. The purpose of Title IX was to give the same advantages or opportunities to females in regards to education and sports, stating that "No person in the United States shall, on the basis of sex, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any education program or activity receiving Federal financial assistance" (U.S. Department of Labor, 2009). Title IX has been academically beneficial to both males and females (Erickson, 2007). This being said, there are still differences that exist between males and females when analyzing university admissions tests.

Mau and Lynn (2001) found that males scored significantly higher on college entrance tests; however, females scored significantly higher grades during their college career. Their findings also indicated that females were more successful than their male

counterparts on assessments based on coursework, whereas males tended to score better on cognitive assessments than females. University admissions tests are found to under-predict the ability of females in university (FairTest, 2007; Young, 1991):

> Standardized tests like the SAT and GRE tend to especially penalize women and many minority students. Females tend to do worse than males on standardized tests but consistently earn better grades than males. Researchers consistently find that adding test scores to the admissions equation results in fewer women and minorities being accepted than if their academic records alone were considered. (Sacks, 1999, p. 7)

FairTest (2007) attributes the following to causing the gender gap on the SAT: the multiple-choice format, the penalty that is given if a student guesses an answer incorrectly, and the time limit on the exam. They state, however, that the testing company's explanation of this score discrepancy is due to the number of females that take the test, as they outweigh the male population. According to Young and Fisler (2000), "Over the past 20 years, the differences in the average scores between men and women on the Verbal (SAT-V) and Mathematical (SAT-M) sections of the test have steadily increased" (Young et al., 2000, p. 401). Young et al. (2000) conducted a study of 69,284 high school seniors who took the SAT in November of 1990. Using the self-reported background information on the SAT, they determined that some of the score differences that were evident between males and females may have been due to other characteristics besides gender:

> As with other tests of cognitive abilities, a number of theories have been advanced that purport to explain sex differences on the SAT, including: (1) biological

> differences in abilities, (2) differences in classroom experiences, and/or (3) bias against women in the development and administration of the test. Sex differences in observed SAT scores are often interpreted as the result of one or more of these factors. However, the current scholarly literature provides little support for any of these theories; in the absence of strong evidence that they can explain differences in test scores, it is necessary to consider other factors such as student background characteristics. (Young et al., 2000, p. 402)

For example, the males in their sample tended to have parents whose socioeconomic status was higher than that of the female parents. Also, the ethnicity of males differed to that of the females in the study. The males contained a higher number of Asian-Americans, the ethnicity that has the highest average SAT math score, when compared to other ethnicities.

Although females tend to score lower on the SAT, according to the U.S. Department of Education (2011) they graduate at a higher rate to their male counterparts. For the past twenty years, women have surpassed men in earning bachelor's degrees (Greene & Greene, 2004): "Division I NCAA universities have reported a 54 percent female to 46 percent male ratio in the undergrad body" (Greene et al., 2004, p. 27). North Carolina State University differs in this aspect from other universities; of the 25,255 undergraduate students attending North Carolina State University in fall 2009, the male undergraduate population (55.5%) surpassed that of the female population (44.5%) (North Carolina State University, 2009). If North Carolina State University chooses to use Score Choice, the question is whether these numbers will alter.

Socioeconomic status. Socioeconomic status is defined by Heller (2002) as "a composite measure that reflects students' family income as well as their parents' level of educational attainment and occupational status" (p. 100). When discussing socioeconomic status and education, the most influential research is that of *Equality of Educational Opportunity*, more commonly referred to as The Coleman Report (1966). This report, published in 1966, analyzed surveys of 570,000 students and 60,000 teachers with the expectation that the outcome of the research would show school quality had the most significant impact on a student's academic success (Hoff, 1999; Towers, 1992). However, the actual outcome of the study did not match the expected outcome: the main discovery was that a school's socioeconomic makeup and a student's family background were the best predictors of a student's success (Hoff, 1999; Towers, 1992). Another study conducted in 1997 on Title I came to the same conclusion as the Coleman Report: that student success is strongly correlated with a school's socioeconomic background (Hoff, 1999).

Social mobility is a democratic ideal upon which America prides itself. However, the American education system does not always provide students with social mobility. According to Carnevale et al. (2003), the ability to gain entrance to highly selected universities is skewed by race and ethnicity; yet even race and ethnicity do not have as much impact on admissions as socioeconomic status does. According to Cabrera and La Nasa (2001):

> Applying to a 4-year institution appears to be particularly challenging for the lowest-SES students. Only 65.5% of the college-qualified, high school graduates from lowest-SES backgrounds actually apply to a 4-year institution. This rate is 16% and

22% below the national rate of similarly qualified eighth graders and the rate for students from high-SES. (p. 121)

Carnevale et al. (2003) conducted a study using the National Educational Longitudinal Study (NELS:88). They used the *Barron's* selectivity measures, which uses factors such as student's SAT/ACT score, high school class rank and GPA to categorize universities into six levels from most selective to least selective. The specific terminology used is: Most Competitive, Highly Competitive, Very Competitive, Competitive, Less Competitive, and Noncompetitive (Carnevale et al., 2003, p. 8). Carnevale et al. (2003) simplified *Barron's* selectivity measures into four tiers:

1. Tier 1 consists of the "Most" and "Highly" competitive institutions. Students are usually in the top 35 percent of their high school class, have an average high school GPA of B or higher, and score about 1240 or higher on the SAT (the SAT consisted of two sections with a maximum score of 1600) or above 27 on the ACT.
2. Tier 2 consists of the "Very" competitive institutions. Students are usually in the middle of their class, have an average high school GPA B- or higher, range of 1146 to 1238 on the SAT or 24-26 on the ACT.
3. Tier 3 consists of the "Competitive" institutions. Students usually have an average high school GPA of C or better, score above 1000 on the SAT or above 21 on the ACT.
4. Tier 4 consists of the "Less" competitive and "noncompetitive" institutions. Students usually have an average high school GPA C or less, score below 1000 on the SAT or below 21 on the ACT. (p. 8)

It is evident from their findings that there is an inverse relationship between SES and the type of higher educational institution a student attends. The higher the SES of a student, the more likely the student will attend a highly selective university. The significance of these findings is immense, as students who attend highly selective universities tend to have a higher graduation rate and are more likely to be accepted into graduate or professional schools after graduation (Carnevale et al., 2003).

Herrnstein and Murray (1994) analyzed data from the National Longitudinal Survey of Youth (NLSY) in their bestselling book *The Bell Curve*. They used the data from the NLSY to analyze socioeconomic status and IQ in relation to variables such as poverty, schooling, unemployment, idleness, and injury. The findings indicated that there was a positive correlation between economic status and societal success; there was also a correlation between ethnicity and success. However, a contested finding was that IQ had a greater impact on success over time than socioeconomic status did.

The SAT has been a source of contention regarding university admissions. The introduction of Score Choice on the SAT has caused some to question whether it will only benefit the affluent (*USA Today*, 2009). Bunin (2009) does not believe that this will be the case, as he states that repeating the SAT does not guarantee an increase in a student's score. Therefore, hiding previous iterations of the SAT may not negatively affect only the lower SES student. Furthermore, Bunin (2009) points out those students with a low SES have the ability to get fee waivers, and thus can take the SAT twice at no charge. Nichols et al. (2008) disagree with Bunin's notion, stating:

> We think that the unnoticed slipping of high-stakes testing into our culture has taken place partly because it hits our poorest, most racially diverse student body hardest and thereby forces the kind of education on the children of the poor that ensures that they cannot compete successfully with the children of the wealthy. The drill-and-test-prep education we see in schools for the poor does not prepare them for knowledge workers' jobs or for the best universities. This makes room for the children of the more privileged. Since the status of children from advantaged families is preserved though high stakes testing, it is easy for these folks to defend their use. (p. 45)

High school attended. Rank (2005) described how certain K-12 schools are disadvantaged due to their location, resulting in the students within those schools receiving less funding than others: "About half of all Black and Latino students attend schools in which three-quarters or more students are poor. Only 5% of White students attend such schools" (Bhargava, Frankenberg, & Le, 2009, p. 14). Public schools are primarily funded by local taxes, such as property tax. Therefore, communities of low wealth do not have funding equivalent to their more affluent counterparts. Lack of funding for these schools results in diminished resources, such as fewer current text books, less access to technology, and higher challenges associated with recruiting and retaining skilled teachers (Rank, 2005, p. 141). The influence of the high school a student attends is considerable. The Coleman report established that the socioeconomic level of the student body within a school was an influential factor on a student's academic success (Towers, 1992). Therefore, if a student attends a high school in a wealthy area of town in which the majority of students are from the

highest SES quartile, then that student has an advantage over a student attending a school in a poor area of town in which the students tend to be from the lowest SES quartile.

In regards to the SAT, Chenoweth (1998) states “The most striking differences the College Board found in average SAT scores was between suburban schools and schools in both urban and rural areas. Suburban schools logged in average scores that were 15 to 20 points above the average, whereas urban and rural schools had average scores that were between 9 and 17 points below the average” (p. 24).

Summary

Chapter Two summarized the literature surrounding the high stakes testing movement. Additionally, the literature review explored the history surrounding the SAT and Score Choice and discussed the variables that will be analyzed in this study, including ethnicity, gender, financial aid, and location of a student’s high school. Chapter Three will discuss the methodology for data collection and analysis that will be used to analyze the data, a combination of multilevel modeling and regression.

Chapter 3: Methodology

Introduction

The purpose of this study is to determine whether there is a significant relationship between repeated test taking on the SAT and ethnicity, gender, need for financial aid, or high school location for North Carolina State University applicants. A further aspect is to investigate the implications the new SAT reporting option, Score Choice, could have on future applicants' scores. Therefore, this study focuses on the following five related research questions:

(1) On average, is there change in student SAT scores associated with repeated testing?

(2) Is there a point at which scores decrease with each additional attempt of the SAT?

(3) Does the relationship between repeated testing and SAT scores depend on:

(a) ethnicity;

(b) gender;

(c) financial aid; and/or,

(d) location of a student's high school?

(4) Does the point at which students' SAT scores decrease with each additional attempt of the SAT depend on ethnicity/gender/financial aid/location of a student's high school?

(5) Is a student's SAT score that would be reported through Score Choice related to ethnicity, gender, financial aid, or location of a student's high school?

The collection and measuring of the variables will be accomplished using two instruments, the College Board's electronic score report and the College Board's SDQ.

Multilevel linear modeling coupled with regression will be the analytic methods of choice. This chapter will describe the methodology that will be used to test the research questions and will be divided into five sections: (1) Setting, (2) participants, (3) instrumentation, (4) data collection, and (5) data analysis.

Setting

North Carolina State University is the setting for this particular study. It is located in Raleigh, North Carolina. It is the largest university in the state and is one of the 16 universities within the University of North Carolina system. Table 3.1 shows the average SAT scores of first time freshman at the 16 universities in the North Carolina University System.

The population of North Carolina in 2010 was approximately 9.5 million, with 25.8% of the residents aged twenty five and above having a bachelor's degree or higher and 83% of residents aged 25 and above having a high school diploma (U.S. Census Bureau, 2009). The State is a mixture of rural and urban areas; however recent years have seen the greatest increase in urban and suburban areas, such as Raleigh, Charlotte and Durham (Stuart, & Baum, 2005).

Table 3.1

Fall 2008 Average SAT of First Time Freshman

	SAT Verbal	SAT Math	Total
Appalachian State University	576	587	1,163
East Carolina University	502	523	1,025
Elizabeth City State University	414	421	835
Fayetteville State University	425	440	865
NC A&T State University	452	468	920
North Carolina Central University	421	432	853
North Carolina State University	569	607	1,176
UNC Asheville	586	572	1,158
UNC - Chapel Hill	643	658	1,301
UNC Charlotte	513	543	1,056
UNC Greensboro	516	523	1,039
UNC Pembroke	455	473	928
UNC Wilmington	568	588	1,156
UNC School of the Arts	568	543	1,111
Western Carolina University	511	528	1,039
Winston-Salem State University	432	442	874
UNC Total	532	551	1,083

Note. SAT scores are rounded. From "The University of North Carolina a multi-campus university: Facts and figures," 2009, http://www.northcarolina.edu/about/facts.htm

Participants

The data used in this study were collected from the Office of Admissions at North Carolina State University. The university has a population of approximately 30,000 students with roughly 6,000 new admits annually. Table 3.2 indicates the total number of students

enrolled at North Carolina State University from 2003 to 2008 as reported by the NCSU Office of University Planning and Analysis.

Table 3.2

Total Enrollment at North Carolina State University: 2003 Academic Year – 2008 Academic Year

	All Students	Undergraduate	Graduate	New Freshmen	New Transfer	New Graduate
2003	29,854	22,971	6,883	3,931	1,089	1,835
2004	29,957	22,754	7,203	3,957	1,066	1,894
2005	30,149	22,879	7,270	4,375	1,061	1,929
2006	31,130	23,730	7,400	4,693	1,108	1,914
2007	31,802	24,145	7,657	4,907	1,062	2,100
2008	32,872	24,741	8,131	4,804	1,093	2,284

Note. From "University Planning and Analysis: Enrollment Data," 2009, http://www2.acs.ncsu.edu/UPA/enrollmentdata/index.htm

The original sample consisted of all SAT scores which were submitted to this university between 2003 and 2008. Therefore, many of the students who submitted SAT scores did not actually apply for admission to the university. In order to assure students' anonymity, their social security number was removed and replaced with a student number (beginning with student #1 and going through student #151,901).

With the aim of analyzing the most accurate data, duplicate scores were eliminated along with scores that did not contain accurate test dates. As many submitted scores multiple times, the original file contained scores of duplicate entries: Multiple records are created when a student submits more than one request to send a report to your institution within the same

reporting period, or if a student takes the SAT and/or Subject Tests more than once and does not provide consistent identifying information (College Board, 2005). Prior to analyzing these data, it was important to eliminate duplicate records to prevent any potential of the data being skewed or inaccurate.

SAT scores were paired with information on the student's ethnicity, gender, financial aid, and location of a student's high school to analyze whether the relationship between time and SAT scores depended upon these variables. Furthermore, this was used to determine whether a student's highest SAT score (the score that would be reported through Score Choice) related to ethnicity, gender, financial aid, or location of a student's high school.

Instrumentation

The College Board's electronic score report was the instrument used to collect the SAT raw scores. When students register to take the SAT, they can indicate up to four schools to which the results of the test will be sent. Supplementary schools may be chosen for an additional fee. Once the student sends these scores, the university can download the scores into their school specific interface and use them as they choose in the admissions process. The process for reporting SAT scores is close to flawless, given that the student provides sufficient identifying information.

Score reports from tests administered prior to March 2005 contained slightly different information from those administered after March 2005. According to the College Board (2009d), prior to March 2005 score reports would contain verbal (known as critical reading after March 2005) and math scores, both on the 200-800 scale. No writing score was reported, as the writing section premiered in March 2005. Students who took the test again

after March 2005 could compare their verbal scores to their critical reading scores, and math scores from the old test with mathematics scores from the new test (College Board, 2009d).

The instrument used to obtain the demographic information for this study was the self-reported questionnaire administered by the College Board at the time the student registered for the SAT. Cassady (2001) found self-reporting of GPA and SAT scores to be significantly related to students' actual scores, when testing a group of 89 undergraduate students at a Midwestern university. In Freeberg's (1988) analysis of the questionnaire, he found it possessed an appropriate level of accuracy:

> As a self-report instrument, the Student Descriptive Questionnaire (SDQ) has, since 1971, enabled college applicants to describe a range of interests, activities, plans, and abilities in both academic and nonacademic areas. This information, obtained as an optional part of the College Board's Admissions Testing Program and forwarded to colleges designated by the student, is intended to supplement the information base and thus improve the quality of decisions made by college admission personnel, school counselors, and program planners. (Freeberg, 1988, p. 1)

There were a variety of socio-demographic items included on the questionnaire such as students' housing plans, employment status (part-time or full-time), religion, citizenship, and more. Based on previous findings, this study will focus specifically on gender, ethnicity, financial aid need, and location of high school.

Level 1 Variable (Within-Student)

Time. For purposes of this dissertation, time will be the variable referring to the number of times a student has taken the SAT. It will be coded as 0 = first time a person has taken the SAT, 1 = the second time, 2 = the third time etc.

Quadratic time. In order to determine whether there is a point at which scores will decrease with each additional attempt of the SAT, quadratic time will be examined. This will show the quadratic trajectory of scores over time.

Level 2 Variables (Between-Student)

All the level 2 variables were self-reported by the students when they registered for the SAT. The variables are each divided into groups corresponding with the possible answers given on the SDQ.

Ethnicity. According to Schmidt (2007), "many of the objections to the SAT stem from the tendency of Blacks, Hispanics, and Native Americans to score lower than Whites, even when compared to people at the same income level" (p. 151). Schmidt (2007) found the average score of a White student tends to be approximately 200 points greater than their Black counterpart. This study will attempt to determine whether these differences could be due to the number of times a student took the test.

There is also a debate regarding whether test-preparation programs do indeed help students increase their scores and whether certain ethnicities have greater accessibility to these programs. According to Alderman and Powers (1983), "to some degree standardized tests may also reflect individual differences in prior test-taking experience" (p. 71). In order

to determine whether taking the SAT multiple times advantages or disadvantages certain ethnicities, the variable of ethnicity will be analyzed.

Table 3.3 reports the percentage of North Carolina State University fall semester first-time-in-college students in each ethnicity from 2003-2008. It is evident that the majority of admitted students who attend North Carolina State University are White. It is imperative that the variable ethnicity be analyzed in conjunction with SAT results in order to determine whether Score Choice could disadvantage certain ethnicities during the application process.

Table 3.3

Ethnic Percentages of North Carolina State University Fall Semester First-Time-In-College Bachelor's Degree Students

Ethnicity	2003	2004	2005	2006	2007	2008
Nonresident Alien	0.6	0.6	0.5	0.7	0.6	1.1
Race and Ethnicity Unknown	.	1.1	2.3	3.2	2.8	2.8
Hispanics of any race	2.3	2.4	2.2	2.4	2.7	2.4
American Indian, Alaska Native	0.7	0.9	0.8	0.5	0.6	0.7
Asian	4.2	4.3	4.6	5.1	5	5.3
Black or African American	9.9	10.7	9.5	8.9	9	9.5
Native Hawaiian, Pacific Islander	.	.	.	.	.	.
White	82.3	79.9	80.2	79.3	79.4	78.2
Two or more races	.	.	.	.	.	.

Note. From "University Planning and Analysis: First Time Freshman Profile," 2011, http://www2.acs.ncsu.edu/UPA/admissions/freshman_profile.htm

The SAT questionnaire divided ethnicity into nine categories: American Indian/Alaskan Native, Asian/Asian American/Pacific Islander, Black/African American, Mexican/Mexican American, Puerto Rican, Other Hispanic/Latino/Latin American, White, Other, and Unreported. For the purpose of this study, eight dummy codes will be created with White as the referent group. The "Other" category is necessary as many students do not identify themselves as being in one of these ethnic groups because many come from mixed ethnic backgrounds. Due to this fact, the "Other" category is necessary to capture these groups. The "Unreported" category exists since it is not a requirement that someone reports their ethnicity when taking the SAT. Furthermore, admissions offices cannot require this information. For this reason, it is necessary to include this "Unreported" category.

Gender. North Carolina State University differs from its UNC System counterparts in that there are more males enrolled at North Carolina State University than females. Table 3.4 indicates that all other UNC System universities have more undergraduate and graduate females enrolled than males. The only exception to this is the UNC School of the Arts, which has more undergraduate males than females; however, this is not a university, so it is not an accurate comparison. There are a variety of reasons that NCSU is the only university in the system that admits and enrolls more males. The main reason is the nature of the degree programs offered at the university. North Carolina State University is known for its College of Agriculture and Life Sciences and its College of Engineering, which historically have been male dominated fields (Foster, 2003; Jagacinski, 1983).

Table 3.4

Fall 2008 UNC System Enrollment by Gender

	Undergraduate			Graduate/First Professional			All Students		
	Female	Male	Total	Female	Male	Total	Female	Male	Total
Appalachian State Univ.	7,423	7,138	14,561	1,404	645	2,049	8,827	7,783	16,610
East Carolina Univ.	12,246	8,728	20,974	4,636	2,067	6,703	16,882	10,795	27,677
Elizabeth City State Univ.	1,866	1,155	3,021	70	13	83	1,936	1,168	3,104
Fayetteville State Univ.	3,817	1,785	5,602	446	169	615	4,263	1,954	6,217
NC A&T State Univ.	4,635	4,194	8,829	973	586	1,559	5,608	4,780	10,388
North Carolina Central Univ.	3,945	2,033	5,978	1,439	618	2,057	5,384	2,651	8,035
North Carolina State Univ.	10,840	13,901	24,741	3,688	4,443	8,131	14,528	18,344	32,872
UNC Asheville	2,081	1,508	3,589	25	15	40	2,106	1,523	3,629
UNC - Chapel Hill	10,474	7,421	17,895	6,268	4,404	10,672	16,742	11,825	28,567
UNC Charlotte	9,496	8,833	18,329	3,084	1,887	4,971	12,580	10,720	23,300
UNC Greensboro	10,963	5,231	16,194	2,652	1,130	3,782	13,615	6,361	19,976
UNC Pembroke	3,484	2,094	5,578	513	212	725	3,997	2,306	6,303

Table 3.4 continued

UNC Wilmington	6,631	4,680	11,311	855	477	1,332	7,486	5,157	12,643
UNC School of the Arts	312	453	765	63	51	114	375	504	879
Western Carolina Univ.	3,739	3,391	7,130	1,252	668	1,920	4,991	4,059	9,050
Winston-Salem State Univ.	4,206	1,769	5,975	369	98	467	4,575	1,867	6,442
UNC Total	96,158	74,314	170,472	27,737	17,483	45,220	123,895	91,797	215,692

Note. From "The University of North Carolina a multi-campus university: Facts and figures," 2009, http://www.northcarolina.edu/about/facts.htm

The question this study will determine is whether the relationship between the number of times students take the SAT and their overall score on the test related to their gender. It will establish if there is a differential benefit to males/females taking the SAT more/fewer times than their counterparts. This type of information is important in order to determine whether certain genders are being disadvantaged due to the number of times they take the SAT. In this study, gender will be divided into three categories: male, female, and unreported. Male will be the referent group and dummy codes will be created to analyze the other two categories: women and unknown.

Financial aid. Students' need for financial aid was assessed and collected on the SDQ by asking the following question: "Do you plan to apply for financial aid at any college?" The findings will indicate whether the relationship between SAT scores and time depends upon a student's need for financial aid. The implications from these findings could be great if inequity is determined.

The purpose of financial aid is to supplement the difference between the cost of a higher educational institution and the amount a family has the ability to contribute (College Board, 2009a). According to the College Board (2009a), there are three main categories of financial aid: grants and scholarships, loans, and work. Students are not required to repay any grants or scholarships that they receive, as these items tend to be awarded based on merit. Contrary, loans must be repaid. The third main category of financial aid, work, refers to "work studies", student who are employed on campus to generate money to help pay for tuition, books, and more.

Currently, low-income students are eligible to receive two fee waivers on the SAT (Bunin, 2009). Those eligible for consideration for fee waivers must be high school students who meet the financial eligibility guidelines, such as participating in the Federal Free and Reduced Lunch/National School Lunch Program (College Board, 2009c). If students meet these guidelines, they can have two SAT fees waived; however, if they feel the need to take the test three or more times, they will have to pay the $49 fee.

The variable of student need for financial aid will be divided into four categories: yes (indicating that they do intend to receive financial aid); no (indicating that they do not intend to receive financial aid); unknown (indicating that they do not know whether they will need to obtain financial aid); and unreported (indicating the student did not answer this question). "Yes" is the referent group and the "No," "Unknown," and "Unreported" categories are dummy variables.

High school location. According to Chenoweth (1998), students from suburban schools scored approximately 15 to 20 points above the average SAT score, "whereas urban and rural schools had average scores that were between 9 and 17 points below the average" (p. 24). Although these results were conclusive, they failed to explain whether students from urban schools took the SAT more or fewer times than those from suburban schools. With the implementation of Score Choice, these would be the only results a university would be privy to. Universities may not receive any information regarding the number of times a student took the SAT; nevertheless, if they know that there is a difference between students who live in urban areas and those who live in suburban areas, then the admissions counselors could make a more informed decision.

The high school location variable was self-reported and was divided into six categories on the SAT questionnaire: large city, medium-sized city, small city or town, suburban, rural, and unreported. Suburban is the referent group with large city, medium-sized city, small city or town, rural, and unreported categories as dummy variables.

Dependent Variables

The dependent variables measured in this study were the verbal and math SAT scores sent to North Carolina State University between 2003 and 2008 and the admissions status associated with the students who submitted their SAT scores. Only verbal and math SAT scores are included in this study, as the written portion of the SAT was only introduced in March 2005.

Data Collection

When students register for the SAT, they are given a questionnaire which asks approximately 43 questions regarding their high school courses, grades, participation in extracurricular activities, type of higher educational institution they are interested in, their degree goals, and more (College Board, 2005). Answers to these questions, paired with students' SAT verbal, math, and essay scores (reported on a scale from 200 to 800), are sent to the universities indicated by the student. Admissions officers are sent electronic copies of all the results, from the SAT and the questionnaire. These files were provided by the NCSU Undergraduate Admissions Office for analysis in this study.

Data Analysis

Regression analysis has been the chosen method for most previous research on the SAT. However, the goal of this study is not to predict SAT scores, but rather to identify

whether a difference is evident in SAT score based upon the number of times a student takes the test. Due to this fact, multilevel linear modeling will be used to answer the research questions one, two, three, and four, and regression will be used to answer the fifth research question.

Prior to the full data analysis, basic statistics will be conducted to determine whether multicollinearity exists. Multicollinearity indicates that independent variables are highly correlated, thus measuring the same thing. As it is not necessary to include multiple variables measuring the same thing, the appropriate variable will be removed, if necessary.

Over the past decade, there has been an increase in the use of multilevel modeling (MLM) techniques, often referred to as hierarchical linear modeling (HLM), in the social science sector (Everson & Millsap, 2004; Schreiber & Griffin, 2004). This increase is in part due to the fact that data in the social sciences are often hierarchical. Much of the data are nested; educational examples include: students nested within schools, classrooms nested within schools, and longitudinal data such as repeated measures taken from students over time (Schreiber & Griffin, 2004). The last example is the method that was used in this study. As the data were not balanced, the number of people who took the SAT once did not match that of the people who took it twice, three times, or more; therefore, MLM was the appropriate method as it permits the analysis of repeated SAT scores across time even though the data are not balanced. Repeated measures analysis of variance on the other hand requires balanced datasets and therefore would be an inappropriate choice for the study.

Another advantage to using MLM is that it analyzes complete data, not just complete cases. Consequently, if a student has taken the SAT four times and he/she fails to self-report

ethnicity on the questionnaire, the other data can still be used when the equation does not include ethnicity. The first step in conducting MLM is to perform a preliminary analysis to guarantee there is adequate variability at Level 1 and Level 2 to merit continuing with analyses (Raudenbush & Bryk, 2002). This preliminary analysis is referred to as the fully unconditional or null model. Results from this analysis will indicate how much of the variability in SAT scores is due to time, which is considered 'within person' variability and how much of the variability is due to the students' own characteristics, which is considered the "between person" variability. The fully unconditional model is:

Level 1: SAT Score$_{it}$ = β_{0it} + r_{it}

Level 2: β_{0i} = γ_{00} + u_{0i}

SAT Score$_{it}$ is the SAT score for person $_i$ at time $_t$

β_{0it} is the average score on the SAT for person $_i$ at time $_t$

r_{it} is the average fluctuation within person (level 1 error)

γ_{00} is the sample average SAT score

u_{0i} is the between person difference from the sample average (level 2 error)

Using SAS computer software, the data was analyzed using the unconditional model to determine whether there was enough variability at each level. Once determined that there was an appropriate amount of variability, a one-way ANCOVA with random effects was conducted to answer the first and second research questions: (1) On average, is there change in student SAT scores associated with repeated testing? (2) Is there a point at which scores decrease with each additional attempt of the SAT? (This will be tested by looking at the quadratic effect). The equation is as follows:

Level 1: $SAT\ Scores_{it} = \beta_{0it} + \beta_{1it}(Time) + \beta_{2it}(Time^2) + r_{it}$

Level 2: $\beta_{0i} = \gamma_{00} + u_{0i}$

$\beta_{1i} = \gamma_{10} + u_{1i}$

$\beta_{2i} = \gamma_{20} + u_{2i}$

In Level 1, β_{0it}, the intercept, is defined as the expected SAT score on the SAT for $person_i$ at the first time the test was taken (time = 0). β_1 is the expected linear change in SAT score associated with taking the SAT additional times. β_2 is the expected quadratic change in SAT score associated with taking the SAT additional times. r_{it} is the average fluctuation within person, or Level 1 error. This is basically the amount that person's SAT scores fluctuate over time. In order to test for quadratic effect, it is necessary to include linear effect in the model. The individual intercepts, β_{0i}, and slopes, β_1 and β_2, become the outcome variables for the Level 2 equations. The average SAT score for the sample when time is equal to zero (first time taking the SAT) is represented by γ_{00}, the average linear change over time for the sample is represented by γ_{10} and the average quadratic change over time for the sample is represented by γ_{20}. The extent to which people vary from the average SAT score of the sample is represented by u_{0i}.

An intercepts and slopes as outcomes equation was conducted to answer the third and fourth research questions: (3) Does the relationship between repeated testing and SAT scores depend on ethnicity, gender, financial aid, or location of a student's high school? and (4) Does the point at which students' SAT scores decrease with each additional attempt of the SAT depend on ethnicity, gender, financial aid, or location of a student's high school? (In

this equation, quadratic time will be used to find the point at which scores decrease with each additional attempt of the SAT).

Level 1: Math/Verbal SAT Scores$_{it}$ = β_{0it} + β_{1it} (Time) + β_{2it} (Time^2) + r_{it}

Level 2: β_{0i} = γ_{00} + γ_{01}(Ethnicity American Indian/Alaskan Native) + γ_{02}(Ethnicity Asian/Asian American/Pacific Islander) + γ_{03}(Ethnicity Black/African American) + γ_{04}(Ethnicity Mexican/Mexican American) + γ_{05}(Ethnicity Puerto Rican) + γ_{06}(Ethnicity Other Hispanic/Latino/Latin American) + γ_{07}(Ethnicity Other) + γ_{08}(Ethnicity Unreported) + γ_{09}(Gender Female) + γ_{010}(Gender Unreported) + γ_{011}(Financial Aid No) + γ_{012}(Financial Aid Unknown) + γ_{013}(Financial Aid Unreported) + γ_{014}(High School in small city/town) + γ_{015}(High School in medium-sized city) + γ_{016}(High School in large city) + γ_{017}(High School in rural area) + γ_{018}(High School unreported) + u_{0i}

β_{1i} = γ_{10} + γ_{11}(Ethnicity American Indian/Alaskan Native) + γ_{12}(Ethnicity Asian/Asian American/Pacific Islander) + γ_{13}(Ethnicity Black/African American) + γ_{14}(Ethnicity Mexican/Mexican American) + γ_{15}(Ethnicity Puerto Rican) + γ_{16}(Ethnicity Other Hispanic/Latino/Latin American) + γ_{17}(Ethnicity Other) + γ_{18}(Ethnicity Unreported) + γ_{19}(Gender Female) + γ_{110}(Gender Unreported) + γ_{111}(Financial Aid No) + γ_{112}(Financial Aid Unknown) + γ_{113}(Financial Aid Unreported) + γ_{114}(High School in small city/town) + γ_{115}(High School in medium-sized city) + γ_{116}(High School in

large city) + γ_{117}(High School in rural area) + γ_{118}(High School unreported) + u_{1i}

$\beta_{2i} = \gamma_{20} + \gamma_{21}$(Ethnicity American Indian/Alaskan Native) + γ_{22}(Ethnicity Asian/Asian American/Pacific Islander) + γ_{23}(Ethnicity Black/African American) + γ_{24}(Ethnicity Mexican/Mexican American) + γ_{25}(Ethnicity Puerto Rican) + γ_{26}(Ethnicity Other Hispanic/Latino/Latin American) + γ_{27}(Ethnicity Other) + γ_{28}(Ethnicity Unreported) + γ_{29}(Gender Female) + γ_{210}(Gender Unreported) + γ_{211}(Financial Aid No) + γ_{212}(Financial Aid Unknown) + γ_{213}(Financial Aid Unreported) + γ_{214}(High School in small city/town) + γ_{215}(High School in medium-sized city) + γ_{216}(High School in large city) + γ_{217}(High School in rural area) + γ_{218}(High School unreported) + u_{2i}

If the previous analysis indicates that the Level 2 predictor reduces the residual variance of β_{1i} or β_{2i} to near zero, then the slope should be constrained and further analysis will be necessary. This determination was based on Singer's (1998) criteria for choosing the appropriate model, by comparing the difference in the Chi squared distributions.

Regression was used to answer the fifth research question: Is a student's SAT score that would be reported through Score Choice related to ethnicity, gender, financial aid, or location of a student's high school? The equation used will be:

SAT Scores = a + b_1(Ethnicity American Indian/Alaskan Native) + b_2(Ethnicity Asian/Asian American/Pacific Islander) + b_3(Ethnicity Black/African American) + b_4(Ethnicity Mexican/Mexican American) + b_5(Ethnicity Puerto

Rican) + b_6(Ethnicity Other Hispanic/Latino/Latin American) + b_7(Ethnicity Other) + b_8(Ethnicity Unreported) + b_9(Gender Female) + b_{10}(Gender Unreported) + b_{11}(Financial Aid No) + b_{12}(Financial Aid Unknown) + b_{13}(Financial Aid Unreported) + b_{14}(High School in small city/town) + b_{15}(High School in medium-sized city) + b_{16}(High School in large city) + b_{17}(High School in rural area) + b_{18}(High School unreported)

In this equation the *B* coefficients, also known as the regression coefficients, represent the contribution each independent variable makes to the prediction of the dependent variable (StatSoft Inc., 2008).

Summary

Chapter Three has summarized the methodological process used to analyze the data for this study. Multilevel modeling and regression analysis were used to evaluate the variables studied. Chapter 4 will present the findings from conducting these statistical methods. Chapter 5 will discuss the findings in relation to the higher educational literature already available regarding SAT scores.

Chapter 4: Results

Introduction

The overall purpose of this study was to determine whether there was a significant relationship between repeated test taking on the SAT and ethnicity, gender, need for financial aid, or high school location for North Carolina State University applicants, and to investigate the implications the new SAT reporting option, Score Choice, could have on future applicants scores. To answer these questions, data from students who applied to North Carolina State University between 2003 and 2008 were analyzed using multilevel modeling and regression analysis. This chapter will provide the results from this data analysis, beginning with overall descriptive statistics followed by the findings from the multilevel modeling and regression analysis. The results will focus interpretations as they relate to the participants, which is mathematically equivalent to a variable-centered interpretation.

Descriptive Statistics

Table 4.1 displays the descriptive statistics associated with the independent variables used in this study. The results show that 65% of the applicants in this study were White. This percentage is slightly lower than the North Carolina State University ethnic percentage of White students reported in Table 3.3. This discrepancy is due to the fact that Table 3.3 reported admitted freshman, whereas this study looked at all applicants over a six year period, thus encompassing a significant number of applicants who were not admitted.

The majority of applicants reported that they would receive financial aid; only 7.42% said they would not receive any aid (merit or need-based). There were slightly more male

applicants than females, which is consistent with Table 3.4 (the UNC System Enrollment by Gender). Most applicants attended either a suburban or rural high school, with slightly more attending suburban. These results were not surprising due to the vast geographical range of North Carolina.

Table 4.1

Descriptive Statistics of Independent Variables

Independent Variable	Frequency	Percentage
Ethnicity		
White	98,932	65.13
American Indian/Alaskan Native	1,322	0.87
Asian/Asian American/Pacific Islander	7,283	4.79
Black/African American	25,359	16.69
Mexican/Mexican American	1,178	0.78
Puerto Rican	915	0.6
Other Hispanic/Latino/Latin American	2,780	1.83
Other	3,550	2.34
Unreported	10,582	6.97
Receive Financial Aid		
No	11,265	7.42
Yes	93,815	61.76
Unknown	31,928	21.02
Unreported	14,893	9.8
Gender		
Male	81,191	53.45
Female	70,435	46.37
Unreported	275	0.18
High School Location		
Small city or town	27,844	18.33
Medium-sized city	22,257	14.65
Large city	17,253	11.36
Suburban	31,668	20.85

Table 4.1 Continued

Rural	31,470	20.72
Unreported	21,409	14.09

As noted in Table 4.2, a total of 151,901 subjects were used in this study. There were 299,376 individual SAT scores reported from this group of 151, 901 applicants, with almost 53% of them taking the SAT once and less than 1% taking it six times. Six was the highest number of times any individual in this study took the SAT and one was the minimum.

Table 4.2

Number of Times North Carolina State University 2003-2008 Applicants took the SAT

Test Number	Frequency	Percentage
1	79,824	52.55
2	48,408	31.87
3	18,612	12.25
4	3,964	2.61
5	860	0.57
6	233	0.15

Table 4.3 displays the number of times North Carolina State University 2003-2008 applicants took the SAT in relation to the independent variables. It is evident that students who stated they would receive financial aid tend to take the SAT more times than the other categories of financial aid need. The number of times males and females take the SAT are quite evenly distributed; likewise, so are the values for high school location.

Table 4.3

Number of Times North Carolina State University 2003-2008 Applicants took the SAT by Independent Variables

	Test Number					
Independent Variable	1	2	3	4	5	6
Ethnicity						
White	48,541	33,732	13,268	2,669	583	139
American Indian/Alaskan Native	819	343	127	26	3	4
Asian/Asian American/Pacific Islander	2,931	2,459	1,357	406	94	36
Black/African American	17,363	6,051	1,537	326	65	17
Mexican/Mexican American	795	286	78	14	5	0
Puerto Rican	547	267	90	10	1	0
Other Hispanic/Latino/Latin American	1,579	896	257	41	4	3
Other	1,988	1,082	378	82	16	4
Unreported	5,261	3,292	1,520	390	89	30
Gender						
Male	43,383	25,753	9,470	2,029	438	118
Female	36,220	22,617	9,132	1,932	420	114
Unreported	221	38	10	3	2	1
Receive Financial Aid						
Yes	52,502	28,904	9,919	1,959	413	118
No	4,804	3,961	1,938	436	101	25
Unknown	15,346	10,781	4,520	1,020	206	55
Unreported	7,172	4,762	2,235	549	140	35
High School Location						
Suburban	15,734	11,092	4,009	668	138	27
Small city or town	14,449	8,821	3,524	808	182	60
Medium-sized city	10,355	7,504	3,331	837	184	46
Large city	8,704	5,659	2,311	480	76	23
Rural	17,454	9,421	3,509	827	212	47
Unreported	13,128	5,911	1,928	344	68	30

Correlation

Because the independent variables are nominal, a phi (or Cramer's *V*) correlation was conducted. According to Morgan (2004), "phi/Cramer's *V* are good choices for statistics when analyzing two nominal variables" (p.99). Cramer's *V* is the appropriate correlation statistic to view on these interactions as all the variables have more than two levels. "If one has a 2x2 cross tabulation, phi is the appropriate statistic. For larger crosstabs, Cramer's *V* is used" (Morgan, 2004, p.99).

Table 4.4

Cramer's V Correlation of Independent Variables

	Gender	Ethnicity	High School Location	Need for Financial Aid
Gender	-	0.04*	0.03*	0.03*
Ethnicity	0.04*	-	0.09*	0.30*
High School Location	0.03*	0.09*	-	0.04*
Need for Financial Aid	0.03*	0.30*	0.04*	-

Note: *p<.001

All the independent variables used in this study were significantly correlated. However, the magnitude of the vast majority of correlations was low. Due to the large number of subjects used in this study (151,901), the fact that the correlations were significant was not an issue. None of the variables overlapped one another more than 4% with the exception of need for financial aid and ethnicity, which had approximately a 9% overlap. In order to determine whether this would be an issue of multi-collinearity during the analysis, the first model was run using both predictors (need for financial aid and ethnicity); then

another analysis was run where only ethnicity was included. From these results, the two models were compared to see if the pattern of effects were any different. Due to the fact that there was not a difference, both variables were used.

Fully Unconditional Model

Level 1: $SAT\ Score_{it} = \beta_{0it} + r_{it}$

Level 2: $\beta_{0i} = \gamma_{00} + u_{0i}$

β_{0it} = Intercept; SAT Score for person i at time t
r_{it} = Within-person variability; level 1 variance
γ_{00} = Grand mean; Sample average SAT score
u_{0i} = Between-person variability; level 2 variance

A preliminary analysis, more specifically a fully unconditional model, allowed for the concurrent analysis of the time-varying (level 1) and time-invariant (level 2) variables. This first step in multilevel modeling (MLM) was used to determine whether there was sufficient within-person (level 1) and between-person (level 2) variability to justify continuing analysis (Raudenbush et al., 2002). Results from this analysis indicated that 86% of the variability in math SAT scores was between people ($\tau_{00} = 8459.18$, $z = 247.88$, $p < .001$) and 85% of the variability in verbal SAT scores was between people ($\tau_{00} = 7978.19$, $z = 246.12$, $p < .001$). Significant within person variability was also found in math SAT ($\sigma^2 = 1333.68$, $z = 269.55$, $p < .001$) and verbal SAT scores ($\sigma^2 = 1353.31$, $z = 269.54$, $p < .001$), with 14% of variability in math and 15% of variability in verbal SAT scores associated with individual differences. Basically these findings indicate that there was a greater difference in scores between individuals in this sample than within those individuals over time.

Research Questions 1 & 2

Question 1: On average, is there change in SAT associated with repeated testing?

Question 2: Is there a point at which scores decrease with each additional attempt of the SAT?

$$\text{Level 1 Scores}_{it} = \beta_{0it} + \beta_{1it}(\text{Time}) + \beta_{2it}(\text{Time}^2) + r_{it}$$

$$\text{Level 2: } \beta_{0i} = \gamma_{00} + u_{0i}$$
$$\beta_{1i} = \gamma_{10} + u_{1i}$$
$$\beta_{2i} = \gamma_{20} + u_{2i}$$

β_{0it} = Intercept; SAT Score for person i at time t
r_{it} = Within-person variability; level 1 variance
γ_{00} = Grand mean; Sample average SAT score
u_{0i} = Between-person variability; level 2 variance
β_{1i} = the expected linear change in SAT score associated with taking the SAT additional times
γ_{10} = the average linear change over time for the sample
u_{1i} = variance from the linear slope
β_{2i} = the expected quadratic change in SAT score associated with taking the SAT additional times
γ_{20} = the average quadratic change over time for the sample
u_{2i} = variance from the quadratic slope

A one-way ANCOVA with random effects was conducted to answer the first and second research questions. The only difference between these two research questions was that the first research question specifically examined linear time, while the second research question examined quadratic time. Math SAT scores and verbal SAT scores were examined separately. Slopes were allowed to vary in these equations; when the slopes were constrained, less variability was explained.

Linear time. The findings, shown in Table 4.5, indicate that there was a significant increase in math SAT scores over time (γ_{10}= 15.70, t = 174.90, p < .001) and 20% of the

within person variability in math SAT scores was accounted for by linear time. Likewise, there was a significant increase in verbal SAT scores over time (γ_{10}= 15.19, t = 164.18, p < .001) and 21% of the within person variability in verbal SAT scores was accounted for by linear time. Therefore, according to these findings it is evident that scores increased with subsequent attempts.

Table 4.5

One-way ANCOVA with Random Effect Linear Time Results

Fixed Effects			
Math SAT Scores, β_0		Verbal SAT Scores, β_0	
Intercept, γ_{00}	536.38***	Intercept, γ_{00}	509.64***
Linear Time, β_1		Linear Time, β_1	
Intercept, γ_{10}	15.70***	Intercept, γ_{10}	15.19***
Random Effects			
Math SAT Scores, τ_{00}	8414.93***	Verbal SAT Scores, τ_{00}	8030.05***
Within-person fluctuation, σ^2	1060.39***	Within-person fluctuation, σ^2	1064.18***

Note: *p<.05, **p<.01, ***p<.001

Quadratic time. The second research question examined quadratic time. The findings indicated there was still a significant increase in math SAT scores over and above the effects of linear time (γ_{10}= 19.95, t = 111.03, p < .001) and above the effects of linear verbal SAT scores over time (γ_{10}= 18.30, t = 100.17, p < .001) when the quadratic effective was added to the model. Furthermore, the quadratic effect was also found significant for both math SAT scores (γ_{20}= -1.86, t = -27.33, p < .001) and verbal SAT scores (γ_{20}= -1.38, t = -19.72, p < .001). Table 4.6 shows the results from the one-way ANCOVA with random effects for quadratic time. 20.78% of the within person variability in math SAT scores was

accounted for by linear and quadratic time and 21.40% of the within person variability in verbal SAT scores.

Table 4.6

One-way ANCOVA with Random Effect Quadratic Time Results

Fixed Effects			
Math SAT Scores, β_0		Verbal SAT Scores, β_0	
Intercept, γ_{00}	535.70***	Intercept, γ_{00}	509.16***
Linear Time, β_1		Linear Time, β_1	
Intercept, γ_{10}	19.96***	Intercept, γ_{10}	18.30***
Quadratic Time, β_2		Quadratic Time, β_2	
Intercept, γ_{20}	-1.86***	Intercept, γ_{20}	-1.38***
Random Effects			
Math SAT Scores, τ_{00}	8400.07***	Verbal SAT Scores, τ_{00}	8020.52***
Within-person fluctuation, σ^2	1056.49***	Within-person fluctuation, σ^2	1063.74***

Note: *p<.05, **p<.01, ***p<.001

It is difficult to understand the quadratic trend without a graphic representation of the results. Just like significant interactions are graphed for interpretation, so to are quadratic effects. Figure 4.1 and 4.2 are below.

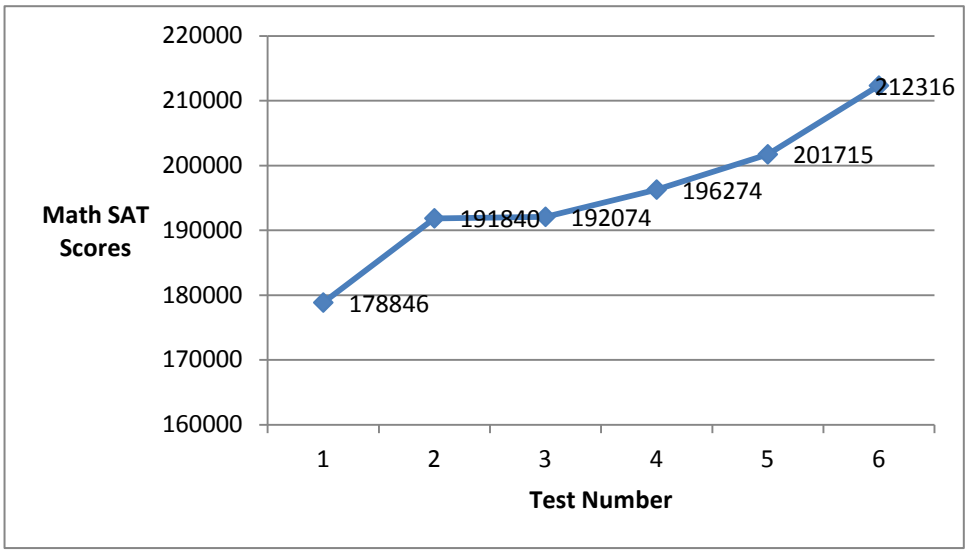

Figure 4.1 *Math SAT Scores*

Figure 4.1 displays the results regarding math SAT scores over time. It is evident that there was a significant increase in math SAT scores over time. The difference in math SAT scores diminishes the more times a student took the exam, with the biggest jump in scores taking place between the first time a student took the test and the second. There is clearly an improvement the more times a student takes the test, but the improvement is diminishing.

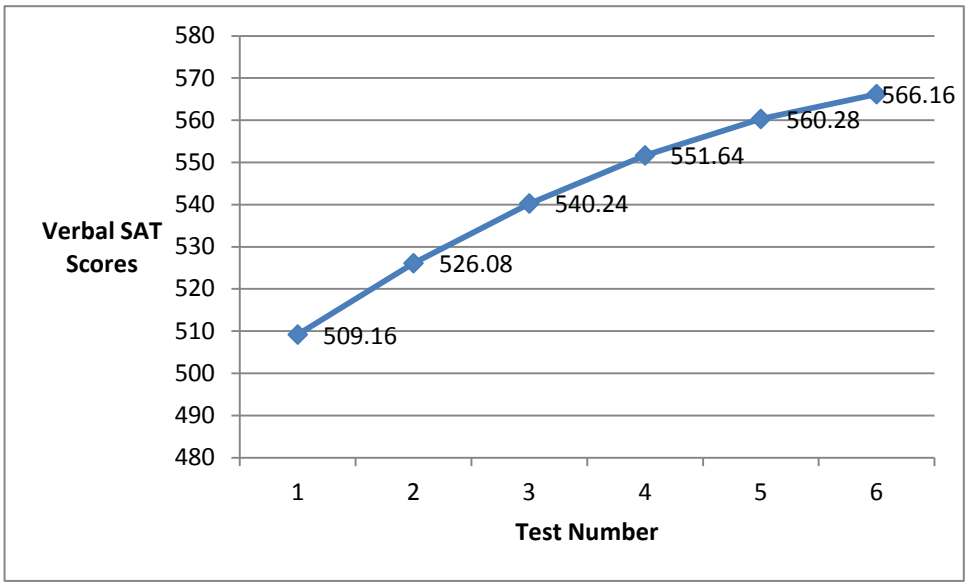

Figure 4.2 *Verbal SAT Scores*

Figure 4.2 illustrates that there was a significant increase in verbal SAT scores over time. It is clear that the difference in verbal SAT scores slightly lessened the more times a student took the exam. However, it never completely leveled off or decreased. Therefore, one could speculate that the more times a student takes the verbal SAT, the better he/she will score.

Research Questions 3 & 4

Question 3: Does the relationship between repeated testing and SAT scores depend on ethnicity, gender, financial aid, or location of a student's high school?

Question 4: Does the point at which students SAT scores decrease with each additional attempt of the SAT depend on ethnicity, gender, financial aid, or location of a student's high school?

Level 1: Math/Verbal SAT Scores$_{it}$ = β_{0it} + β_{1it} (Time) + β_{2it} (Time^2) + r_{it}

Level 2: $\beta_{0i} = \gamma_{00} + \gamma_{01}$(Ethnicity American Indian/Alaskan Native) + γ_{02}(Ethnicity Asian/Asian American/Pacific Islander) + γ_{03}(Ethnicity Black/African American) + γ_{04}(Ethnicity Mexican/Mexican American) + γ_{05}(Ethnicity Puerto Rican) + γ_{06}(Ethnicity Other Hispanic/Latino/Latin American) + γ_{07}(Ethnicity Other) + γ_{08}(Ethnicity Unreported) + γ_{09}(Gender Female) + γ_{010}(Gender Unreported) + γ_{011}(Financial Aid No) + γ_{012}(Financial Aid Unknown) + γ_{013}(Financial Aid Unreported) + γ_{014}(High School in small city/town) + γ_{015}(High School in medium-sized city) + γ_{016}(High School in large city) + γ_{017}(High School in rural area) + γ_{018}(High School unreported) + u_{0i}

$\beta_{1i} = \gamma_{10} + \gamma_{11}$(Ethnicity American Indian/Alaskan Native) + γ_{12}(Ethnicity Asian/Asian American/Pacific Islander) + γ_{13}(Ethnicity Black/African American) + γ_{14}(Ethnicity Mexican/Mexican American) + γ_{15}(Ethnicity Puerto Rican) + γ_{16}(Ethnicity Other Hispanic/Latino/Latin American) + γ_{17}(Ethnicity Other) + γ_{18}(Ethnicity Unreported) + γ_{19}(Gender Female) + γ_{110}(Gender Unreported) + γ_{111}(Financial Aid No) + γ_{112}(Financial Aid Unknown) + γ_{113}(Financial Aid Unreported) + γ_{114}(High School in small city/town) + γ_{115}(High School in medium-sized city) + γ_{116}(High School in large city) + γ_{117}(High School in rural area) + γ_{118}(High School unreported) + u_{1i}

$$\beta_{2i} = \gamma_{20} + \gamma_{21}(\text{Ethnicity American Indian/Alaskan Native}) + \gamma_{22}(\text{Ethnicity Asian/Asian American/Pacific Islander}) + \gamma_{23}(\text{Ethnicity Black/African American}) + \gamma_{24}(\text{Ethnicity Mexican/Mexican American}) + \gamma_{25}(\text{Ethnicity Puerto Rican}) + \gamma_{26}(\text{Ethnicity Other Hispanic/Latino/Latin American}) + \gamma_{27}(\text{Ethnicity Other}) + \gamma_{28}(\text{Ethnicity Unreported}) + \gamma_{29}(\text{Gender Female}) + \gamma_{210}(\text{Gender Unreported}) + \gamma_{211}(\text{Financial Aid No}) + \gamma_{212}(\text{Financial Aid Unknown}) + \gamma_{213}(\text{Financial Aid Unreported}) + \gamma_{214}(\text{High School in small city/town}) + \gamma_{215}(\text{High School in medium-sized city}) + \gamma_{216}(\text{High School in large city}) + \gamma_{217}(\text{High School in rural area}) + \gamma_{218}(\text{High School unreported}) + u_{2i}$$

An intercepts and slopes as outcomes equation was conducted to answer the third and fourth research questions. The third research question specifically analyzed linear time, whereas the fourth research question looked at the addition of quadratic time to the model. Both models were run with the financial aid variable included and also without the financial aid variable in order to determine whether the inclusion of the variable would alter the results due to multicollinearity. Additionally, the financial aid variable was listed as a limitation in this study because the variable did not indicate whether the financial aid received was need-based or merit-based. However, when comparing the pattern of effects there was no difference between the models including financial aid and those that did not contain this variable; therefore, both financial aid and ethnicity were used in the models, given the relatively low association between them.

Linear time. The findings for the third research question are displayed in Table 4.7. They indicate that math SAT scores (γ_{10}) and verbal SAT scores (γ_{10}) tended to increase with subsequent attempts. There were significant ethnic ($\gamma_{01}, \gamma_{02}, \gamma_{03}, \gamma_{04}, \gamma_{05}, \gamma_{06}, \gamma_{07}, \gamma_{08}$), gender ($\gamma_{09}, \gamma_{110}$), need for financial aid ($\gamma_{011}, \gamma_{012}, \gamma_{013}$), and high school location ($\gamma_{014}, \gamma_{015}, \gamma_{016}, \gamma_{017}, \gamma_{018}$) differences in the average math SAT scores between students. Likewise, there were significant ethnic ($\gamma_{01}, \gamma_{02}, \gamma_{03}, \gamma_{04}, \gamma_{05}, \gamma_{06}, \gamma_{07}, \gamma_{08}$), gender ($\gamma_{110}$), need for financial aid ($\gamma_{011}, \gamma_{012}, \gamma_{013}$), and high school location ($\gamma_{014}, \gamma_{015}, \gamma_{017}, \gamma_{018}$) differences in the average verbal SAT scores between students; with the exception of females (γ_{09}) and those who attended a high school located in a large city (γ_{016}) which were not significant.

Table 4.7

Intercepts and Slopes as Outcomes Linear Results

Fixed Effects			
Math SAT Scores, β_0		Verbal SAT Scores, β_0	
Intercept, γ_{00}	591.49***	Intercept, γ_{00}	554.63***
Ethnicity		Ethnicity	
American Indian/Alaskan Native, γ_{01}	-57.60***	American Indian/Alaskan Native, γ_{01}	-59.63***
Asian/Asian American/Pacific Islander, γ_{02}	20.80***	Asian/Asian American/Pacific Islander, γ_{02}	-33.38***
Black/African American, γ_{03}	-110.63***	Black/African American, γ_{03}	-103.98***
Mexican/Mexican American, γ_{04}	-55.29***	Mexican/Mexican American, γ_{04}	-55.00***
Puerto Rican, γ_{05}	-58.78***	Puerto Rican, γ_{05}	-50.88***
Other Hispanic/Latino/Latin American, γ_{06}	-48.17***	Other Hispanic/Latino/Latin American, γ_{06}	-45.12***
Other, γ_{07}	-37.57***	Other, γ_{07}	-34.40***
Unreported, γ_{08}	-21.73***	Unreported, γ_{08}	-13.02***
Gender		Gender	
Female, γ_{09}	-27.19***	Female, γ_{09}	0.02
Unreported, γ_{010}	-54.28***	Unreported, γ_{010}	-47.17***
Receive Financial Aid		Receive Financial Aid	
No, γ_{011}	10.27***	No, γ_{011}	2.29*
Unknown, γ_{012}	8.09***	Unknown, γ_{012}	3.61***
Unreported, γ_{013}	4.35***	Unreported, γ_{013}	-3.47**
High School Location		High School Location	
Small city or town, γ_{014}	-18.94***	Small city or town, γ_{014}	-16.84***
Medium-sized city, γ_{015}	-3.28***	Medium-sized city, γ_{015}	5.14***
Large city, γ_{016}	-7.96***	Large city, γ_{016}	0.42

Table 4.7 Continued

Rural, γ_{017}	-32.94***	Rural, γ_{017}	-33.30***
Unreported, γ_{018}	-23.68***	Unreported, γ_{018}	-18.86***
Linear Time, β_1		Linear Time, β_1	
Intercept, γ_{10}	13.55***	Intercept, γ_{10}	8.19***
Ethnicity		Ethnicity	
American Indian/Alaskan Native, γ_{11}	6.54**	American Indian/Alaskan Native, γ_{11}	5.75*
Asian/Asian American/Pacific Islander, γ_{12}	1.44	Asian/Asian American/Pacific Islander, γ_{12}	4.24***
Black/African American, γ_{13}	9.07***	Black/African American, γ_{13}	10.47***
Mexican/Mexican American, γ_{14}	10.87***	Mexican/Mexican American, γ_{14}	14.58***
Puerto Rican, γ_{15}	5.99*	Puerto Rican, γ_{15}	12.80***
Other Hispanic/Latino/Latin American, γ_{16}	10.64***	Other Hispanic/Latino/Latin American, γ_{16}	12.92***
Other, γ_{17}	8.53***	Other, γ_{17}	6.08***
Unreported, γ_{18}	6.26***	Unreported, γ_{18}	6.41***
Gender		Gender	
Female, γ_{19}	-1.51***	Female, γ_{19}	0.19
Unreported, γ_{110}	2.08	Unreported, γ_{110}	10.30
Receive Financial Aid		Receive Financial Aid	
No, γ_{111}	-6.03***	No, γ_{111}	-4.86***
Unknown, γ_{112}	-2.72***	Unknown, γ_{112}	-1.75**
Unreported, γ_{113}	-0.48	Unreported, γ_{113}	-0.17
High School Location		High School Location	
Small city or town, γ_{114}	4.03***	Small city or town, γ_{114}	3.82***
Medium-sized city, γ_{115}	4.09***	Medium-sized city, γ_{115}	3.09***
Large city, γ_{116}	2.68***	Large city, γ_{116}	2.81***

Table 4.7 Continued

Rural, γ_{117}	4.88***	Rural, γ_{117}	5.38***
Unreported, γ_{118}	4.29***	Unreported, γ_{118}	3.84***
Random Effects		Random Effects	
Math Level 2 Variability, τ_{00}	7931.60***	Verbal Level 2 Variability, τ_{00}	7998.93***
Covariance, τ_{01}	-848.34***	Covariance, τ_{01}	-674.91***
Math Linear Slope, τ_{11}	269.25***	Verbal Linear Slope, τ_{11}	238.72***
Within-person fluctuation, σ^2	525.01***	Within-person fluctuation, σ^2	447.87***

Ethnicity

The results show that in contrast to White students, the referent (comparison) ethnic group, all ethnicities ($\gamma_{01}, \gamma_{02}, \gamma_{04}, \gamma_{05}, \gamma_{06}, \gamma_{07}, \gamma_{08}$) on average scored significantly lower on the math portion of the SAT the first time they took it, with the exception of Asian/Asian American/Pacific Islanders (γ_{03}) who scored almost 21 points higher than Whites. However, Asian/Asian American/Pacific Islanders (γ_{13}) were the only students found to have a non-significant change in math SAT scores over time. There was a positive relationship between all other ethnicities ($\gamma_{11}, \gamma_{12}, \gamma_{14}, \gamma_{15}, \gamma_{16}, \gamma_{17}, \gamma_{18}$) and changes in math SAT scores over time, indicating that students' scores on the math SAT tended to increase each time they retook the test. Since all ethnicities, with the exception of Asian/Asian American/Pacific Islanders, had a similar trend (significant positive relationship) in regard to their change over time, the Mexican/Mexican American variable was used as a representative of the group and can be viewed in Figure 4.3. Black/African American (γ_{03}) students scored on average the lowest of all ethnicities on the math portion of the SAT the first time they took the test; scoring approximately 111 points lower than their White counterparts. However, their score increased with each sitting of the math SAT at a similar rate to other ethnicities.

The points on the figures in this section are predicted points. They are not based on cell means; therefore, graphs may contain predicted points for students who took the SAT six times even though no students in this sample took the SAT six times within that category.

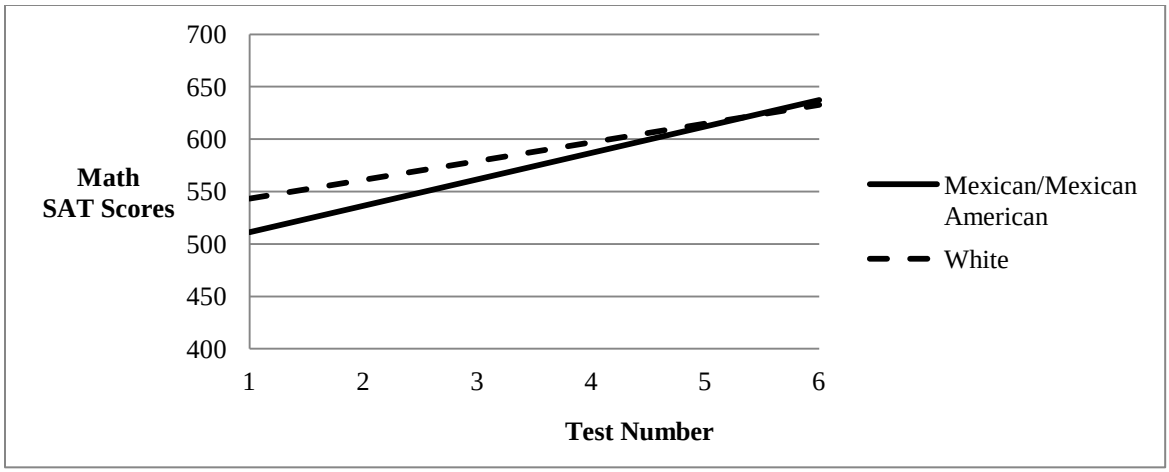

Figure 4.3 *Ethnicity and the Relationship between Linear Time and Math SAT scores*

All ethnicities ($\gamma_{01}, \gamma_{02}, \gamma_{03}, \gamma_{04}, \gamma_{05}, \gamma_{06}, \gamma_{07}, \gamma_{08}$) on average scored significantly lower on the verbal portion of the SAT the first time they took it than their White counterparts. Black/African American (γ_{03}) students scored on average the lowest of all ethnicities on the verbal portion of the SAT the first time they took the test, scoring approximately 104 points lower than White students. There was a significant positive relationship with all ethnic groups ($\gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{14}, \gamma_{15}, \gamma_{16}, \gamma_{17}, \gamma_{18}$) and changes in verbal SAT scores over time, indicating that scores on average increased the more times a student took the verbal SAT for all ethnic groups. That is, on average, scores increased over time for all ethnic groups. As before, the Mexican/Mexican American group was used as a representation of the trend for all the non-white ethnic groups (see Figure 4.4).

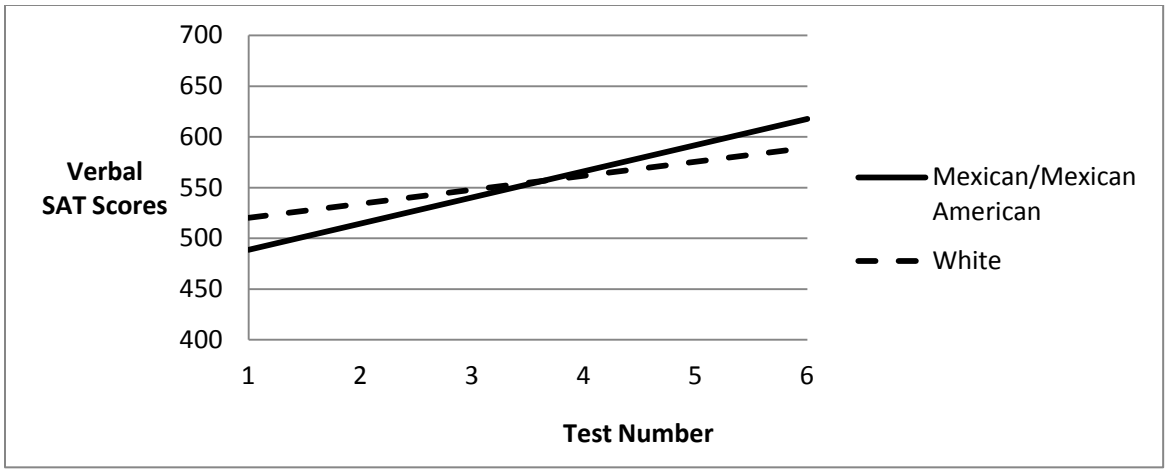

Figure 4.4 *Ethnicity and the Relationship between Linear Time and Verbal SAT scores*

Gender

In comparison to males (the referent value), females (γ_{09}) and those who did not report their gender (γ_{010}) on average scored significantly lower on the math SAT the first time they took the test. Females on average scored 27 points lower than males, while students with an unreported gender scored approximately 54 points lower. There was a significant negative relationship with females (γ_{19}) and changes in math SAT scores over time, which is displayed in Figure 4.5. It is evident that if a female and a male take the math SAT the same number of times, on average the male will score higher. The slopes of the lines displayed in Figure 4.5 appear to be parallel, showing that females and males tend to increase their scores at a similar rate each additional time they take the math SAT. Despite subsequent attempts on the SAT, on average, females will not catch-up to their male counterparts. There was a non-significant relationship in math SAT scores and time for those who did not report their gender (γ_{110}).

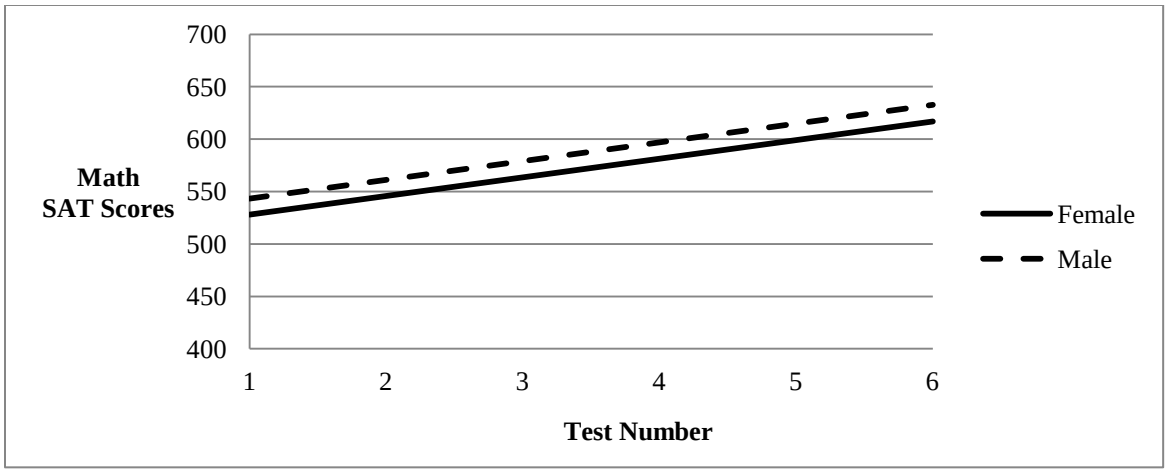

Figure 4.5 *Gender and the Relationship between Linear Time and Math SAT scores*

Females (γ_{09}), scored on average .02 points higher than males the first time they took the verbal portion of the SAT; however, this result was found to be non-significant. Students who did not report their gender (γ_{010}) on average scored significantly lower on the verbal SAT the first time they took the test than males, scoring 47 points lower. Neither females (γ_{19}) nor those who did not report their gender (γ_{110}) had a significant difference from men between math SAT scores and time.

Financial Aid

Students who reported they would not receive financial aid (γ_{011}), their need for financial aid was unknown (γ_{012}), or their need for financial aid was unreported (γ_{013}) on average all scored significantly higher on the math portion of the SAT the first time they took the test than the referent group (those who would receive financial aid). However, there was a significant negative relationship between students who reported they would not receive financial aid (γ_{111}) and those who reported their need for financial aid as unknown (γ_{112}) and

the math portion of the SAT over time. Since the trend of these two groups was similar, the results for students who reported they would not receive financial aid was graphed as a representative, see Figure 4.6. The graph shows that when compared to the referent value, those who reported they would not receive financial aid tended to score higher initially; however, over time their math SAT score did not increase at the same rate as students who reported they would receive financial aid. Although there was a negative relationship between those whose need for financial aid was unreported (γ_{113}) and their math SAT score over time, the results were non-significant.

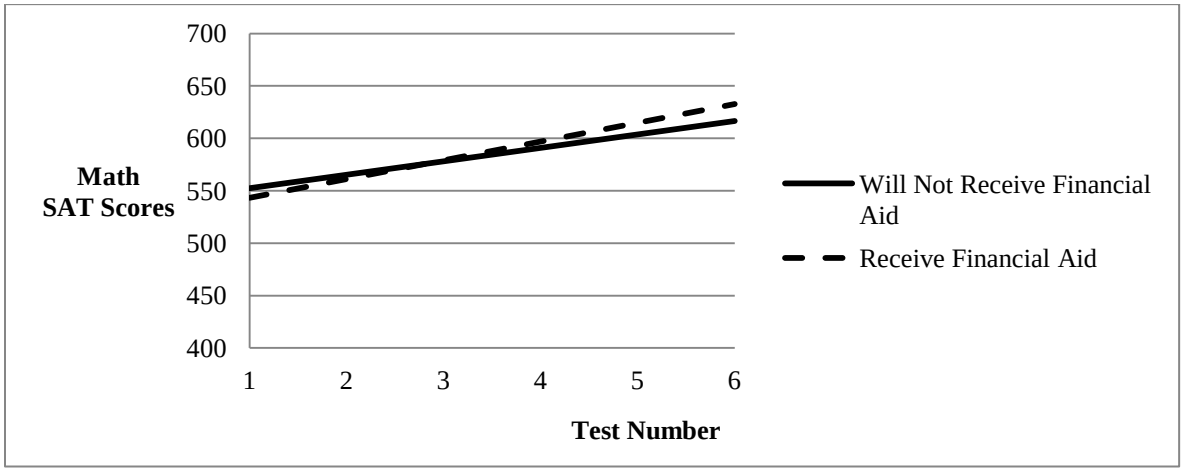

Figure 4.6 *Need for Financial Aid and the Relationship between Linear Time and Math SAT scores*

Students who reported they would not receive financial aid (γ_{011}), or their need for financial aid was unknown (γ_{012}) on average scored significantly higher on the verbal portion of the SAT the first time they took it than the referent group, students who would receive

financial aid. Students whose need for financial aid was unreported (γ_{013}) scored significantly lower than those who reported they would receive financial aid. However, there was a significant negative relationship between students who reported they would not receive financial aid (γ_{111}) and those who reported their need for financial aid as unknown (γ_{112}) and the verbal portion of the SAT over time. Since the trend of these two groups was similar, the results for students who reported they would not receive financial aid was graphed as a representative (see Figure 4.7). The graph shows that when compared to the referent value, those who reported they would not receive financial aid tended to score higher initially; however, over time their verbal SAT scores failed to increase at the same rate as students who reported they would receive financial aid. Although there was a negative relationship between those whose need for financial aid was unreported (γ_{113}) and their verbal SAT score over time, the results were non-significant.

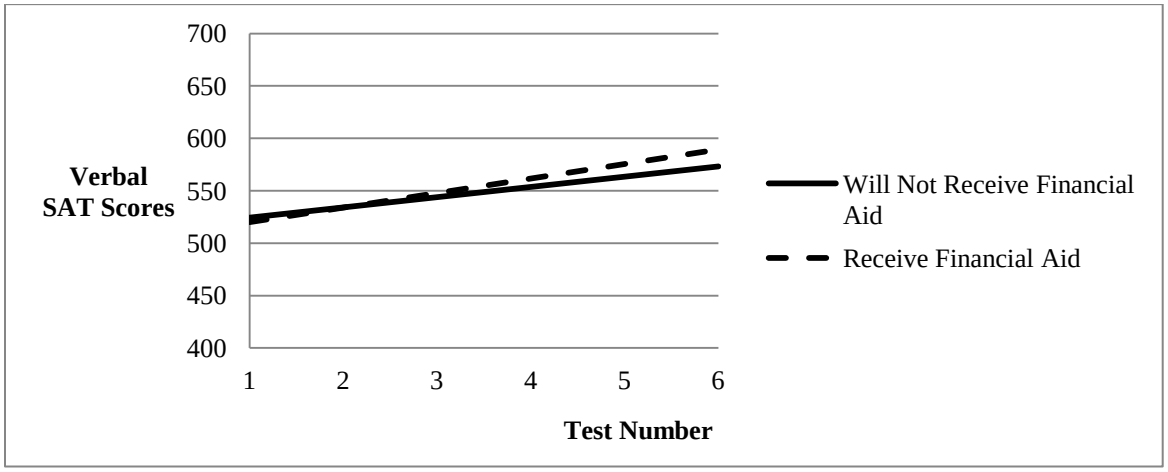

Figure 4.7 *Need for Financial Aid and the Relationship between Linear Time and Verbal SAT scores*

High School Location

In comparison to suburban, the referent value, students whose high school was located in a small city or town (γ_{014}), medium-sized city (γ_{015}), large city (γ_{016}), rural area (γ_{017}), or whose high school location was unreported (γ_{018}) all scored significantly lower on the math SAT the first time they took the test. Rural high school students scored the lowest, with their first test score on average being almost 33 points lower than those who attended a suburban high school. There was a significant positive relationship between high school location ($\gamma_{114,}$ $\gamma_{115,}$ $\gamma_{117,}$ γ_{118}) and changes in math SAT scores over time. Rural was chosen as a representative of the high school location variables since the trend for these variables was the same. Figure 4.8 shows that those who attend high school in a rural location increased their scores with each sitting of the test; however, even after taking the math SAT six times, they still would not achieve an equivalent score to someone who attended high school in a suburban setting and took the math SAT the same number of times. However, they do increase their score over time at a higher rate to their suburban counterpart.

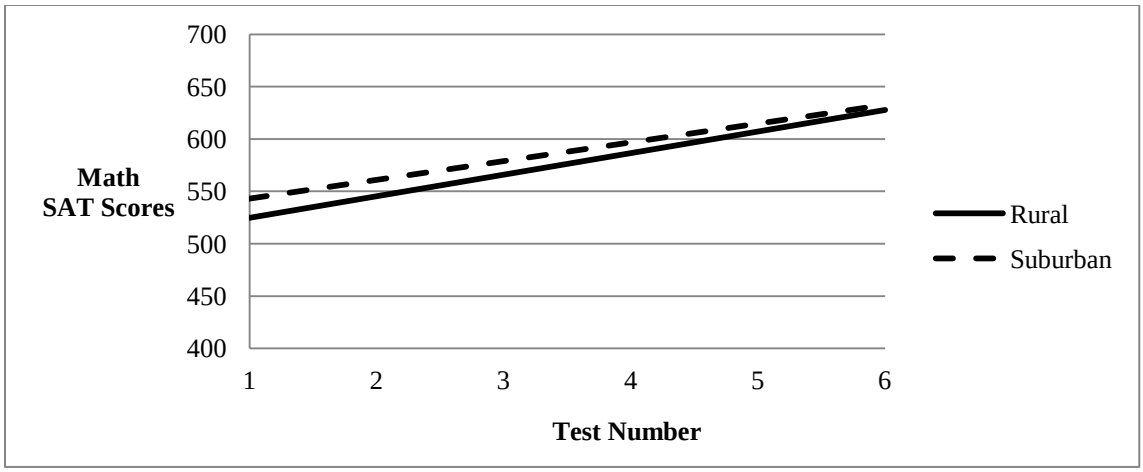

Figure 4.8 *High School Location and the Relationship between Linear Time and Math SAT scores*

Students whose high school was located in a small city or town (γ_{014}), rural area (γ_{017}), or whose high school location was unreported (γ_{018}) tended to score significantly lower on the verbal SAT the first time they took the test than those whose high school was in a suburban location. On average, rural high school students scored the lowest, with the average of their first test score being approximately 33 points lower than those who attended a suburban high school. However, in comparison to the referent group (suburban high school location), students who attended high school in a medium-sized city (γ_{015}) scored significantly higher on the verbal portion of the SAT the first time they took the test. Those whose high school was located in a large city (γ_{016}) also scored higher on the verbal SAT the first time they took it than suburban high school students; however, the results were found to be non-significant. There was, however, a significant positive relationship between all high school locations (γ_{114}, γ_{115}, γ_{117}, γ_{118}) and changes in verbal SAT scores over time. The rural

group was used as a representative of the trend for the non- suburban high school groups (see Figure 4.9).

Figure 4.9 shows that students who attended high school in a suburban setting not only scored higher on the verbal SAT the first time they took the test in comparison to those who attended high school in other locations, but also continued to increase their score over time to the point where other locations couldn't compete. The graph implies that regardless of the number of times a student from a rural high school takes the verbal SAT, their score will on average still be significantly lower than that of a student who attended a suburban high school.

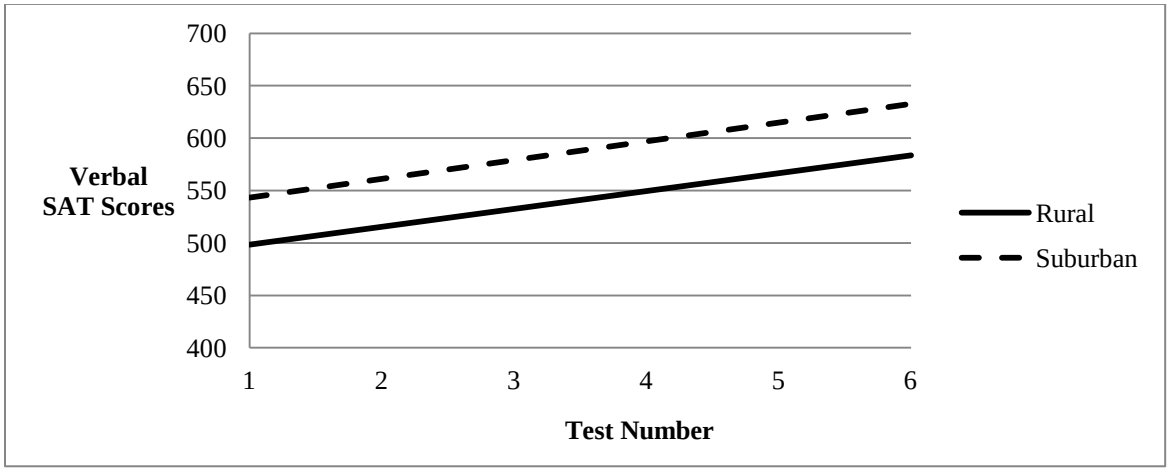

Figure 4.9 *High School Location and the Relationship between Linear Time and Verbal SAT scores*

Quadratic time. When the quadratic effective was added to the model, there continued to be a significant increase in linear verbal SAT scores over time (γ_{10}) and math

SAT scores over linear time (γ_{10}). There was also a significant quadratic effect for both verbal SAT scores (γ_{20}) and math SAT scores (γ_{20}). Table 4.8 shows the results from the intercepts and slopes as outcome model when quadratic time was included. With the addition of quadratic time to the model, there remained significant ethnic ($\gamma_{01,}$ $\gamma_{02,}$ $\gamma_{03,}$ $\gamma_{04,}$ $\gamma_{05,}$ $\gamma_{06,}$ $\gamma_{07,}$ γ_{08}), gender ($\gamma_{09,}$ γ_{110}), need for financial aid ($\gamma_{011,}$ $\gamma_{012,}$ γ_{013}), and high school location ($\gamma_{014,}$ $\gamma_{015,}$ $\gamma_{016,}$ $\gamma_{017,}$ γ_{018}) differences in the average math SAT scores the first time a student took the test in comparison to the referent group (White males who attended a suburban high school and would receive financial aid). Likewise, there remain significant ethnic ($\gamma_{01,}$ $\gamma_{02,}$ $\gamma_{03,}$ $\gamma_{04,}$ $\gamma_{05,}$ $\gamma_{06,}$ $\gamma_{07,}$ γ_{08}), gender (γ_{110}), need for financial aid ($\gamma_{011,}$ $\gamma_{012,}$ γ_{013}), and high school location ($\gamma_{014,}$ $\gamma_{015,}$ $\gamma_{017,}$ γ_{018}) differences in the average verbal SAT scores between students; with the exception of females (γ_{09}) and those who attended a high school located in a large city (γ_{016}) which were not significant.

Table 4.8

Intercepts and Slopes as Outcomes Quadratic Results

Fixed Effects			
Math SAT Scores, β_0		Verbal SAT Scores, β_0	
Intercept, γ_{00}	589.24***	Intercept, γ_{00}	552.85***
Ethnicity		Ethnicity	
American Indian/Alaskan Native, γ_{01}	-57.65***	American Indian/Alaskan Native, γ_{01}	-60.03***
Asian/Asian American/Pacific Islander, γ_{02}	19.72***	Asian/Asian American/Pacific Islander, γ_{02}	-35.30***
Black/African American, γ_{03}	-110.07***	Black/African American, γ_{03}	-103.47***
Mexican/Mexican American, γ_{04}	-53.74***	Mexican/Mexican American, γ_{04}	-53.69***
Puerto Rican, γ_{05}	-59.37***	Puerto Rican, γ_{05}	-51.67***
Other Hispanic/Latino/Latin American, γ_{06}	-47.53***	Other Hispanic/Latino/Latin American, γ_{06}	-45.09***
Other, γ_{07}	-37.78***	Other, γ_{07}	-35.57***
Unreported, γ_{08}	-23.87***	Unreported, γ_{08}	-15.35***
Gender		Gender	
Female, γ_{09}	-26.85***	Female, γ_{09}	0.23
Unreported, γ_{010}	-51.11***	Unreported, γ_{010}	-46.35***
Receive Financial Aid		Receive Financial Aid	
No, γ_{011}	10.53***	No, γ_{011}	2.42*
Unknown, γ_{012}	7.29***	Unknown, γ_{012}	3.25***
Unreported, γ_{013}	2.71*	Unreported, γ_{013}	-4.71***
High School Location		High School Location	
Small city or town, γ_{014}	-19.23***	Small city or town, γ_{014}	-17.43***
Medium-sized city, γ_{015}	-3.80***	Medium-sized city, γ_{015}	4.15***
Large city, γ_{016}	-8.51***	Large city, γ_{016}	-0.48

Table 4.8 Continued

Rural, γ_{017}	-32.92***	Rural, γ_{017}	-33.40***
Unreported, γ_{018}	-23.13***	Unreported, γ_{018}	-18.38***
Linear Time, β_1		Linear Time, β_1	
Intercept, γ_{10}	23.04***	Intercept, γ_{10}	16.42***
Ethnicity		Ethnicity	
American Indian/Alaskan Native, γ_{11}	8.53	American Indian/Alaskan Native, γ_{11}	9.93*
Asian/Asian American/Pacific Islander, γ_{12}	1.78	Asian/Asian American/Pacific Islander, γ_{12}	7.30***
Black/African American, γ_{13}	11.69***	Black/African American, γ_{13}	13.05***
Mexican/Mexican American, γ_{14}	5.82	Mexican/Mexican American, γ_{14}	9.48
Puerto Rican, γ_{15}	12.66	Puerto Rican, γ_{15}	24.03***
Other Hispanic/Latino/Latin American, γ_{16}	8.96*	Other Hispanic/Latino/Latin American, γ_{16}	15.24***
Other, γ_{17}	10.69***	Other, γ_{17}	12.59***
Unreported, γ_{18}	14.41***	Unreported, γ_{18}	15.70***
Gender		Gender	
Female, γ_{19}	-3.53***	Female, γ_{19}	-1.16
Unreported, γ_{110}	-18.91	Unreported, γ_{110}	14.48
Receive Financial Aid		Receive Financial Aid	
No, γ_{111}	-8.96***	No, γ_{111}	-7.42***
Unknown, γ_{112}	-1.02	Unknown, γ_{112}	-1.72
Unreported, γ_{113}	3.74*	Unreported, γ_{113}	2.79
High School Location		High School Location	
Small city or town, γ_{114}	4.37**	Small city or town, γ_{114}	5.57***
Medium-sized city, γ_{115}	3.98**	Medium-sized city, γ_{115}	4.95**
Large city, γ_{116}	4.34*	Large city, γ_{116}	6.18***
Rural, γ_{117}	5.13***	Rural, γ_{117}	6.04***

Table 4.8 Continued

Unreported, γ_{118}	4.33**	Unreported, γ_{118}	4.15*
Quadratic Time, β_2		Quadratic Time, β_2	
Intercept, γ_{20}	-3.97***	Intercept, γ_{20}	-3.61***
Ethnicity		Ethnicity	
American Indian/Alaskan Native, γ_{21}	-0.68	American Indian/Alaskan Native, γ_{21}	-1.64
Asian/Asian American/Pacific Islander, γ_{22}	0.47	Asian/Asian American/Pacific Islander, γ_{22}	-0.44
Black/African American, γ_{23}	-1.39**	Black/African American, γ_{23}	-1.45
Mexican/Mexican American, γ_{24}	2.02	Mexican/Mexican American, γ_{24}	2.24
Puerto Rican, γ_{25}	-3.36	Puerto Rican, γ_{25}	-6.15*
Other Hispanic/Latino/Latin American, γ_{26}	0.64	Other Hispanic/Latino/Latin American, γ_{26}	-1.40
Other, γ_{27}	-0.80	Other, γ_{27}	-2.58*
Unreported, γ_{28}	-2.75***	Unreported, γ_{28}	-3.27***
Gender		Gender	
Female, γ_{29}	0.82**	Female, γ_{29}	0.55
Unreported, γ_{210}	8.33*	Unreported, γ_{210}	-1.07
Receive Financial Aid		Receive Financial Aid	
No, γ_{211}	1.37*	No, γ_{211}	1.27*
Unknown, γ_{212}	-0.49	Unknown, γ_{212}	0.18
Unreported, γ_{213}	-1.25*	Unreported, γ_{213}	-0.77
High School Location		High School Location	
Small city or town, γ_{214}	0.26	Small city or town, γ_{214}	-0.32
Medium-sized city, γ_{215}	0.51	Medium-sized city, γ_{215}	-0.25
Large city, γ_{216}	-0.48	Large city, γ_{216}	-1.21
Rural, γ_{217}	0.24	Rural, γ_{217}	0.07

Table 4.8 Continued

Unreported, γ_{218}	-0.00	Unreported, γ_{218}	0.17
Random Effects			
Math Level 2 Variability, τ_{00}	7942.32***	Verbal Level 2 Variability, τ_{00}	7970.50***
Covariance (Intercept & Linear), τ_{10}	-1188.16***	Covariance (Intercept & Linear), τ_{10}	-929.25***
Math Linear Slope, τ_{11}	811.65***	Verbal SAT Linear Slope, τ_{11}	624.96***
Covariance (Intercept & Quadratic), τ_{20}	113.04***	Covariance (Intercept & Quadratic), τ_{20}	107.64***
Covariance (Linear & Quadratic), τ_{21}	-119.85***	Covariance (Linear & Quadratic), τ_{21}	-101.61***
Math Quadratic Slope, τ_{22}	20.09***	Verbal Quadratic Slope, τ_{22}	19.00***
Within-person fluctuation, σ^2	492.76***	Within-person fluctuation, σ^2	450.74***

Note: *p<.05, **p<.01, ***p<.001

Ethnicity

The results show that when quadratic time was added to the model, the relationship between changes in math SAT scores and American Indian/Alaskan Native (γ_{12}), Mexican/Mexican American (γ_{14}), and Puerto Rican (γ_{16}) ethnicities changed. When the model did not take into account quadratic time, the relationship between changes in math SAT scores and these ethnicities (vs. White) was significant; however, with the addition of quadratic time, this relationship became non-significant. All other relationships remained the same, and there was still not a significant relationship between Asian/Asian American/Pacific Islanders (γ_{13}) and White and changes in math SAT scores over time. There continued to be a positive relationship between Black/African American (γ_{13}), Other Hispanic/Latino/Latin American (γ_{16}), Other (γ_{17}) and Unreported (γ_{18}) ethnicities and changes in math SAT scores over time vs. White. With the addition of quadratic time, the relationship between verbal SAT scores and Mexican/Mexican American (γ_{14}) ethnicity vs. White became non-significant. All other ethnicities ($\gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{15}, \gamma_{16}, \gamma_{17}, \gamma_{18}$) continued to have a significant relationship with changes in verbal SAT scores over time vs. White.

Consistent with the linear-only model, there was a significant relationship between females (γ_{19}) and changes in math SAT scores over time vs. males; however, there was not a significant relationship between unreported gender (γ_{110}) and changes in math SAT scores over time vs. males. Likewise, there was not a significant relationship between female (γ_{19}) and unreported (γ_{110}) gender and changes in verbal SAT scores over time vs. males.

With the addition of quadratic time, the relationship between students with an unknown financial aid status (γ_{112}) vs. those who need financial aid and changes in both math and verbal SAT scores over time became non-significant. Conversely, unreported financial aid status (γ_{113}) and changes in math SAT scores became significant with the addition of quadratic time. All other results remained the same as the linear-only model. The relationship between those who do not receive financial aid (γ_{111}) and changes in both math and verbal SAT scores over time stayed significant. The relationship between unreported financial aid status (γ_{113}) and changes in verbal SAT scores over time vs. those who need financial aid remained non-significant with the addition of quadratic time. Consistent with the linear-only model, there was a significant positive relationship between both math and verbal SAT scores over time and the high school location of a student (γ_{114}, γ_{115}, γ_{117}, γ_{118}).

The only ethnicities that had a significant negative quadratic effect on the math SAT were Black/African American (γ_{23}) and unreported (γ_{28}) vs. White. Black/African American was chosen as a representative of these two variables since the trend for these variables was the same. Figure 4.10 shows that Black/African American students' math SAT score tended to increase the first four times they took the test. Any iteration of the test after the fourth on average caused a decrease in the math SAT score. All other ethnicities (γ_{21}, γ_{22}, γ_{24}, γ_{25}, γ_{26}, γ_{27}) had non-significant quadratic effects when compared to Whites, indicating that they had a similar slope of change over time (like Whites). It is important to note that this refers to the slope only; it does not refer to the level or intercept of the score. Separate models were conducted for Blacks and Whites. The average is represented in Figure 4.10.

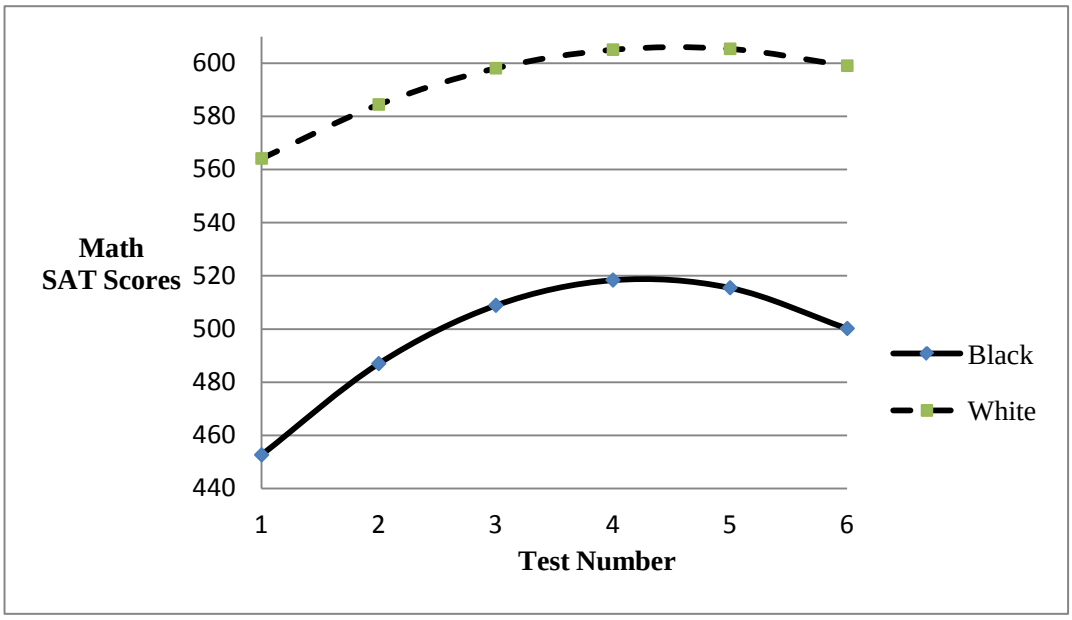

Figure 4.10 *Ethnicity and the Relationship between Quadratic Time and Math SAT Scores (the lines represent average trends based on predicted points)*

Puerto Rican (γ_{25}), other (γ_{27}), and unreported (γ_{28}) were the only ethnicities that had a significant negative quadratic effect vs. White on the verbal portion of the SAT. Therefore, Puerto Rican was chosen as a representative, since the trend was the same across these variables. Figure 4.11 shows that the point at which Puerto Ricans should stop repeating the verbal SAT tends to be the fourth time. It is evident that scores increased the first three times the verbal SAT was taken and the score tended to decrease at the fourth attempt. This is in line with the referent group, as the point at which Whites' should stop repeating the verbal SAT tends to also be the fourth time. There was a significant variance around the slope, indicating that there will be some students whose SAT scores deviate from the predicted trajectory.

All other ethnicities ($\gamma_{21}, \gamma_{22}, \gamma_{23}, \gamma_{24}, \gamma_{26}$,) had non-significant quadratic effects, indicating that there was no difference compared to Whites in the point at which taking the SAT additional times would result in the verbal SAT score decreasing. This implies that that these students should take the test as many times as possible because their score on average would increase with each sitting.

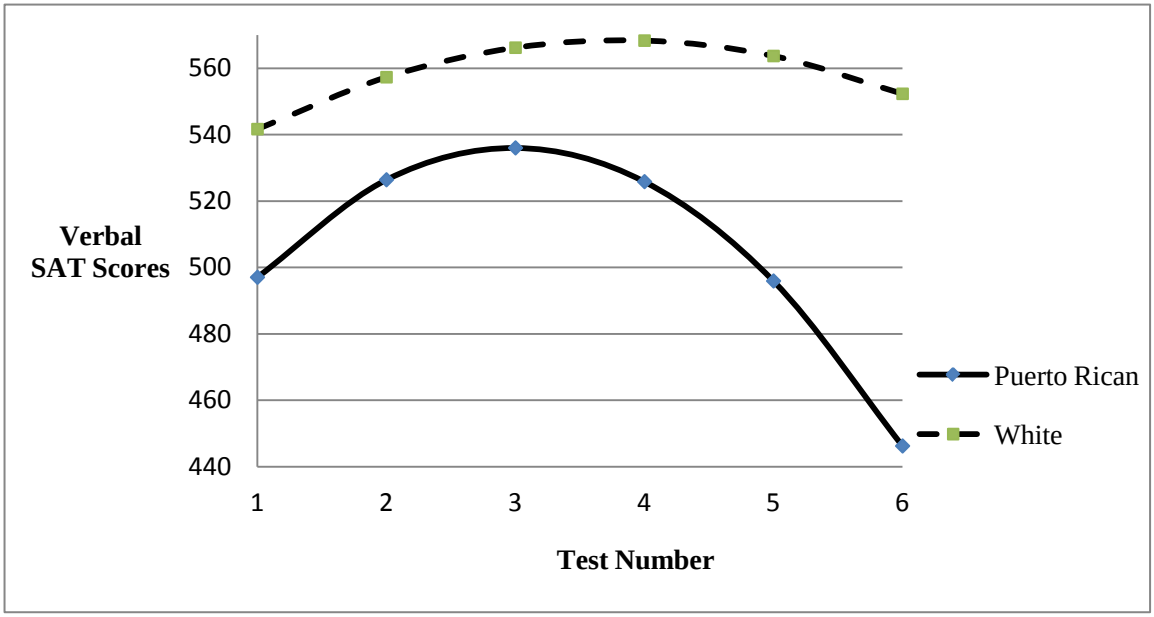

Figure 4.11 *Ethnicity and the Relationship between Quadratic Time and Verbal SAT Scores (the lines represent average trends based on predicted points)*

Gender

There was a significant positive relationship with females (γ_{29}) and students whose gender was unreported (γ_{210}) and changes in math SAT scores over quadratic time vs. men, which is displayed in Figure 4.12 with females being the representative gender group. It is

evident that, in general, females who took the math SAT more than five times actually decreased their score. Therefore, the sixth time was the point at which taking the SAT additional times resulted in the math SAT score decreasing for females. These results were almost parallel to the results from the referent group, male. This premise is further supported by the fact that the variance around the math slope was non-significant, indicating that the majority of the students in this sample will follow the predicted path displayed in Figure 4.12.

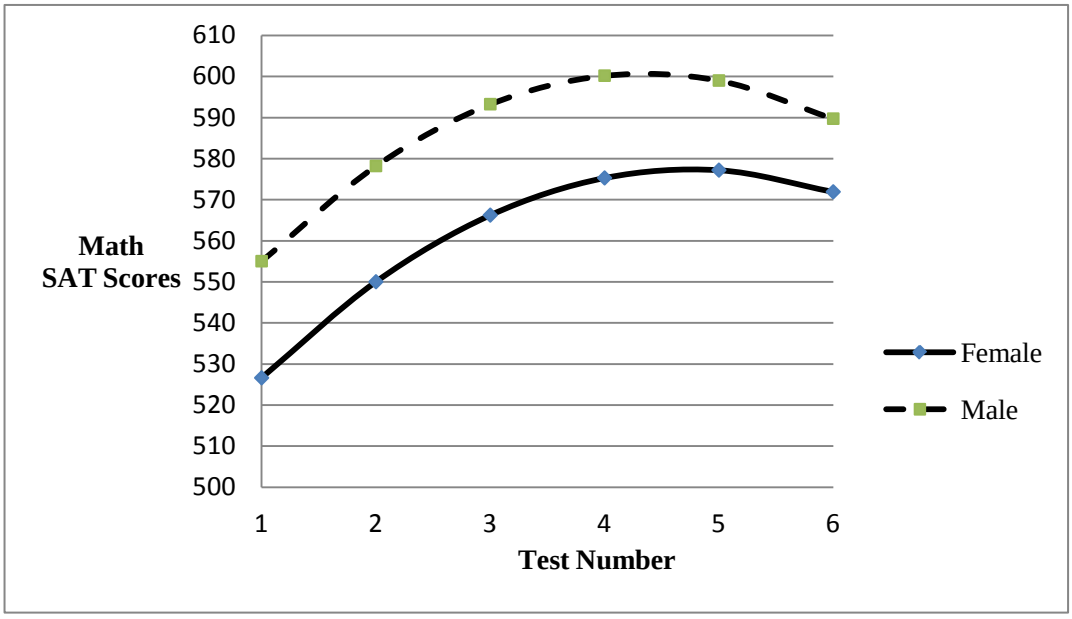

Figure 4.12 *Gender and the Relationship between Quadratic Time and Math SAT Scores (the lines represent average trends based on predicted points)*

Females (γ_{29}) and students whose gender was unreported (γ_{210}) had non-significant quadratic effects on the verbal SAT. This indicates that there were no gender differences in

the point at which the verbal SAT score would decrease with each additional attempt of the SAT.

Financial Aid

Students who reported they would not receive financial aid (γ_{211}) had a significant positive quadratic effect on the math SAT vs. those who need financial aid. Figure 4.13 shows that when students who stated they would not receive financial aid took the math SAT more than five times, their score tended to decrease with an additional attempt. Figure 4.13 indicates that students' taking the math SAT more than five times will most likely result in a decrease in their score. Those who had an unreported need for financial aid (γ_{213}) had a significant negative quadratic effect on the math SAT, whereas students with an unknown need for financial aid (γ_{212}) had a non-significant quadratic effect vs. students' who did need financial aid. This indicates that there was no difference in the point at which repeating the SAT would cause a decrease in the math SAT compared to students who needed financial aid.

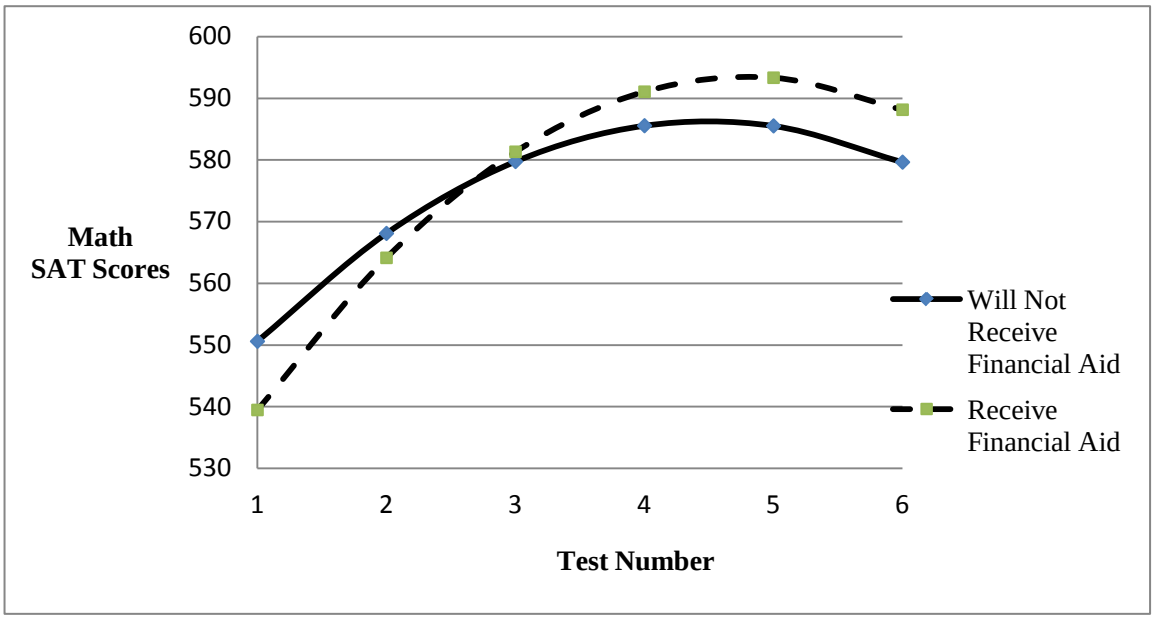

Figure 4.13 *Need for Financial Aid and Relationship between Quadratic Time and Math SAT Scores. (The lines represent average trends based on predicted points)*

Students who reported they would not receive financial aid (γ_{211}) had a significant positive quadratic effect vs. those who need financial aid on the verbal SAT. Figure 4.14 shows that when students who stated they would not receive financial aid took the verbal SAT more than four times, their score tended to decrease with each additional attempt. Therefore, one could conclude that these students should never take the verbal SAT more than four times. Those who had an unknown need for financial aid (γ_{212}) and those with an unreported need for financial aid (γ_{213}) had a non-significant quadratic effect.

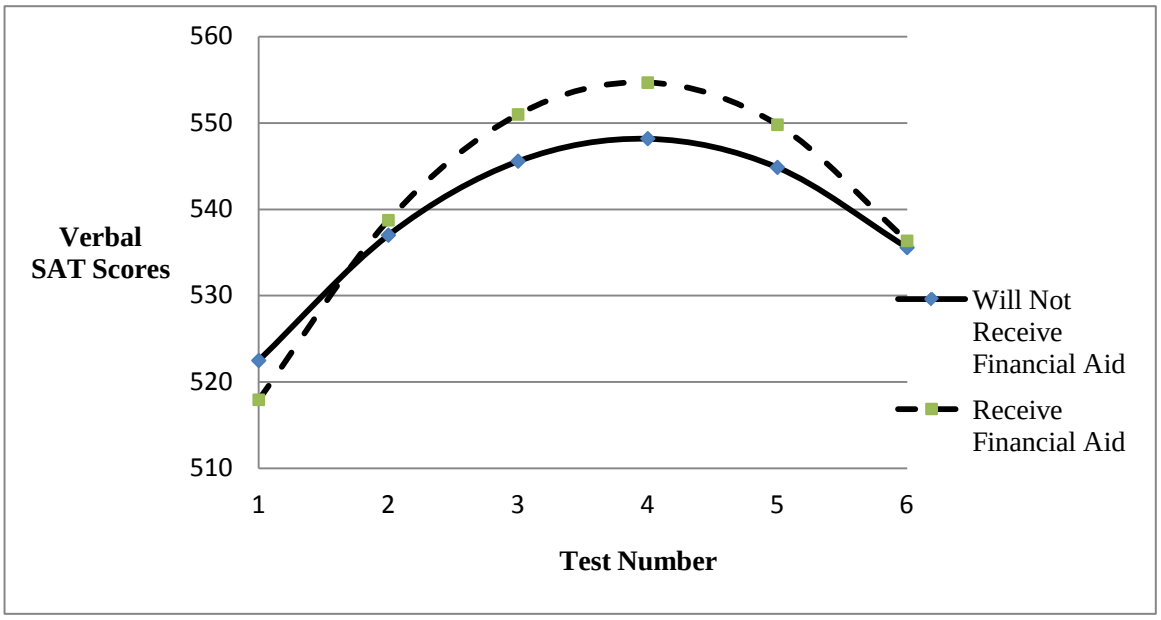

Figure 4.14 *Need for Financial Aid and the Relationship between Quadratic Time and Verbal SAT Scores (the lines represent average trends based on predicted points)*

High School Location

There were no significant quadratic effects regarding high school location (γ_{214}, γ_{215}, γ_{216}, γ_{217}, γ_{218}) vs. suburban on either the math SAT or the verbal SAT.

Random Effects

There was significant between-person variability in math SAT scores (τ_{00}) and verbal SAT scores (τ_{00}). The covariance between math SAT scores and the intercept (τ_{10}) and the covariance between verbal SAT scores and the intercept (τ_{10}) were significant, indicating the intercept and the slope were dependent upon one another. There was a significant amount of variance around both the slope of math SAT scores (τ_{11}) and the slope of verbal SAT scores

(τ_{11}). There was also significant residual variance within students in math SAT scores (σ^2) and verbal SAT scores (σ^2).

The significance of the random effects in the equation did not alter when quadratic time was added to the model. There continued to be significant between-person variability in math SAT scores (τ_{00}) and verbal SAT scores (τ_{00}).The covariance between linear time and the math SAT intercept (τ_{01}) and the covariance between linear time and the verbal SAT intercept (τ_{01}) were significant, indicating the intercept and the slope were dependent upon one another. The covariance between quadratic time and the math SAT intercept (τ_{20}) and the covariance between quadratic time and the verbal SAT intercept (τ_{20}) were significant. The covariance of linear time with quadratic time for the math (τ_{21}) and verbal (τ_{21}) portion of the SAT were significant; indicating that the score a student initially received on the SAT is related to the final score they achieved. There was not a significant amount of variance around the linear slope (τ_{11}) or quadratic slope (τ_{22}) of math SAT scores but there was a significant variance around the linear slope (τ_{11}) and quadratic slope (τ_{22}) of verbal SAT scores.

The linear model accounted for of the 6.24% between-person variability in math SAT scores and -0.23% in verbal SAT scores. It accounted for 60.63% of the within-person variability in math SAT scores and 66.91% of verbal SAT scores. When quadratic time was added to the model it accounted for 6.11% of the between-person variability in math SAT scores and <0.01% in verbal SAT scores. It accounted for 63.05% of the within-person variability in math SAT scores and 66.69% of verbal SAT scores.

Research Question 5

Question 5: Is a student's score that would be reported through Score Choice related to ethnicity, gender, financial aid, or location of a student's high school?

SAT Scores = a + b_1(Ethnicity American Indian/Alaskan Native) + b_2(Ethnicity Asian/Asian American/Pacific Islander) + b_3(Ethnicity Black/African American) + b_4(Ethnicity Mexican/Mexican American) + b_5(Ethnicity Puerto Rican) + b_6(Ethnicity Other Hispanic/Latino/Latin American) + b_7(Ethnicity Other) + b_8(Ethnicity Unreported) + b_9(Gender Female) + b_{10}(Gender Unreported) + b_{11}(Financial Aid No) + b_{12}(Financial Aid Unknown) + b_{13}(Financial Aid Unreported) + b_{14}(High School in small city/town) + b_{15}(High School in medium-sized city) + b_{16}(High School in large city) + b_{17}(High School in rural area) + b_{18}(High School unreported)

A linear regression was used to answer the fifth research question. The variable SAT scores used in this equation refer to the combined math and verbal SAT scores. The overall model was significant, $R = .47$, $F(18,151,882) = 2341.64$, $p < .001$, showing that the combination of the independent variables (high school location, gender, need for financial aid and ethnicity) significantly predicted students' SAT scores. Students' high school location, gender, need for financial aid and ethnicity accounted for 21.7% of the variance in scores they receive on the SAT if using the Score Choice option. Therefore, to answer the research question, a student's highest SAT score, the score that would be reported through Score Choice, was, indeed, related to ethnicity, gender, financial aid, and location of a student's high school.

Each of the independent variables had a statistically significant effect on SAT scores except Asian ethnicity and high school location of large city. Neither of these variables was statistically significant, demonstrating they were not differentially related to SAT scores compared to the referent group. The constant (referent group) used in this equation was a student who was white, male, attended a suburban high school, and answered "yes" to the question about whether they needed financial aid. If a student meeting these criteria took the SAT and used the Score Choice option, this equation predicted they would score 1159.64 on their highest sitting of the SAT. If this student attended a high school in a small city or town, rural area, or "other" location as opposed to a suburban one, he would score lower on the SAT (their score would decrease by 30.04 points, 60.54 points, and 41.89 points respectively), when holding need for financial aid, gender, and ethnicity constant. However, if they attended a high school in a medium city or town, their score would increase by 10.21 points. There was not a significant statistical relationship between SAT score and high school location of large city vs. suburban.

These results suggest that the highest score a student received in an individual sitting of the SAT (the score reported through Score Choice) was related to their gender, when controlling for ethnicity, need for financial aid, and high school location. This specifically shows that females would receive almost 27 points lower on their SAT score than their male counterparts when utilizing the Score Choice option. Those who reported their gender as "unknown" tend to score 107 points lower than males.

These results also indicated that ethnicity had a significant influence on SAT scores reported via Score Choice when controlling for gender, need for financial aid, and high

school location. White students scored higher than all other ethnicities and Black students scored the lowest at almost 214 points lower than their White counterparts. Students with an ethnicity of Black also scored almost 100 points less than any other ethnicity on the SAT, with American Indian/Alaskan Natives, Mexican/Mexican Americans, Puerto Ricans, Other Hispanic/Latino/Latin Americans, Other, and Unreported ethnicities scoring respectively about 115 points, 106 points, and 104 points, 83 points, 66 points, and 26 points less than their White counterparts. There was not a significant statistical difference between SAT scores and ethnicity of Asian/Asian American/Pacific Islander compared to the referent group (White).

Need for financial aid was found to have a significant influence on SAT scores reported via Score Choice, when controlling for gender, ethnicity, and high school location. The magnitude of the score difference was relatively small, as a student who answered that they did not need financial aid was predicted to receive about nine points higher on their SAT score, via Score Choice, than a student who reported they needed financial aid. Likewise, a student who was unknown regarding whether they had a need for financial aid would get almost twelve points higher on their SAT and a student who didn't answer this question on the SAT questionnaire would receive about four points higher on their SAT.

Table 4.9

Regression Results

	B	SE	*t*-value
Constant (White, Male, Suburban, Yes Financial Aid)	1159.64	1.11	1048.67***
High School Location			
Large city	-2.99	1.54	-1.94
Medium city or town	10.21	1.42	7.18***
Small city or town	-30.04	1.34	-22.50***
Rural	-60.54	1.30	-46.60***
Other	-41.89	1.44	-29.07***
Gender			
Female	-26.50	0.84	-31.64***
Unknown	-107.19	9.82	-10.92***
Ethnicity			
American Indian/Alaskan Native	-115.14	4.50	-25.58***
Asian/Asian American/Pacific Islander	-1.30	1.98	-0.66
Black/African American	-213.89	1.17	-183.00***
Mexican/Mexican American	-106.14	4.76	-22.29***
Puerto Rican	-104.48	5.40	-19.36***
Other Hispanic/Latino/Latin American	-83.34	3.13	-26.60***
Other	-65.89	2.78	-23.71***
Unreported	-26.27	1.94	-13.56***
Receive Financial Aid			
No	9.14	1.64	5.58***
Unknown	11.55	1.06	10.84***
Unreported	4.36	1.67	2.61**

Adjusted R^2 = 0.22 (n = 151,900, p <0.05)

Note: *p<.05, **p<.01, ***p<.001

Summary

Chapter 4 has summarized the results from conducting multilevel modeling and regression analysis on the data. Chapter 5 will discuss these findings in relation to the

research questions and the higher education literature. It will also discuss some limitations associated with the current study and areas for further research in respect to university admissions and high stakes testing.

Chapter 5: Discussion

Introduction

The following chapter is divided into four subsections. The first discusses the main findings and how they relate to each research question. The second section examines the implications of these findings for higher educational institutions and recommendations. The third section shows the limitations of the current study. The fourth section suggests potential areas of study for future researchers. The final section provides a conclusion.

The overall purpose of this study was to determine whether there is a significant relationship between repeated test taking on the SAT and ethnicity, gender, need for financial aid, or high school location for North Carolina State University applicants. A secondary goal was to investigate what implications the new SAT reporting option, Score Choice, could have on future applicants' scores. This is an area of research that has lacked empirical data because the Score Choice policy was implemented with the incoming freshman class of 2010.

Review of the Findings

The initial analysis, the fully unconditional model, indicated that there was significant between-person and within-person variability in both math and verbal SAT scores. This shows that there was, indeed, a change in SAT scores over time. Furthermore, the results indicated that there was greater variability in scores between individuals, as opposed to changes in individual scores. This shows that there was a greater difference in SAT scores over time between Student A and Student B than changes in Student A's individual scores

over time. From these findings, the conclusion may be drawn that it is important to know how many times a student has taken the SAT, as their scores do significantly change over time. However, it is more important to take into account other factors regarding the identification of a student, as there was a significant difference between individual students' scores.

The first research question asked if on average there was a change in SAT scores associated with repeated testing. It is evident from the results that the answer to this question is yes; there was a significant increase in both math and verbal SAT scores over time (where time refers to the test number). Burke (1997) indicates that the reason SAT scores increase over time may be due to retesting. The implication of this finding is large, as it implies that the number of times students take the SAT could significantly impact their score. With the implementation of Score Choice, universities may not be privy to the information regarding the number of times a student has taken the SAT, thus may compare students who have taken the test a plethora of times to those who have only taken it once. According to Gleason (2010), "the exact mission of land-grant institutions is open to discussion, but there seems to be a consensus that they have a special obligation to provide high-quality education for citizens of their home state, as well as to focus attention on the economic development and social welfare of that state" (p. A88). Therefore, as a land-grant institution, North Carolina State University has an obligation to serve the citizens of North Carolina. In order to give NC citizens an equitable chance at obtaining a higher education from this institution, it is important that the admissions office take into account the fact that the number of times someone takes the SAT is significantly related to their final score. Currently, North Carolina

State University has an admission policy requiring applicants send all their SAT scores; however, with the implementation of Score Choice, North Carolina State University will only receive the scores that students specifically send them. Moreover, since the release of information by the College Board requires student consent, the university will have no way of knowing whether a student complied with their policy.

The second research question asked if there was a point at which scores decrease with each additional attempt of the SAT. The findings indicate that the answer is no, there is a continued improvement (see Figure 4.1 and Figure 4.2). On average, the students in this study scored 532.70 points on the math portion of the SAT the first time they took it and 509.16 points on the verbal portion of the SAT the first time they took the test. These scores incrementally increased each additional time a student took the test. However, it is also evident that there were diminishing returns associated with the number of times a student took the SAT. Basically, a student's potential score increase tended to be the greatest between their first and second sitting of the math or verbal portion of SAT. According to the results, on average their score would tend to increase between their fifth and the sixth sitting of the test, but the magnitude of the increase was much smaller than that between their first and second sitting. It is important to note that these trends are based on general quadratic patterns and do not yet look at ethnic, gender, need for financial aid, or high school location differences.

The question that arises from these results is whether, in reality, this would matter. Is it worth a student taking the SAT six times when, on average, this would constitute about a 57-point increase in their verbal SAT score and a 53.3 increase in their math SAT score from

the first time they took it? Most would probably agree that an overall 110-point total SAT score increase could potentially heighten their chances of being admitted to a university. However, would it be more beneficial for this student to save his/her money (as there is a fee for taking the SAT) and only take the SAT twice because according to the results, between the first and second sitting of the SAT is when their scores will have the greatest increase? It also begs the question whether the fee waiver policy is enough. It allows for underprivileged students to receive two fee waivers for the SAT. Would this be enough, or should they receive more?

In order to determine whether the difference in scores was due to students' ethnicity, gender, financial aid, or location of their high school, research question three was posed: Does the relationship between repeated testing and SAT scores depend on ethnicity, gender, financial aid, or location of a student's high school? The results indicate that there were indeed significant gender, ethnic, need for financial aid, and high school location differences in the average math and average verbal SAT scores between students. These findings are consistent with the literature: a student's gender (Mau et al., 2001; Young et al., 2000), ethnicity (Schmidt, 2007), socioeconomic status (Bowen et al., 1998), and location of high school attended (Chenoweth, 1998) can influence SAT scores.

There was a significant relationship between all ethnicities (vs. White) and verbal SAT scores. These findings were aligned with the literature regarding high stakes testing. Schmidt (2007) found the average score of a White student tend to be approximately 200 points greater than their Black counterpart. This study rendered very similar results, as Whites scored on average scored almost 111 points higher on the math portion of the SAT

and almost 104 points higher on the verbal portion of the SAT the first time they took the test vs. Blacks. The combination of these scores result in approximately a 215-point difference between Whites' and their Black counterparts the first time they took the SAT (combination of math and verbal). Furthermore, Schmidt (2007) stated "many of the objections to the SAT stem from the tendency of Blacks, Hispanics, and Native Americans to score lower than Whites, even when compared to people at the same income level" (p. 151). Again, this study found similar results. Black/African American students scored the lowest on both the math and verbal portion of the SAT the first time they took the test. American Indian/Alaskan Natives, Mexicans/Mexican Americans, Puerto Ricans, and Other Hispanics/Latinos/Latin Americans also all scored significantly lower than their White counterparts the first time they took the SAT.

The findings indicate that math SAT scores for Asian/Asian American/Pacific Islanders did not significantly increase vs. Whites with subsequent attempts. All other ethnicities' scores increased with additional attempts vs. Whites. This implies that students with ethnicities other than Asian/Asian American/Pacific Islander will most likely increase their math scores by taking the SAT multiple times. Likewise, all ethnicities scores increased over time on the verbal portion of the SAT.

The graphs (Figure 4.3 and Figure 4.4) comparing the representative ethnicity Mexican/Mexican American students to White students, the referent value, show that Mexican/Mexican American students increased their SAT scores over time at a higher rate than their White counterparts. Mexican/Mexican American students obtained a lower initial SAT score, but their score increased at a higher rate to White students', eventually surpassing

the White students' scores. A Mexican/Mexican American student on average would have to take the math SAT six times and the verbal SAT four times to achieve the same SAT score as a White student.

In general, males scored higher than females on the math portion of the SAT the first time they took the test. Furthermore, females tended to increase their score at a similar rate to males over time, thus preventing females from achieving the same score as a male who took the math SAT the same number of times (see Figure 4.5). These findings are in line with the previous research regarding entrance exams and gender. Mau et al. (2001) found that males scored significantly higher on university entrance exams; FairTest (2007) and Young (1991) showed that admission tests tend to under predict females university ability. There has been a long debate regarding females' ability in the math arena. Quaid (2008) reported on a study conducted by Janet Hyde which analyzed the results from seven million children's annual math tests in 2002. The findings from this particular study showed that the math ability of both genders was equivalent, even in high school. Hyde stated that "girls have now achieved gender parity in performance on standardized math tests" (Quaid, 2008, p.12). Hyde attributed the discrepancy between male and female SAT scores to the volume of females taking the test stating, "More of them take the test, which is needed to get into college. The highest-performing students of both genders take the test, but more girls lower on the achievement scale take it, skewing the average" (Quaid, 2008, p. 12). There are differing explanations regarding why on average females tend to score lower than their male counterpart on the SAT, but the overall conclusion is the same – females tend to score lower.

Students who reported they would not receive financial aid or whose need for financial aid was unknown on average scored significantly higher on the math and verbal SAT the first time they took the test than those who did need financial aid. However, once the results were graphed (Figure 4.6 and Figure 4.7), it was evident that over time students who did need financial aid surpassed the students who did not need it or whose need was unknown. These results seem somewhat contradictory to the literature that concludes students from wealthier backgrounds tend to do better on the SAT (Rampell, 2009). One explanation for these results is that the variable used to determine need for financial aid did not distinguish between merit-based aid and needs-based aid. Therefore, students who did receive aid do not necessarily represent a lower socioeconomic background, as the aid they received may have been in the form of scholarships.

There was a significant relationship between math and verbal SAT scores and the high school location of a student. Students whose high school was located somewhere other than a suburban location were found to receive a lower math SAT score than those who attended a suburban high school the first time they took the test. Likewise, students whose high school was located in a small city or town, rural area, or whose high school location was unreported all scored significantly lower on the verbal SAT than those whose high school was in a suburban location the first time they took the test. However, those who attended high school in a medium-sized city scored significantly higher on the verbal portion of the SAT the first time they took the test. Rural high school students scored the lowest of all high school locations on both the verbal and math portion of the SAT the first time they took the test.

Although the SAT scores of students increased over time regardless of the high school location, it was evident that even after taking the SAT (math and verbal) six times, a student from a rural high school could not attain an equivalent score to a suburban high school student (see Figure 4.8 and Figure 4.9). According to Fischer and Hebel (2006), students from rural schools "on average, score below those in suburban schools in both mathematics and reading on standardized tests, and they are less likely to earn high-school diplomas" (p. A20). One reason verbal SAT scores tend to be lower for rural students may be because English is a second language to many rural students: "A comparison of recent immigrants in rural areas suggests that, compared to their more urban counterparts, they are more likely to be Hispanic (and Mexican-origin in particular)" (Jensen, 2006, p. 7). The focus of the verbal portion of the SAT is English grammar and includes questions on sentence improvement, paragraph improvement, and identifying sentence errors. For someone who doesn't primarily speak English, this portion could be especially challenging:

> A distinguishing feature of immigration since the late 19th century has been its decided urban orientation. Throughout the 1970s and 1980s, immigrants entered and settled in gateway cities such as New York, Miami, Chicago, and Los Angeles. In recent years, however, that appears to be changing. A flurry of research and popular observation has confirmed that immigrants are beginning to settle in new destinations and gateway communities. Although many of these are smaller metropolitan areas, a great many small towns and rural areas are seeing an influx of new immigrants. (Jensen, 2006, p. 8)

Question four looked at whether the point at which students SAT scores decreased with each additional attempt of the SAT depended upon ethnicity, gender, financial aid, or location of a student's high school. The results showed that some of the quadratic effects were significant. This indicates that there were some ethnic, gender, need for financial aid, and high school location differences in the point at which scores decreased with each additional attempt of the SAT.

The only ethnicities that had a significant quadratic effect on the math SAT were Black/African American and unreported, with Black/African American students scores tending to decrease after the fourth time they took the test vs. Whites. Puerto Ricans, Other, and Unreported were the ethnicities that had a significant quadratic effect on the verbal SAT vs. Whites, with Puerto Rican scores tending to decrease after the third time they took the verbal SAT. All other ethnicities were found to have a non-significant quadratic effect vs. Whites, indicating that their scores would tend to increase each time they repeated the SAT. Therefore, it would tend to be beneficial for a White student to take the math portion of the SAT five or six times, whereas a Black/African American student should not take the test more than four times, as their score would most likely decrease as opposed to increase after their fourth attempt.

Females and students whose gender was unreported had non-significant quadratic effects on the verbal SAT vs. males, but significant effects on the math SAT. Figure 4.12 shows that in general, females tend to score less on the math SAT the sixth time they take it in comparison to the fifth score they received. However, there was no point of at which the verbal SAT scores decreased with each additional attempt of the SAT vs. males.

Figure 4.13 and Figure 4.14 show the relationship between quadratic time and math/verbal SAT scores for students who would not receive financial aid. Both figures indicate that in general, these students will increase their score until the sixth time they take the SAT, at which point their score will most likely decline.

The trend that is noticeable when considering the quadratic effect is that students' scores increased the first three times they took the SAT, regardless of ethnicity, gender, need for financial aid, and high school location. Even when there were significant quadratic effects found, a student had to take the SAT in excess of three times to get to the point where the score would begin to decrease with each additional attempt. This finding is especially important in light of the implementation of Score Choice. The best advice someone could give a student is to take the SAT a minimum of three times, as it is likely their score will increase each time. Nevertheless, Bunin (2009) stated that repeating the SAT does not guarantee an increase in score. This equation looked at the SAT scores of the sample as a whole, so it is possible that some students' increased their scores while others may have decreased; however, the results indicate that it is likely an individual will increase their scores the first three times they take the SAT. This implies that Score Choice may be a disadvantage for a student who does not have the means to pay for multiple sittings of the test. Furthermore, it implies that the two fee waivers available to low socioeconomic students may not be enough, when others may have the opportunity to increase their score by taking the test more times due to their financial ability.

The final research question asked whether a student's SAT score that would be reported through Score Choice is related to ethnicity, gender, financial aid, or location of a

student's high school. The answer to this question is yes; a student's highest SAT score, the score that would be reported through Score Choice, is indeed related to ethnicity, gender, financial aid and location of a student's high school. According to the results, a White male who went to suburban high school and doesn't know whether he needs financial aid would on average report a score of 1171 to his university of choice when using the Score Choice option on the SAT. However, a Black student who reported his/her gender as unknown, grew up in a rural town, and needs financial aid would tend to report a score of 778 when using the Score Choice option on the SAT. These results are not extraordinarily surprising, as much of the literature indicates that there is a link between these variables and SAT scores (Chenoweth, 1998; FairTest, 2007; Herrnstein & Murray, 1994; Mau and Lynn, 2001; U.S. Department of Education, 2011; Young and Fisler, 2000). What is important to note is that the only score being analyzed in this research question is the highest score a student obtained during one sitting of the SAT (the score that would be reported via Score Choice).

Research suggests that males tend to score higher on standardized tests or college entrance exams than females (Mau et al., 2001; Sacks, 1999). The results imply this would still remain the case if Score Choice were used. Likewise, high school location would also align with the literary findings, that students from suburban schools tend to score the highest on the SAT and those from rural schools have lower average scores (Chenoweth, 1998), even when Score Choice was utilized.

Some of the variance in students' highest SAT scores was found to be related to ethnicity. The results show that when Score Choice is used, Whites and Asian/Asian American/Pacific Islanders on average remained the highest scoring ethnicities, while

Black/African Americans were again the lowest, scoring over 200 points lower than their White counterparts. Black/African Americans scores were followed consecutively (from lowest to highest) by American Indian/Alaskan Natives, Mexican/Mexican Americans, Puerto Ricans, Other Hispanic/Latino/Latin Americans, Other, and Unreported ethnicity. The order of these ethnicities in regards to their SAT scores is very similar to the findings obtained in the third research question. This indicates that Black/African American students tend to score lower than other ethnicities on the SAT regardless of how many times they take the test. However, if a university is unaware of the number of times a student has taken the SAT, it may be disadvantageous to a student. It is important to note that all the variables combined only explained 22% of the variance in this equation.

From a practical standpoint, the difference in SAT scores reported through Score Choice based on a student's need for financial aid, although significant, is quite minimal. Those who needed financial aid tended to score approximately 10 points lower than those who did not need it, and 12 points below those who stated it was unknown whether they would need financial aid. Because this variable does not distinguish between merit-based financial aid and need-based financial aid, it is hard to determine whether these results would remain the same if this distinction was made within the variable.one was to only look specifically at need-based.

Interestingly, there was not a significant relationship between Asian ethnicity vs. White and the SAT score reported via Score Choice, or large city as high school location vs. suburban and the SAT score reported via Score Choice. Neither of these differences was statistically significant, indicating they were not related to SAT scores.

Currently, the North Carolina State University Undergraduate Admissions Office requires applicants to submit all their SAT scores. The Undergraduate Admissions Office "superscores" the SAT scores by combining students' highest math SAT score with their highest verbal SAT score to get the highest possible combined score. This practice is popular in many admissions offices; it gives students the maximum score by mixing and matching their highest verbal and highest math, regardless of whether they were taken during different test sittings. Out of interest, a paired samples *t* test was conducted to see how the current North Carolina State University practice would fair against Score Choice. The paired samples *t* test revealed a statistically reliable difference between the score a student would report using Score Choice (the highest combination of math and verbal scores during one sitting) (M = 1085, SD = 183.51) and the combination of an applicant's highest scores from each portion of the SAT to obtain the highest possible final SAT score (M = 1091, SD = 184.90), $t(151{,}900) = 186.70$, $p < .001$, $\alpha = .05$. Therefore, it appears that the practice of taking the highest math and highest verbal score across all sittings of the SAT is actually more beneficial to a student than the score submitted by Score Choice.

Implications for Practice and Recommendation

Overall, the results from this study illustrate there is a significant relationship between repeatedly taking the SAT and ethnicity, gender, need for financial aid, and high school location. These findings coupled with the availability of the new SAT reporting option, Score Choice, illuminate the need for a university or college admissions office to fully comprehend the nature of the SAT scores received during the admissions process. It is evident that the use

of Score Choice could result in an unfair advantage for certain groups of students because there is a significant increase in both math and verbal SAT scores over time.

Currently, the North Carolina State University Undergraduate Admissions Office requires applicants to submit all their SAT scores and they "superscore" by combining the highest math SAT score with the highest verbal SAT score across all sittings of the SAT to get the highest total combined score. The findings from a paired samples *t* test indicated that the practice of "superscoring" would benefit an applicant more than the score submitted by Score Choice. The results from this study imply this practice may also unfairly advantage certain groups of students, as the number of times students take the SAT impacts their overall score. However, it does give the university the opportunity to make decisions based on all SAT scores, not just the highest one. No scores are hidden, which provides North Carolina State University additional knowledge and gives the university the option to adjust admissions criteria based on the findings from this study.

Higher educational institutions should fully comprehend the impact Score Choice may have on their applicant pool. Since there is a positive relationship between SAT scores and the number of times students take the SAT, university admissions personnel must understand this prior to making the decision to allow students to hide scores from some of their SAT sittings, which would be the case if they permit Score Choice be utilized by their applicants. In order to give students the same opportunity, the institution should only compare students who have taken the SAT the same number of times. Additionally, the results indicate significant differences in SAT scores over time between the study's variables: ethnicity, gender, need for financial aid, and high school location.

Limitations

The current study attempted to determine whether there were was a significant relationship between repeatedly taking the SAT and ethnicity, gender, need for financial aid, and high school location. Furthermore, it attempted to investigate the implications the new SAT reporting option, Score Choice, could have on future North Carolina State University applicants' scores. This study is distinct since the Score Choice reporting option is relatively new, so minimal empirical research has been done on this subject. It is also unique in that it specifically looked at a large land-grant institution. However, this study is not without limitations. The first limitation is that this study looked at SAT scores over time, but the amount of time was not specified. Therefore, two people could have taken the SAT three times. One may have taken it three times within a year, while the other may have taken it three times over three years. It would add another dimension to the results if this difference was analyzed in future studies

Additionally, some of the applicants used in this study may have initially sent North Carolina State University their SAT scores but may have failed to send additional scores if they decided against attending North Carolina State University. Therefore, they may have taken the SAT more times than were reflected in this study. The assumption can also be made that applicants who initially scored well on the SAT likely did not take the SAT additional times. It should be noted that this variance was taken into account in the MLM methodology used in this study. The covariance of linear time with quadratic time for the math and verbal portion of the SAT indicated that the score a student initially received was related to the final score they achieved. This may explain why less than 4% of the students in

this study took the SAT four or more times. This fact could also be considered a limitation since a far greater number of students took the SAT between one and three times than those who took it more than three times. For future research, it would be beneficial to group students who have taken the exam four or more times into one group, instead of three separate but small groups.

Another limitation and potential area for future research is that of SAT scores for athletes. It would be interesting to see whether athletes differ from other students regarding their SAT scores and the number of times they take the test. Currently, athletes must meet certain standards to gain NCAA eligibility. The eligibility standards for freshman athletes include a sliding scale for high school GPA and SAT scores. The lower the student's high school GPA, the higher the SAT score must be in order to meet eligibility. This could impact the number of times an athlete chooses to take the SAT. This was not taken into account in this study.

The variable used in this study as a proxy for socioeconomic status, "financial aid need," has limitations, as it does not distinguish between need-based and merit-based financial aid. The question asked on the SDQ was "Do you plan to apply for financial aid at any college?" There were three potential answers to the question: yes, no, and unknown. However, it does not take into account whether students who received financial aid answered "yes" on the SDQ. It may have been more accurate to have obtained information regarding parental education level or expected family contribution, but due to the low response rates, it was determined that these would be inappropriate variables to use in the MLM models. For

purposes of this study, the financial aid variable was the best available option. However, for future research, another variable may be more optimal.

The fact that data from only one university were analyzed is also a limitation. The data used in this study are specific to North Carolina State University; therefore, the findings may not be generally applicable to higher educational institutions that are not a similar size or nature (land-grant). It would be beneficial to run a similar analysis on a variety of universities with different characteristics

Finally, this study only analyzed one portion of the admissions application, the SAT; it did not delve into other criteria that are often considered, such as students' GPA, recommendations, or extracurricular activities. The focus of this study was very specific; however, this may also be viewed as a limitation.

Future Research

There is currently a lack of empirical data surrounding the Score Choice policy recently implemented by the College Board. Due to the fact that the policy was put into operation with the Fall 2010 incoming freshman class, there were no data available to analyze when this study was initiated. However, future research will have the ability to take data directly affected by the Score Choice policy and could compare these results to see the impact of this policy more accurately.

It would also be beneficial for future studies to use data from a broader array of universities. This study specifically focuses on a large land-grant university. It would be interesting to see if the themes found in this study are true across other higher educational institutions, such as private universities or smaller colleges. Another area of interest could be

whether the findings vary over curriculums. For example, do engineering students differ from education students?

The ACT is another admissions standardized test that is used by many universities. According to the ACT website (2010), “The ACT test assesses high school students' general educational development and their ability to complete college-level work.” The ACT also has a score reporting policy similar to the SAT’s Score Choice. It would be interesting to compare both these tests in regards to the score reporting policies.

Summary

Overall, this study indicated that there was a significant relationship between repeatedly taking the SAT and ethnicity, gender, need for financial aid, and high school location. Furthermore, the new SAT reporting policy, Score Choice, could, indeed, impact North Carolina State University undergraduate applicants’ scores. On average, there was a change in student SAT scores over time when looking at the North Carolina State University applicant pool between the years 2003 and 2008. Additionally, there was not a disadvantage to taking the SAT multiple times, as student average scores did not decrease across the selected population. This indicated that the more times students took the SAT, the greater the potential to increase their score. The findings also indicated that the differences in the point at which scores decreased with each additional attempt of the SAT were ethnic, gender, and need for financial aid. The SAT score that would be reported through Score Choice was found to be related to ethnicity, gender, financial aid, and location of a student’s high school.

The above summary of results implies that Score Choice will enable inequity during the admissions process. However, the current practice of combining a student's highest score on each portion of the SAT used by many colleges and universities, thereby resulting in the greatest overall score, may actually create greater inequity than the Score Choice option. Going into this study, more discrepancy was expected regarding socioeconomic status and SAT scores. It was assumed that those who had more money would take the SAT many more times than their poorer counterpart. With this assumption, it was thought that Score Choice would make a large impact on the applicant pool. However, this study showed that the majority of students take the SAT twice. This implies that the fee waivers offered to financially needy students to pay for the cost of taking the SAT twice might be enough.

Upon further analysis of the data, it is also evident that the results from this study would likely vary significantly if viewed at a different type of higher educational institution. North Carolina State University is a large public university. One could argue that the type of student that attends North Carolina State University would most likely vary from a small liberal arts college, for example. Therefore, the number of times a student takes the SAT at different schools could vary, consequently altering the impact of Score Choice.

Based on this study, it is evident that Score Choice could potentially skew admissions officers' decisions regarding an applicant due to the lack of scores provided to the college or university if a student chooses this option. Because of this potential, it would appear that the most equitable decision for a college or university would be to require all student SAT scores during the admissions process. Furthermore, this would also be the most advantageous to the student if the college or university utilizes the practice of "superscoring" SAT scores.

References

ACT Inc. (2010). Retrieved from About ACT website: http://www.act.org/aboutact/

Alderman, D. L. & Powers, D. E. (1983). Effects of test familiarization on SAT performance. *Journal of Educational Measurement, 20*(1), 71-79.

Amrein, A. L., & Berliner, D. C. (2010, March 28). High-stakes testing & student learning. Retrieved from Education Policy Analysis Archives website: http://epaa.asu.edu.

Amrein, A. T. & Berliner, D. C. (2003). The effects of high-stakes testing on student motivation and learning. *Educational Leadership, 60*(5), 32-38.

Antonio, A. L. (2003). Diverse student bodies, diverse faculties. *Academe, 89*(6), 14-17.

Bhargava, A., Frankenberg, E., & Le, C. Q. (2008). Still looking to the future: Voluntary K-12 school integration. Retrieved from http://www.naacpldf.org/

Black Issues in Higher Education (2002). Florida court rejects NAACP challenge to university admissions. *Black Issues in Higher Education, 19*(4), 10.

Blair, J. (2000). Plan to ban race in admissions to Fla. colleges clears regents. *Education Week*, *19*(24), 21.

Boger, J. C., & Orfield, G. (Eds.) (2005). *School resegregation: Must the South turn back?* Chapel Hill, NC: The University of North Carolina Press.

Boldt, R. F., Centra, J. A., & Courtney, R. G. (1986, January 1). The validity of various methods of treating multiple SAT scores. Retrieved from http://professionals.collegeboard.com/

Borek, J. (2008). A nation at risk at 25. *Phi Delta Kappan*, *89*(8), 572-574.

Borman, G. D., Stringfield, S. C., & Slavin, R. E. (2001). *Title I: Compensatory education at the crossroads.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Bowen, W. G., & Bok, D. (1998). *The shape of the river.* Princeton, NJ: Princeton University Press.

Bowman, D. H. (2002). College Board to drop policy letting students select best test scores. *Education Week*, *21*(25), 14.

Bracey, G. W. (2001). Test scores in the long run. *Phi Delta Kappan*, *82*(8), 637-638.

Brannon, P. M., Morgan Dean, R. Q., & Morgan Dean, J. C. (2002). Our land grant mission in the twenty-first century. *Human Ecology*, *30*(1), 1.

Brigham, C. C. (1923). *A study of American intelligence.* Princeton, NJ: Princeton University Press.

Bunin, L. (2009, March 13). Score choice helps everyone. *USA Today*. Retrieved from http://www.lib.ncsu.edu

Burke, E. F. (1997). A short note on the persistence of retest effects on aptitude scores. *Journal of Occupational & Organizational Psychology*, *70*(3), 295-301.

Cabrera, A. F. & La Nasa, S. M. (2001). On the path to college: Three critical tasks facing America's disadvantaged. *Research in Higher Education*, *42*(2), 119-149.

Camara, W. J. (2006, September 7). Letter from College Board's Vice President of Research. Retrieved from http://www.collegeboard.com

Carnevale, A. P., & Rose, S. J. (2003, March). Socioeconomic status, race/ethnicity, and selective college admissions. A century foundation paper. Retrieved from http://www.tcf.org.

Carter, S. L. (1992). *Reflections of an affirmative action baby.* New York, NY : Perseus Books.

Cassady, J. C. (2001, January 24). Self-reported gpa and sat: a methodological note. Retrieved from http://PAREonline.net

Chenoweth, K. (1998). The College Board decries preparation gap. *Black Issues in Higher Education*, *15*(15), 24.

Clarke, M., Haney, W., & Madaus, G. (2000, January). High Stakes Testing and High School Completion. Retrieved from National Board on Educational Testing and Public Policy, Boston, MA. Web site: http://www.bc.edu

Coleman, J. S. (1966). *Equality of educational opportunity.* Retrieved from http://dx.doi.org/

College Board. (2011). About the SAT. Retrieved from http://professionals.collegeboard.com

College Board. (2005). For colleges: Internet score delivery and disk record layout/description for SAT reasoning test and SAT subject tests score reports. Retrieved from http://www.collegeboard.com.

College Board. (2008, August 26). SAT scores stable as record numbers take test. Retrieved from http://professionals.collegeboard.com/

College Board. (2008). Retaking the SAT. Retrieved from http://professionals.collegeboard.com

College Board. (2009a). How financial aid works. Retrieved from http://www.collegeboard.com

College Board. (2009b). New SAT Score-reporting Feature Allows Students to Choose the Scores They Send to Colleges, Universities and Scholarship Programs. Retrieved from http://www.collegeboard.com

College Board. (2009c). SAT program fee-waiver service. Retrieved July 10, 2009, from http://www.collegeboard.com

College Board. (2009d). SAT scoring before March 2005. Retrieved from http://www.collegeboard.com

College Entrance Examination Board (1999). Concordance Between SAT I and ACT Scores for Individual Students. Retrieved from http://professionals.collegeboard.com

Coleman, J. S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., & York, R. (1966). *Equality of educational opportunity.* Washington D.C.: U.S. Government Printing Office.

Cornell University Law School. (2009a). Regents of the University of California v. Bakke (No. 7811). Retrieved from http://www.law.cornell.edu

Cornell University Law School. (2009b). Gratz v. Bollinger (02-516). Retrieved October from http://www4.law.cornell.edu.

Crouse, J., & Trusheim, D. (1988). *The case against the SAT*. Chicago, IL: University of Chicago Press.

Desmond, C. & Goldman, E. (2008). Talking about a revolution: Looking to the past to save our future. *New England Journal of Higher Education*, *22*(5), 18-19.

Economic Research Service. (2010). State fact sheets: North Carolina. Retrieved from http://www.ers.usda.gov

Edwards, H. T. (2004). The journey from Brown v. Board of Education to Grutter v. Bollinger: From racial assimilation to diversity. *Michigan Law Review*, *102*(5), 944-978.

Eisenhower, D. D. (1961). President Eisenhower's State of the Union message. *President Eisenhower's State of the Union Message*, 1. Retrieved from The American Presidency Project website: http://www.presidency.ucsb.edu

Epstein, J. P. (2009). Behind the SAT-optional movement: Context and controversy. *Journal of College Admission*, Summer (204), 8-19.

Erickson, J. (2007). Women make progress under Title IX, but barriers persist. *National NOW Times*, *39*(2), 6-7.

Everson, H. T. & Millsap, R. E. (2004). Beyond individual differences: Exploring school effects on SAT scores. *Educational Psychologist*, *39*(3), 157-172.

Executive Summaries of the Reports of the Kellogg Commission on the Future of State and Land-Grant

Fairleigh Dickinson University. (2001). On the importance of diversity in higher education. Retrieved from http://www.fdu.edu

FairTest. (2007, August 20). Gender bias in college admissions tests. Retrieved from http://www.fairtest.org.

Farley, K. (2009, January 14). Colleges can request all SAT scores. *The Dartmouth.* Retrieved from http://thedartmouth.com

Fischer, K., & Hebel, S. (2006). The geographic have-nots: Urban centers and rural regions. *Chronicle of Higher Education*, 53(11), A20.

Fleming, J. & Garcia, N. (1998). Are standardized tests fair to African Americans?: Predictive validity of the SAT in black and white institutions. *The Journal of Higher Education, 69*(5), 471-495.

Foderaro, L. (2009, March 1). Well-regarded public colleges get a surge of bargain hunters. *The New York Times*. Retrieved from http://www.nytimes.com

Foster, B. (2003). Profiling female teachers of agricultural education at the secondary level. *Journal of Career and Technical Education, 19.* Retrieved from http://scholar.lib.vt.edu

Frederiksen, N. (1994). *The influence of minimum competency tests on teaching and learning.* Princeton, NJ: Educational Testing Service, Policy Information Center.

Freeberg, N. E. (1988). Analysis of the revised student descriptive questionnaire, phase I: Accuracy of student-reported information. Retrieved from http://professionals.collegeboard.com/

Freedle, R. O. (2008). Correcting the SAT's ethnic and social-class bias: A method for reestimating SAT scores. *Harvard Educational Review, 73*(1), 1-43.

Frey, C. (2006). UW looking beyond test scores, GPA. Retrieved from http://seattlepi.nwsource.com

Geraghty, M., & Guernsey, L. (1996, September 6). Average scores on sat and act examinations increase slightly trends in sat scores. Retrieved from The Chronicle of Higher Education website: http://chronicle.com

Gleason, B. (2010). World-class greatness at a land-grant university near you? *Chronicle of Higher Education, 57*(6), A88.

Goral, T. (2008). The SAT is dead; Long live the SAT. *University Business, 11*(8), 10.

Gose, B. & Selingo, J. (2001). The SAT's greatest test. *Chronicle of Higher Education, 48*(9), A10.

Graves, L. (2008, June 24). Kids can pick which SAT scores a college sees: The new policy very likely will be popular with students, but what do admissions officers think? *U.S. News.* Retrieved from http://www.usnews.com

Greene, H. & Greene, M. (2004). The widening gender gap. *University Business, 7*(9), 27-29.

Greenleaf, W. J. (1934). The land grant colleges. *Congressional Digest, 13*(8/9), 198.

Hamilton, L. S., Stecher, B. M., & Klein, S. P. (Eds.) (2002). *Making sense of test-based accountability in education.* Santa Monica, CA: RAND Corporation.

Hatch, N. O. (2008, June 29). A better measure than the SAT. *Washington Post*, B07.

Hawkins, D. A. (2004). The state of college admission 2003-2004. Retrieved from http://www.admissionpossible.org

Heller, D. E. (2002). *Conditions of access: Higher education for lower income students.* Westport, CT: Praeger Publishers.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in american life.* New York: Free Press.

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

Hoff, D. J. (1999). Echoes of the Coleman Report. *Education Week*, *18*(28), 33-35.

Hoover, E. (2008). At admissions conference, talk of standardized tests dominates. *The Chronicle of Higher Education*, *55*(7), A26.

Hoover, E. (2009). 'Score Choice': A tempest in a teapot? *Chronicle of Higher Education*, *55*(20), A4.

Innerst, C. (1996, August 23). SAT scores rise, but critics call tests 'ruined': College Board defends quality. *The Washington Times*. Retrieved from http://www.lexisnexis.com/hottopics

Jagacinski, C. M. (1983, April). Engineering Careers: Women in a male-dominated field. Retrieved from http://www.eric.ed.gov

Jaschik, S. (2009, March 26). The impact of dropping the SAT. Retrieved from http://www.insidehighered.com

Jensen, L. (2006). New Immigrant Settlements in Rural America: Problems, Prospects, and Policies. Retrieved from http://www.carseyinstitute.unh.edu/

Kaplan, D. A., Henig, S., & Philips, M. (2008). The new SAT score policy: Tiny loophole, big shock? *Newsweek*, *152*(24), 12.

Koretz, D. (2008). *Measuring up: What educational testing really tells us.* Cambridge, MA: Harvard University Press.

Ladenson, R. F. (1975). Rawls' principle of equal liberty. *Philosophical Studies*, *28*(1), 49-54.

Lawlor, S., Richman, S., & Richman, C. L. (1997). The validity of using the SAT as a criterion for black and white students' admission to college. *College Student Journal*, *31*(4), 507-515.

Lemann, N. (2000). *The big test: The secrets of the American meritocracy.* New York, NY: Farrar, Straus and Giroux.

Marchant, G. J. (2004). What is at stake with high stakes testing? A discussion of issues and research. *The Ohio Journal of Science*, *104*(2), 2-7.

Marchant, G. J. & Paulson, S. E. (2001). State comparisons of SAT scores: Who's your test taker? *NASSP Bulletin, 85,* 62-73.

Manzo, K. K. (2008). Panel urges reduced use of college-admission exams. *Education Week, 28*(6), 13.

Mau, W. & Lynn, R. (2001). Gender differences on the Scholastic Aptitude Test, the American College Test and college grades. *Educational Psychology*, *21*(2), 133-136.

McDowell, G R. (2001). *Land-grant universities and extension into the 21st century.* Ames, Iowa: Iowa State University Press.

McGaghie, W. C. & Kreiter, C. D. (2005). Holistic versus actuarial student selection. *Teaching & Learning in Medicine*, *17*(1), 89-91.

Miners, Z. (2008). Students allowed to choose top SAT scores. *District Administration, 55*(9), 16.

Morgan, G. A. (2004). *SPSS for introductory statistics: Use and interpretation.* Hove, East Sussex: Psychology Press.

National Association of State Universities and Land-Grant Colleges. (1995). *The land-grant tradition.* Washington, DC: NASULGC.

National Commission on Excellence in Education. (1983, April). A Nation at Risk: The Imperative for Educational Reform. Retrieved from http://www.ed.gov/pubs

Nealy, M. J. (2008). Admission counseling commission endorses optional SAT/ACT admissions requirements. *Diverse: Issues in Higher Education*, *25*(18), 12-13.

Nichols, S. L., & Berliner, D. C. (2007). *Collateral damage: How high-stakes testing corrupts America's schools.* Cambridge, MA: Harvard Education Publishing Group.

Nichols, S. L. & Berliner, D. C. (2008). Why has high-stakes testing so easily slipped into contemporary American life? *Education Digest, 74*(4), 41-47.

North Carolina State University. (2009). Centennial campus. Retrieved from North Carolina State University website: http://centennial.ncsu.edu/

North Carolina State University. (2011). Discovery begins at North Carolina State University. Retrieved from North Carolina State University website: http://www.ncsu.edu

North Carolina State University. (2009). Extension, engagement & economic development. Retrieved from North Carolina State University website: http://www.ncsu.edu/extension

North Carolina State University. (2010). Scholarships and Other Opportunities. Retrieved from North Carolina State University website: http://admissions.ncsu.edu

North Carolina State University. (2009). University Planning and Analysis: Enrollment Data. Retrieved from North Carolina State University website: http://www2.acs.ncsu.edu/

North Carolina State University. (2011). University Planning and Analysis: First Time Freshman Profile. Retrieved from North Carolina State University website: http://www2.acs.ncsu.edu/UPA/admissions/freshman_profile.htm.

Obama, B. (2009). Remarks by the President at the National Academy of Sciences Annual Meeting. Retrieved from: http://www.whitehouse.gov

Office of Technology Assessment. (1992). *Testing in American schools: Asking the right questions.* Washington, DC: U.S. Government Printing Office.

Quaid, L. (2008). Girls are equal to boys on state math tests, study finds. *Education Week, 27*(44), 12.

Rampell, C. (2009, August 27). Sat scores and family income. *The New York Times*. Retrieved from http://www.nytimes.com/

Rank, M. R. (2005). *One nation, underprivileged.* New York, NY: Oxford University Press.

Rapp, C. (2000). Bushism in One Florida. *National Review, 52*(4), 24-26.

Raudenbush, S. W., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods.* Thousand Oaks, CA: Sage Publications.

Ravitch, D. (2000). *Left back: A century of battles over school reform.* New York, NY: Simon and Schuster.

Rawls, J. (1999). *A theory of justice: Revised edition.* Cambridge, MA: Harvard University Press.

Rigol, G. W. (2003). Admissions decisions-making models: How U.S. institutions of higher education select undergraduate students. Retrieved from http://professionals.collegeboard.com

Rimer, S. & Lewin, T. (2008). SAT changes policy, opening rift with colleges. *New York Times,* 14.

Rooney, C. & Schaeffer, B. (1998, September). Test Scores Do Not Equal Merit: Enhancing Equity & Excellence in College Admissions by Deemphasizing SAT and ACT Results. Retrieved from http://www.eric.ed.gov

Sackett, P. R., Borneman, M. J., & Connelly, B. S. (2008). High stakes testing in higher education and employment: Appraising the evidence for validity and fairness. *American Psychologist, 63*(4), 215-227.

Sacks, P. (1999). *Standardized minds: The high price of America's testing culture and what we can do to change it.* Cambridge, MA: Perseaus Books.

Schmidt, P. (2007). *Color and money.* New York, NY: Palgrave MacMillan.

Schmidt, P. (2010). Sandra Day O'Connor revisits and revives affirmative-action controversy. *The Chronicle of Higher Education*. Retrieved from http://chronicle.com/article/Sandra-Day-OConnor-Revisit/63523/.

Schreiber, J. B. & Griffin, B. W. (2004). Review of multilevel modeling and multilevel studies in the Journal of Educational Research (1992-2002). *Journal of Educational Research, 98*(1), 24-33.

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational & Behavioral Statistics, 23*(4), 323-355.

Soares, J. (2007). *The Power of Privilege: Yale and America's Elite Colleges.* Palo Alto, CA : Stanford University Press.

Stanford University (2011, November 11). Freshman requirements & process. Retrieved from Stanford University Office of Undergraduate Admission website: http://admission.stanford.edu

StatSoft Inc. (2008). Multiple regression. Retrieved from http://www.statsoft.com

Straus, V. (2010, June 8). Do colleges superscore act and sat equally? Retrieved from The Washington Post website: http://voices.washingtonpost.com/answer-sheet/sat-and-act/superscoring-act-vs-sat.html

Stuart, A. W., & Baum, L. (2005). Contemporary migration in North Carolina. Retrieved from http://www.ncmuseumofhistory.org

Teare, C. (2009). The Russian roulette of SAT scores. *The Chronicle of Higher Education, 55*(36), A27.

Thacker, L. (2005). *College unranked: Ending the college admissions frenzy.* Cambridge, MA: Harvard University Press.

The University of North Carolina. (2009). The University of North Carolina a multi-campus university: Facts and figures. Retrieved from http://www.northcarolina.edu

Thomas, J. Y., & Brady, K. P. (2005). The elementary and secondary education act at 40: Equity, accountability, and the evolving federal role in public education. *Review of Research in Education*, *29*(3), 51-67.

Towers, J. M. (1992). Twenty-five years after the Coleman Report: What should we have learned? *Clearing House*, *65*(3), 138-140.

University of Texas at Austin. (2008). About Hopwood. Retrieved from http://tarlton.law.utexas.edu/hopwood/

University of Missouri-Kansas City School of Law. (2009). Regents of the University of California v. Bakke. Retrieved from http://www.law.umkc.edu

USA Today. (2009, March 13) Defining the SAT downward. *USA Today*, 8a.

U.S. Census Bureau. (2011, October 13). Retrieved from State and county quickfacts: North Carolina website: http://quickfacts.census.gov

U.S. Department of Education, (2004). *Achieving Diversity: Race-neutral alternatives in American education.* Washington, DC: U.S. Department of Education, Office for Civil Rights.

U.S. Department of Education. (2011). Degrees earned. Retrieved from National Center for Education Statistics website: http://nces.ed.gov

U.S. Department Of Education. (2012, February 12). Expected family contribution (EFC). website: http://www.fafsa.ed.gov

U.S. Department of Education. (2009). The federal role in education. Retrieved from http://www.ed.gov

U.S. Department of Labor (2009). Title IX, Education Amendments of 1972. Retrieved from http://www.dol.gov

U.S. Environmental Protection Agency (2009, September 10). Demographics. Retrieved from http://www.epa.gov

Vigdor, J. L. (2003). Retaking the SAT. *The Journal of Human Resources*, *38*(1), 1-33.

Webster, T. J. (2001). A principal component analysis of the U.S. News & World Report tier rankings of colleges and universities. *Economics of Education Review*, *20*(3), 235-244.

Western Interstate Commission for Higher Education. (2008). Knocking at the college door: Projections of high school graduates by state and race/ethnicity 1992-2022. Retrieved from http://www.wiche.edu/

West-Faulcon, K. (2009). The river runs dry: When Title VI trumps state anti-affirmative action laws. *University of Pennsylvania Law Review*, *157*(4), 1075-1160.

Wightman, L. F. (1997). An empirical analysis of the consequences of abandoning race as a factor in law school admissions decisions. *Forest Ecology and Management*, *72*(1), 1-53.

Williams, J. (2007, June 29). Don't mourn brown v. board of education. *The New York Times*. Retrieved from http://www.nytimes.com

Winfield, L. F. (1990). School competency testing reforms and student achievement: Exploring a national perspective. *Educational Evaluation and Policy Analysis, 12*, 157-73.

Young, J. W. (1991). Gender bias in predicting college academic performance: A new approach using item response theory. *Journal of Educational Measurement*, *28*(1), 37-47.

Young, J. W. & Fisler, J. L. (2000). Sex differences on the SAT: An analysis of demographic and educational variables. *Research in Higher Education*, *41*(3), 401-416.