How We Justify and Maintain the White, Male Academic Status Quo through the Use of Biased College Admissions Requirements

By

Theodore Micceri, Ph.D.
Coordinator of Statistical Research
Office of Planning and Analysis

Office of the Provost
University of South Florida

Tampa, FL 33620

February, 2007

Table of Contents

Executive Summary

Prompted by the recent growth in the proportions of minority entrants from Community Colleges (CCTs) into the SUS, and an earlier scan which suggested consistent bias against females and minorities on entrance examinations, the current study analyzed over one million SUS applicants to determine whether any biases in entrance test scores tended to favor either whites over minorities or males over females.

Background

Today, few question the economic value of higher education to individuals. Day & Newburger (2002, p. 2) show that individuals holding an associate's and a bachelor's degree average respectively 26% and 72% more annual income than those with a high school diploma. Therefore, anything that interferes with an individual's opportunity to attain a degree costs both individuals and society (in the form of tax dollars and well documented societal benefits, Wellman, 2002). Between 1996 and 2003, the Florida SUS experienced a 33% increase in new undergraduate matriculants (70% among FTICs) during a time when state funding in absolute dollars increased by 39%, but in constant HEPI dollars grew by only 11%. Faced with rising student numbers and fewer real dollars to serve each student, SUS schools appear to have chosen to reduce access more than service. Therefore, entrance requirements became more stringent as institutions attempted to reduce the size of their matriculating FTIC cohorts (Vogel, 2006). Usually, in higher education, more stringent entrance requirements translate to higher entrance test score requirements. Supporting this, during the time in question, minority native (FTIC) matriculants to SUS institutions, despite rapid growth among Florida high school graduates, remained flat (moving from 36.4% to 38.0%), while their entrance rates as CCTs increased by 30% (from 27.8% to 37.2%). Female representation among CCTs also grew more than 2.5 times as much as it did among FTIC students. Florida community colleges have an open door policy for Florida high school graduates; thus, tests play no part in admissions.

A substantial literature documents problems for the above groups on standardized tests such as the ACT and the SAT (Micceri, 2001, lists and reports on a large number of these). Such measures were initially championed by the American Eugenics Society to "purify" the race of low-grade and degenerate groups such as minorities and the poor (Gibson, 2001). Despite thousands of studies conducted by higher education institutions and many others supported by testing services, it is almost impossible to find relationships between entrance examinations and any outcomes other than first semester grades, which is the only outcome to which test makers themselves claim their product relates (Fairtest, 2006; Elert, 1992). First semester grades fail to provide any indication of long-term retention or graduation (Micceri, 2001). The current study sought to address the research question in a more through fashion, and from a more valid research perspective than most historical work on the topic.

Methods

This study addressed the following research question: "Do any consistent score differences occur on Standardized Tests between different sexes or race/ethnicities who exhibit the same historic academic performance as measured by high school GPA?"

Using data from the Florida SUS Admissions files for all First Time In College (FTIC) applicants to all eleven SUS institutions for the Academic Years 1997-98 through 2005-06, mean test scores (SAT, Quantitative & Verbal Subscores, ACT) were computed for each subgroup at the 1/10th of a point high school GPA level (e.g. 3.6 is one group, 3.7 another). If tests measure what they purport to measure, academic performance, one would expect only random deviations to occur between groups who exhibit almost identical academic performance as measured by GPA, a reliable, multi-year, multi-source estimate of academic performance. Consistent deviations from the expected random must therefore reflect bias. Subgroups included females and males within each major racial/ethnic subgroup: African American, Asian, Hispanic, white and Other.

Results and Discussion

Among the 1.1 million student scores submitted to analysis at the same high school GPA, where one would expect group means of large groups (smallest being over 500) to cancel out and produce random effects (50% higher, 50% lower), whites consistently outscored minorities by an average difference of over 60 points on the SAT.[1] Whites outscored every minority group. An even greater discrepancy occurred between males and females, with males holding an average advantage over females of 75 points on the SAT at the same high school GPA. Considering that ETS (2001) claims a 60 point difference shows a "...true difference in ability," and since females outperform males on every academic performance measure one can choose (graduation, grades, etc.), and in every discipline, even Engineering (Micceri, 2005, p. 10), one must ask: "A true difference on what?" Apparently, a "true difference" in the ability to score high on multiple-choice, standardized tests.

The consistent advantage that whites and males have on such tests suggests a discriminatory bias that ultimately translates into economic, social and status advantages for whites and males over minorities and females. This also assures that faculty in universities remain primarily white and male, because most university faculty earn their doctorates at flagship universities, and elite doctoral programs tend to prefer applicants from other elite programs, who tend to use test scores as entrance requirements. The effects of the discriminatory bias is well documented, most recently by Haycock & & Gerald (2006, p. 3), who state about public flagship institutions: “Even as the number of low-income and minority high school graduates in their states grows, often by leaps and bounds, these institutions are becoming disproportionately whiter and richer."

[1] Because some 65% of SUS applicants take the SAT, discussion is limited to the SAT total score for simplicity's sake. Similar effects occurred for all scores and subscores in all analyses.

How We Maintain the Academic Status Quo Through the Use of Biased Admissions Requirements

This paper attempts to show how an trend that began during World War II (WWII) helps ensure that society's upper class of wealth and power gains advantages when it comes to enrolling in Florida's public State University System (SUS) institutions. Higher Education has historically been, at least until the proletariat revolutions of the 20th Century, the exclusive territory of the aristocracy and wealthy. A place they could send their young to get away from the dangers of interacting with city workers[2] and thereby gain traditional cultural knowledge/capital. For example, between 1890 and 1900, fewer than 5% of Americans aged 18 to 21 years attended higher education institutions (Goldin & Katz, 1999, p. 41). By 1970, this percentage had risen to 70%, meaning that large numbers of "commoners" were rubbing elbows with the elite, partly because of the 1965 Higher Education Act which stated that colleges couldn't turn away applicants merely because their families were poor. Such interaction with society's riffraff had traditionally been avoided among the aristocracy through the use of exclusive (expensive) private colleges. Theoretically, public universities, which are paid for by the taxes of all, should offer even those of the lower classes the opportunity for higher education. However, in the latter part of the 20th Century and the early years of the 21st Century, as the United States experiences increasing wealth disparity (Sahadi, 2006; Witte & Henderson, 2004), we see a disturbing tendency to exclude “commoners”. Haycock & Gerald (2006, p. 3) state, regarding public flagship institutions: “Even as the number of low-income and minority high school graduates in their states grows, often by leaps and bounds, these institutions are becoming disproportionately whiter and richer." The primary tool used by higher education to discriminate against commoners is a selection bias that results from the use of stringent admissions requirements (read higher test score requirements). In the Florida State University System (SUS) the “more stringent admissions” arguably resulted from a 70% increase in First Time in College (FTIC) matriculations during a time when constant dollar funding only increased by 11%[3] (1996 to 2003). Although standardized tests were initially pushed to the fore by Connant as an Admissions requirement in an attempt to reduce class bias, in reality, they perpetuate the class bias effect, as this study will show. Stimulated by earlier research conducted in an attempt to understand the recent movement of underprivileged minorities from direct entry into the Florida SUS to transfer from community colleges, in addition to Gibson’s (2001, p 1.) claim that: “...The SAT measures, above all else, class, sex, and race...”, this study addresses the research question: Do any consistent score differences occur on Standardized Tests between different sexes or race/ethnicities for students exhibiting the same historic academic performance levels as measured by High School Grade Point Average (HSGPA)?

[2] Thus, most major state public universities locate in small, comparatively affluent towns like Albany, Austin, and Athens rather than New York City, Dallas or Atlanta.

[3] Using the Higher Education Price Index (HEPI), to provide a more realistic adjustment estimate than that of the Consumer Price Index (CPI).

Background

Between 1996 and 2003, the Florida Public State University System (SUS) experienced a 70% increase in new First Time In College (FTIC) students attending one or more of the institutions, and a total increase from 54,000 to 72,000 (33.3%) among new undergraduates annually. During this same period of time, the total amount of money allocated to higher education by the State to higher education in Florida increased by 39.3%, which appears adequate. However, if one adjusts for inflation using HEPI, this represents only an 11.2% increase (SHEEO, 1997; Palmer, 2004; Commonfund, 2006). Thus, Florida SUS institutions were faced with the choice of either reducing service or reducing access and they apparently chose the latter. As a result, it has become increasing difficult for high school students to gain admission to an SUS institution (Vogel, 2006). During this time, minority enrollment directly from high school into the SUS 4-year institutions remained relatively flat, growing by 1.6 percentage points (36.4% to 38.0%), despite their increasing percentages in the overall Florida population and among high school graduates (CDC, 2006). At the same time, their entry to the SUS as Community College Transfers (CCTs) increased by 9.4 percentage points (27.8% to 37.2%). While African American representation among FTIC students remained the same between 1996 and 2003 (18.3% and 18.4%), they showed a 30 percent representation increase among CCTs, from 9.8% to 12.7%. Students classified as "Other" more than doubled in CCT representation, from 2.9% to 6.2%, while Hispanic students increased from 15.1% to 18.3%. Females remained more stable in both populations, however, they too showed a 2.5 times greater increase among CCT students (2.4%) than FTIC students (0.9%).

This increase in minority and female percentages among CCT transfers prompted an investigation of possible bias in admissions that might adversely influence minorities and females entry to SUS institutions. A quick analysis of 164,378 SUS fall FTIC applicants (2001 to 2003) comparing mean SAT scores between whites and minorities and between sexes at each 10th of a HSGPA point across racial/ethnic groups and sex[4] showed a consistent bias favoring whites and males. In 181 of 189 race/ethnicity comparisons whites scored higher than the comparison group. At the same GPA point, for example, 3.6 or 3.7, the overall mean difference between whites and minorities averaged 62 points across years. There was an even more extreme bias favoring males over females than whites over minorities. Males exhibited higher mean SAT scores than females in all 63 comparisons with an average advantage of 75 points across years (Borman, Workman, Miller & Micceri, 2006). On the topic of differences between individuals ETS (2001) states: "The user can be reasonably confident that a score difference of around 60 points or more indicates a true difference in ability between two test takers." This should be more true of group means than individual scores. Therefore, if a "true difference" occurs between males and females exists, one must ask: what "true difference," because females outscore males in performance everywhere in education. Thus, it can't be academic performance. This background and findings prompted the current, broader and more intensive study.

[4] In this paper, the correct term sex will be used in place of the common, but erroneous term gender. Gender is a strictly linguistic term having three groups in English: masculine, feminine and neuter. People are male or female, not neuter.

The History of Standardized Tests as Admissions Criteria

First, it is useful to understand how standardized tests came to replace the traditional essay as a primary selection tool for college admissions. Among the first to push for the use of IQ tests as criteria for officer selection in the military were Terman and Yerkes, executives in the American Eugenics Society (AEC), who wanted to purify the race of low-grade and degenerate groups such as minorities and the poor (Gibson, 2001). From an opposite perspective, following the research of the 1930s and 1940s which showed genius appearing among all classes, James Bryant Conant, at the time, Harvard's president, believed that in the half century leading up to 1940, the U.S. had gone from being a classless society to one that was falling under the control of a hereditary aristocracy. Conant hoped to use the SAT to select students for college who had virtue and talent, assuming that the two went hand-in-hand. Regarding the SAT he wrote: "...we have before us a new type of social instrument whose proper use may be the means of salvation of the classlessness of the nation..." Conant hoped to use this new social instrument (the SAT) to sort and slot the entire populace on the basis of their test-defined intelligence in the name of creating a perfected, classless and democratic America modeled on Plato's Republic.

Wars provide opportunity for restructuring society and Conant, realizing this, moved quickly to establish the current testing regime after WW II began. Just after Pearl Harbor, he replaced the old essay tests for college admission with the SAT for all applicants at Harvard. In 1943, a revised SAT was administered to more than 300,000 people nationwide for officer-selection purposes. Immediately after the war, Conant, through an adept series of bureaucratic maneuvers, arranged for all the leading education tests and testing organizations in the country to be merged into a new, private, non-profit entity that would effectively hold a monopoly in the field, the Educational Testing Service (ETS). Because the service is private, it does not answer to the populace, as do government-run testing organizations in most countries. Because these aptitude tests were free to colleges, represented far less work on the colleges' part than traditional essays, and were thought to be objective and valid, they spread rapidly as a college selection criterion, aided by ETS' nationwide marketing. Conant's purpose was to create an objective measure that could identify his aristocracy of virtue and talent even when located in poorer communities (Lemann, 1999a, p. 53). Unfortunately, these tests have not proven objective and precisely the opposite of what Conant hoped has turned out to be true due to a consistent bias favoring affluence. As renowned demographer Harold Hodgkinson (Hodgkinson, 1999) states: "SATs predict one thing beautifully, but it's not the grades students will earn as freshmen; it's the household income of the test-takers. For every $10,000 increase in household income, math and verbal scores go up a minimum of nine points." One might ask how this could occur for measures that purport to be objective. One primary factor is a set of biases in the tests that consistently favor the affluent over lower-income individuals.

Biases Common to All Forms of Standardized Tests

Although many mistakenly think them objective, standardized tests exhibit numerous types of bias against specific subpopulations of test-takers. As a result, tests consistently underestimate specific groups' knowledge or performance. In measurement, biases arise in some curious ways. For example, when researching his non-language IQ tests, Raven (1939) discovered that the physical location of correct and erroneous answers

consistently influenced (biased) student errors. The following are a few of the most obvious and common types of bias inherent to standardized tests.

1. Students from households using non-standard English face a consistent disadvantage. No matter what their racial/ethnic group, poorer people in the United States tend to grow up in such environments. As an example of how such biases can lead to totally erroneous conclusions, early in the 20th century, when widespread testing began, immigrants from Poland and Italy were considered stupid and inferior because they scored poorly on tests in their second language, English.
2. Regarding cultural biases, test creators tend to grow up in middle-class to upper class American households. The test items such individuals write tend to reflect their personal and cultural experiences. Those from poorer households, different countries, cultures or religions frequently lack experience with the topics discussed on tests. This type of bias first became apparent in the 1890's when researchers suggested the use of field trips to the country for city public school kids because city kids scored lower than country kids on standardized tests. Most of the test makers at that time had grown up on farms, and city kids knew little about animals or farming environments, which were common topics on tests created by former farm kids. As an Australian once said: "If an average American were to take a test of required skills that was created by a Bushman, they would test at the moron level."
3. The great majority of high-stakes and standardized tests are speeded. This tends to work against those who are perfectionists, those who process information slowly, and against those who are less likely to "take risks" (probably more generally true of females than males). Many people will not put down an answer until they are positive it is correct (perfectionists). Obviously, in a speeded test, such individuals have trouble answering enough questions to attain a high score. Additionally, some process information more slowly than others. In most life situations other than driving race cars, this is not very relevant. However, when taking a timed test, it negatively biases estimates of knowledge, intellect and talent.
4. Large numbers of students today suffer from test anxiety. This begins to show in about the third grade and may cause a student to perform poorly on tests. Prior performance being an influence in test anxiety, students who are subject to items #1, #2, #3 or #6 will tend to perform less well and therefore be subject to greater test anxiety, which may reduce their scores.
5. Standardized tests and almost all tests of any type tend to heavily reward short-term memory skills. Although this is a useful skill to have, it is far from the only skill needed when attempting to solve problems either in college, or in the real world. Almost all tests are biased against those who lack strong short-term memory skills no matter at what level their long-term memory skills may be.
6. A fairly large proportion of the population has what is termed a learning disability by the educational community. This may represent any of a number of different ways of looking at the world, but in most cases, these different perspectives or methods of processing information, associate with poorer performance in school and almost always with erratic test performance, itself an indicator of a learning disability.
7. As a rule, tests reflect a very limited perspective on a single type of intellectual process that may be termed logical/analytical, and which rarely requires any higher-order thinking by a respondent. Performance in school is also typically evaluated

using tests of similar level, although well-trained teachers use many other sources of information when assigning grades. Unfortunately, many are not well trained in measurement. As a result, it is not uncommon for those who think differently, and who may be the greatest geniuses among us, to not exhibit this specific type of intelligence as measured by either a teacher's instruments in schools or standardized tests (This bias frequently relates to #6, the learning ability/disability issue.). Robbins (1987) notes some well-known examples from history:

> Albert Einstein's parents were sure he was retarded because he spoke haltingly until the age of nine and even after that would respond to questions only after a long period of deliberation. He performed so badly in his high school courses, except mathematics that a teacher told him to drop out, saying, "You will never amount to anything Einstein." Charles Darwin did so poorly in school that his father told him, "You will be a disgrace to yourself and all your family." Thomas Edison was called "dunce" by his father, "addled" by his high school teacher and was told by his headmaster that he "would never make a success of anything." Henry Ford barely made it through school with the minimum grasp of reading and writing. Sir Isaac Newton was so poor in school that he was allowed to continue only because he was a complete flop at running the family farm. Pablo Picasso was pulled out of school at the age of ten because he was doing so badly. His father hired a tutor to prepare him to go back to school, but the tutor gave up on the hopeless pupil. Giacomo Puccini, the Italian opera composer, was so poor at everything as a child, including music, that his first music teacher gave up in despair, concluding that the boy had no talent.

Several of the forms of bias noted above, tend to influence less affluent groups most, which include most minorities. For example, 2005 median incomes for males were: Other, $27,041, Hispanics, $27,380, Blacks, $34,433, whites, $46,807 and Asians, $48,693 (Webster & Bishaw, 2006). Many minority individuals, particularly those in the lower socio-economic classes, rarely experience Standard English in their homes or communities, and frequently not even in the low-performing schools they attend (realize that the definition of low-performing almost always results from standardized tests that are language biased against the enrolled students and frequently the teachers as well). We can probably assume that the long-term gains on standardized measures shown by less affluent minorities in research such as that on small class size effects result largely from improved reading skills in Standard English (Mosteller, 1995; Illig, 1996).

FairTest (2006) documents a purposeful use of bias #2 above regarding scores on the SAT: "The SAT is designed solely to predict students' first year college grades. Yet, despite the fact that girls earn higher grades throughout both high school and college, they consistently receive lower scores on the exam than do their male counterparts. In 1994, girls averaged 41 points lower than boys on the Math section of the test, and 4 points lower on the Verbal section."[5]

[5] The reason these differences are lower than those reported in the current study is that the analyses conducted here control for HSGPA. Because girls consistently earn higher grades than boys, comparing the average girl's test scores to the average boy's effectively compares those having a 3.63 mean HSGPA (females) with those having 3.46 (males) for 217,743 SUS FTIC enrollees between summer 2000 and spring 2006.

The article reports that the sex gap favoring boys persists across all other demographic characteristics, including family income, parental education, grade point average, course work, rank in class, size of high school, size of city, etc. A study by Phyllis Rosser (1989), *The SAT Gender*[6] *Gap: Identifying the Causes*, found that the vast majority of questions exhibiting large gender differences in correct answer rates are biased in favor of males, despite females' superior academic performance. Rosser found that girls generally did better on questions about relationships, aesthetics and the humanities, while boys did better on questions about sports, the physical sciences and business.

This conclusion is supported by an earlier study by ETS researcher Carol Dwyer, who provides some historical perspective on the gender gap in her 1976 report. She notes that it is common knowledge among test-makers that gender differences can be manipulated by simply selecting different test items. Dwyer cites as an example the fact that, for the first several years the SAT was offered, boys scored higher than girls on the Math section but girls achieved higher scores on the Verbal section. ETS policy-makers determined that the Verbal test needed to be "balanced" more in favor of boys, and added questions pertaining to politics, business and sports to the Verbal portion. Since that time, boys have outscored girls on both the Math and Verbal sections. Dwyer notes that no similar effort has been made to "balance" the Math section, and concludes that "It could be done, but it has not been, and this suggests that either a conscious or unconscious form of sexism underlies this pattern. When girls show the superior performance, 'balancing' is required; when boys show the superior performance, no adjustments are necessary (Fairtest, 2006)."

Such information calls into question both the motives and trustworthiness of test developers, to whom we, as a nation are entrusting hundreds of millions of dollars each year plus the future of our youth. That group also reports research showing that other biasing elements consistently work against females, including, but not limited to: the multiple-choice format itself, speededness, and corrections for guessing.

In view of the preceding, it is not surprising that certain groups, particularly females, the less affluent and second language learners, score consistently lower than other groups. Thus, historical evidence suggests that such measures, rather than being objective estimates of talent or knowledge, are almost exactly the opposite.

The Lack of Relationship Between Admissions Tests and College Performance

Below is a summary of results from a typical sample of the hundreds of studies conducted at various colleges and universities regarding the relationship between prior academic variables (GPA and tests) and college performance. Such studies overall findings may be summarized for both undergraduate and graduate students by the following statement:

> Relationships between tests and any performance measures in college becomes essentially zero when one controls for more important and predictive factors such as prior performance (GPA), affluence (which may be imprecisely estimated by part-time/full-time enrollment) and sex.

[6] The word sex is generally used in this document rather than gender, because gender is a strictly linguistic term consisting of three categories in English: masculine, feminine and neuter. Having never met either a feminine or a neuter, this author prefers to use sex: male or female.

Even test makers themselves do not claim that standardized tests measure either achievement or school outcomes (Bracey, 1997). Test makers such as ETS and ACT do claim a low-level relationship with first semester college grades (Murphy, 2000; Elert, 1992).

Several studies suggest that test score relationships with performance, as measured by College GPA, differ between male and female students (Tusue & Whitaker, 1999; House, 1998; House, Gupta & Xiao, 1997), and between majority and minority students (Bieker, 1996; Lindle & Reinhart, 1998), or are influenced by other historical characteristics such as parental education (Stricker, & Rock, 1993).

Many studies report positive simple relationships (correlations) between tests (ACT, SAT, GRE or GMAT) and first semester grades in college, either graduate or undergraduate. However, even these relatively unimportant relationships are usually quite low generally ranging between ***r*** = .15 and ***r*** = .25. Elert (1992) summarizes numerous studies on SAT's relationship with first semester grades: "The best known record of prediction by the SAT, reported in a 1978 ETS survey of studies, was at a New Jersey college where the 1978 SAT-Verbal would have been matched as a predictor by random chance only 59% of the time. The worst result was reported at a university in Indiana where chance would have predicted grades as well as the 1972 SAT-Verbal 99.96% of the time (Breland & Minsky 149, p. 153)." A few other related studies include: Astin, 1993; Paszczyk, 1994; MSCHE, 1997; Chernyshenko, *et al*; 1999; House, Gupta & Xiao, 1997; Bieker, 1996.

Regarding this, Alexander Astin: "In a very practical sense, the student's ability to stay in college is a more appropriate measure of his success than is his freshman GPA. Although it is true that good grades will help him gain admission to graduate school, to win graduate fellowships, and even to secure certain types of jobs, they are irrelevant to any of these outcomes if the student drops out of college before completing his degree requirements." Astin found that using SAT scores to predict who will graduate resulted in 3.2% prediction for men and 2.9% for women. This means that for over 97% of the cases, random selection would predict the odds of remaining in school as well as the SAT. "Whether or not the student will drop out of college after the freshman year," Astin noted, "can be predicted with only a low degree of accuracy" (cited in Elert, 1992).

Elert (1992) supports the findings of many studies by stating that previous grades are about twice as good as the SAT at predicting academic achievement as measured by first semester grades. The main justification for requiring the tests for admissions is that although the SAT is an inferior predictor relative to high school grades, it can increase the accuracy of prediction when used in combination with them. However, research indicates that inclusion of the SAT increases early grade prediction by an average of only 5%. The major reason that the benefits are so low is that the SAT provides redundant information. Elert (1992) notes:

> Marginal as they are, the predictions of first year grades are the test's most accurate forecasts. Correlations between scores and grades in later years, and overall college average, are lower still. One study found that the ability of college admission tests to predict grades declined consistently from one semester to the next throughout eight semesters (Humphreys). The virtual disappearance of the aptitude tests' ability to predict beyond the freshman year has been explained by some commentators as a result of the nature of

> advanced study. Multiple choice testing dominates introductory courses, they argue, but intermediate and advanced courses demand a broader range of performance.

Crouse & Trusheim (1988) conducted a detailed statistical analysis of the SAT's predictive shortcomings. Using data from the National Longitudinal Study (NLS) of the high school class of 1972, they calculated the number of additional correct admissions using high school rank (HSR) alone and with the SAT. With four different measures of undergraduate success, they calculated that using the SAT in admissions adds between 0.1 and 2.7 additional correct forecasts per 100 applicants.

Many other studies, a few of which are noted, report one or more of the following: (1) tests fail to predict a more pertinent measure of success in college than first term grades (retention or graduation), (2) tests relate negatively with time-to-degree (higher scorers take longer to graduate) or (3) that tests' prediction capacity pales in comparison with other key variables (Bangura, 1995; Adelman, 2006; Waugh, Micceri & Takalkar, 1994) or graduate level (Xiao, 1998; Onasch, 1994; Sternberg & Williams, 1997; Wright & Palmer, 1994; Morrison & Morrison, 1995).

The Florida State University System (Florida SUS, 1995) which consisted at that time of nine universities (UF, FSU, FAMU, USF, UCF, FAU, FIU, UNF, UWF) conducted an extensive retention/graduation study that prompted the setting of sliding scale admission standards. Their study showed that once an applicant has a High School GPA of 3.0 or higher, test scores bear no relationship to retention/graduation. Below that 3.0 cut off, it may be worthwhile considering a student for UF or FSU if they show high test scores. The following quotes from (Florida SUS, 1995) speak to the issue of tests and admissions to college from the perspective of Florida in their first paragraph and from a literature review in their second:

> The greatest single predictor of success in College is High School Grades.
>
> Two methods are most frequently used to quantify high school performance: High School GPA and class standing or rank. Class standing is reported as a standard more often than is GPA.

To give some idea of how extensive such research has been, among the highest relationships in the testing literature between the SAT and anything other than another such test is a correlation of $r = .66$ reported by French & John (1967) between SAT scores and uric acid levels in the blood. Such ETS-sponsored studies show just how widely ETS has searched for something their product predicts empirically.

Methods

This study addressed the following research question:

> Do any consistent score differences occur on Standardized Tests between different sexes or race/ethnicities who exhibit the same historic academic performance as measured by HSGPA?

To assure adequate sample sizes in each cell, data from the Florida SUS Admissions files for all First Time In College (FTIC) applicants to all eleven SUS institutions for the Academic Years (AY) 1997-98 through 2005-06 were used to address the research question.

Variables and Data Analysis

High School GPA – HSGPA values are those reported in Admissions Files by SUS institutions. Students in high school obtain extra GPA points for taking AP, IB and Honors courses, thus the possible range of GPA values is from 0.0 to 5.0. For these analyses, in order to assure adequate sample sizes in each cell (Table 1), the GPA range was limited to 2.5 to 4.5 because some groups become comparatively rare at certain GPA levels (for example, Asians below 2.5 and most groups above 4.5.).

Race/Ethnicity – Self-reported classification obtained from SUS institutions. Analyses were limited to Asian, Black, Hispanic, Other and white, non-Hispanic. Other includes all students (American Indian, Other, etc.) who were not classified as unknown. Those classified as unknown were excluded from analysis.

Sex – Self-reported classification obtained from SUS institutions. Students were classified as male or female, all unknown cases were excluded from analysis.

Tests – Four different tests and subtests were submitted to analysis:

- SAT Total Scores
- SAT Quantitative Subscore
- SAT Verbal Subscore
- ACT Composite Score

The Florida SUS Admissions Files – The Florida Master Admissions Databases are housed at the Northwest Regional Data Center in Tallahassee, Florida, and are part of the Florida Department of Education. The Florida SUS Master Admissions Files provides relevant academic and demographic data for all applicants to any SUS institution.

GPA values were rounded to the nearest 10th of a point and mean test scores for the four tests were computed for each race/ethnic subgroup and for both sexes separately for all students within each group at each HSGPA level. SAS version 9.1 was used to compute the subgroup score values, and between group differences were obtained using Microsoft Excel 2004. Charts were created using Microsoft Excel.

Limitations

It is possible that GPA values differ from one school or school district to another. The possible influence of such effects on findings was addressed by assuring large samples for each cell (smallest among the major groups was 96 among Asians at 2.5 on the ACT). Large samples should weaken such effects on analyses.

One might ask the question of whether there should be a relationship between HSGPA and tests. This is the only claim regarding the validity of these tests for college selection that the test makers themselves assert (Fairtest, 2006). Further, as Table 1 and Figure 1 show, there is a consistent monotonic relationship between HSGPA to the tenth of a point and test score values that remains invariable across all tests and all race/ethnicity/sex subgroups.

Results and Discussion

Sample

The total sample of FTIC students included 1,094,414 cases, of whom 698,054 had SAT scores and 396,360, ACT Scores. Including only larger cell sizes (the smallest cell size for the SAT was over 530 cases, and for the ACT, 96 cases), by limiting GPA values of 2.5 to 4.5 reduced the total sample of SAT scores to 628,946, and ACT to 358,586 (Table 1).

Relationship between Mean Test Scores and HSGPA

Figure 1 depicts the monotonic nature of the relationship between GPA values at the 10th of a point and SAT total scores. Table 1 shows that for all groups and subgroups considered in these analyses, mean test scores increased monotonically as mean HSGPA increases. The table shows that for Other students the same consistent differences tend to occur and the same monotonic relationship between GPA and test scores. Because the test point increments are small (1/10th of a point), and given the monotonic nature of these variables relationship, if the test scores were unbiased, one would expect a relatively random distribution of means between and across groups for a given HSGPA score (e.g. 3.1). Thus, any group should have an equal chance of scoring higher or lower than any other group at a given GPA point. However, the results below show consistent differences favoring males and whites at almost all GPA scale points. Further, the mean differences between groups range from lows of about 20 points on the SAT, to highs of about 140 points. As was noted earlier, ETS (2001) suggests: “The user can be reasonably confident that a score difference of around 60 points or more indicates a true difference in ability between two test takers.”

Figure 1
Monotonic Relationship Between Test Scores and High School GPA (N=628,948)

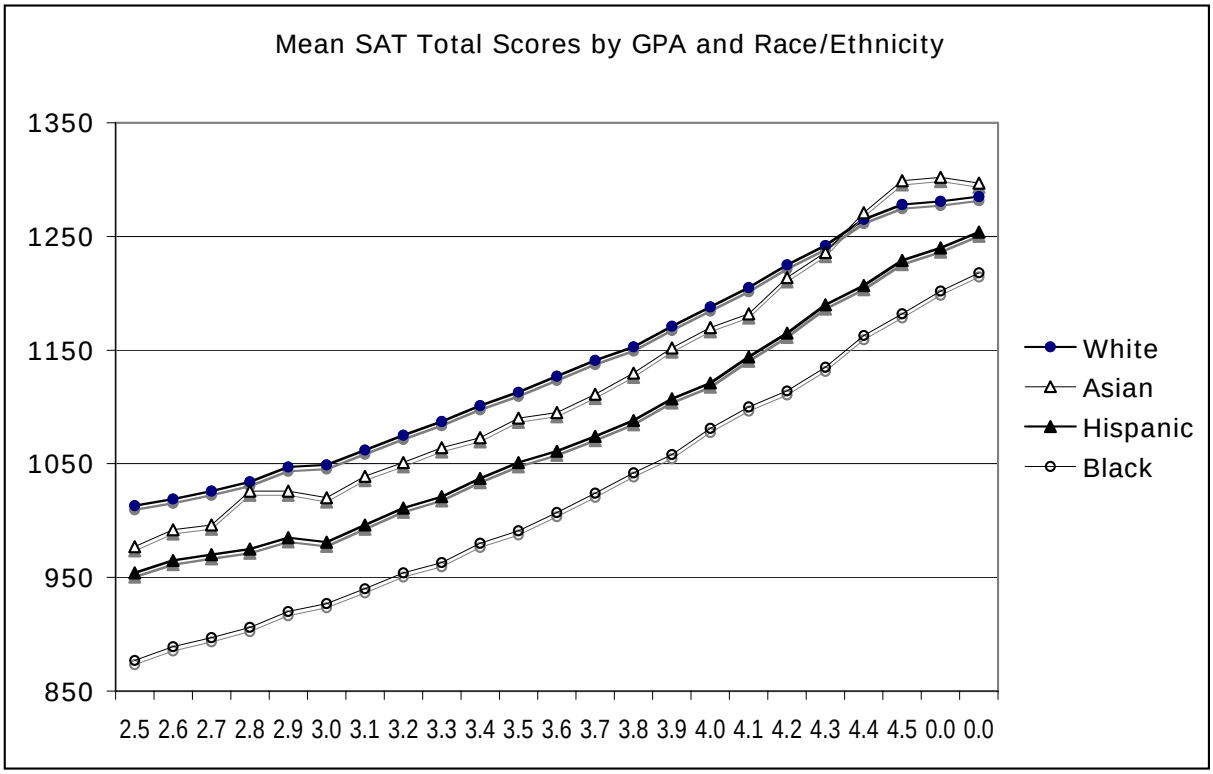


Table 1

Mean Scores on SAT and ACT Tests for All FTIC Applicants, Summer 1997 through Spring 2006

HS	Asian				Black				Hispanic				Other				White			
GPA	Female		Male		Female		Male		Female		Male		Female		Male		Female		Male	
	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
SAT Total Score																				
2.5	218	946	314	998	2347	843	2526	909	1299	914	1625	986	196	914	335	959	4313	971	6108	1042
2.6	232	942	367	1023	2787	857	2523	925	1612	931	1928	994	224	924	350	981	5326	978	7019	1050
2.7	325	951	472	1027	3082	866	2582	934	1882	931	2117	1004	182	930	309	961	6125	986	7682	1058
2.8	330	988	550	1049	3524	876	2773	945	2302	939	2363	1011	298	934	398	1005	7606	997	8916	1066
2.9	467	987	631	1055	3785	889	2779	963	2621	952	2666	1017	231	933	332	989	8784	1010	9616	1081
3.0	666	989	748	1047	4532	903	3006	964	3561	955	3130	1011	685	954	783	988	10131	1016	10550	1082
3.1	667	999	847	1070	4682	915	2961	978	3680	967	3027	1032	323	959	354	1030	10994	1028	10952	1097
3.2	723	1011	811	1087	4446	930	2540	996	3640	980	2951	1048	418	982	405	1043	11253	1041	10528	1111
3.3	811	1028	951	1095	4613	942	2384	1005	3872	994	2939	1057	422	998	421	1034	12500	1053	11512	1124
3.4	838	1042	868	1103	4310	959	2065	1022	3839	1006	2824	1079	430	1018	382	1078	12641	1068	11018	1140
3.5	906	1062	887	1119	4010	967	1907	1040	3793	1023	2737	1089	406	1027	351	1075	12871	1083	10803	1148
3.6	977	1063	950	1127	3758	988	1625	1053	3814	1037	2496	1098	442	1073	332	1112	13001	1096	10398	1165
3.7	1017	1084	866	1142	3172	1005	1363	1067	3496	1049	2164	1114	343	1044	247	1083	12247	1111	9548	1180
3.8	1072	1102	941	1163	2903	1023	1150	1091	3406	1065	2136	1126	408	1072	292	1134	12830	1127	9310	1190
3.9	1042	1126	852	1185	2625	1043	995	1098	3249	1087	1948	1140	281	1076	192	1114	12364	1144	8911	1210
4.0	1125	1140	800	1211	2080	1067	800	1118	2719	1099	1593	1158	430	1148	337	1161	11251	1162	7555	1226
4.1	1013	1160	730	1211	1875	1086	657	1140	2377	1121	1269	1187	170	1134	125	1145	10609	1181	6870	1241
4.2	908	1192	660	1244	1333	1097	447	1165	1906	1143	976	1209	156	1148	82	1192	8713	1203	5504	1259
4.3	929	1217	599	1266	1116	1126	374	1161	1578	1170	802	1230	113	1144	63	1202	8483	1219	5089	1281
4.4	779	1243	521	1314	763	1154	232	1193	1199	1189	654	1241	100	1179	45	1228	7099	1243	4029	1302
4.5	583	1277	424	1330	539	1175	149	1206	879	1208	431	1273	56	1187	32	1248	4951	1255	2802	1319
SAT Quantitative Score																				
2.5	218	475	314	518	2347	413	2526	456	1299	448	1625	496	196	470	335	508	4313	478	6108	528
2.6	232	473	367	526	2787	420	2523	464	1612	456	1928	499	224	478	350	518	5326	481	7019	532
2.7	325	478	472	533	3082	425	2582	470	1882	458	2117	506	182	476	309	505	6125	486	7682	537
2.8	330	499	550	547	3524	430	2773	476	2302	460	2363	509	298	476	398	531	7606	492	8916	541
2.9	467	497	631	551	3785	438	2779	484	2621	467	2666	512	231	481	332	530	8784	499	9616	550
3.0	666	499	748	544	4532	443	3006	487	3561	467	3130	511	685	491	783	528	10131	502	10550	550
3.1	667	506	847	556	4682	450	2961	493	3680	475	3027	523	323	492	354	547	10994	510	10952	558
3.2	723	513	811	569	4446	458	2540	503	3640	481	2951	530	418	500	405	552	11253	516	10528	567
3.3	811	522	951	572	4613	465	2384	509	3872	490	2939	536	422	517	421	555	12500	522	11512	573

HS	Asian				Black				Hispanic				Other				White			
GPA	Female		Male		Female		Male		Female		Male		Female		Male		Female		Male	
	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
3.4	838	532	868	576	4310	473	2065	518	3839	497	2824	549	430	526	382	576	12641	530	11018	582
3.5	906	541	887	584	4010	479	1907	528	3793	506	2737	553	406	529	351	580	12871	538	10803	586
3.6	977	542	950	590	3758	489	1625	536	3814	513	2496	560	442	552	332	594	13001	545	10398	596
3.7	1017	555	866	601	3172	498	1363	543	3496	519	2164	569	343	541	247	585	12247	553	9548	603
3.8	1072	564	941	614	2903	507	1150	555	3406	527	2136	576	408	549	292	609	12830	562	9310	610
3.9	1042	576	852	624	2625	517	995	560	3249	541	1948	584	281	565	192	608	12364	570	8911	619
4.0	1125	586	800	639	2080	530	800	568	2719	547	1593	593	430	601	337	629	11251	580	7555	629
4.1	1013	596	730	639	1875	538	657	584	2377	559	1269	607	170	585	125	616	10609	590	6870	635
4.2	908	614	660	656	1333	546	447	594	1906	569	976	621	156	594	82	628	8713	600	5504	646
4.3	929	625	599	667	1116	562	374	591	1578	584	802	631	113	591	63	632	8483	609	5089	656
4.4	779	637	521	688	763	574	232	612	1199	591	654	634	100	603	45	663	7099	621	4029	667
4.5	583	652	424	694	539	585	149	615	879	603	431	651	56	612	32	653	4951	626	2802	675
SAT Verbal Score																				
2.5	218	471	314	479	2347	430	2526	453	1299	466	1625	491	196	444	335	451	4313	493	6108	515
2.6	232	470	367	497	2787	437	2523	461	1612	475	1928	495	224	447	350	463	5326	497	7019	518
2.7	325	473	472	494	3082	440	2582	464	1882	473	2117	499	182	455	309	456	6125	500	7682	522
2.8	330	488	550	501	3524	446	2773	469	2302	479	2363	501	298	457	398	475	7606	505	8916	525
2.9	467	490	631	505	3785	451	2779	478	2621	485	2666	505	231	452	332	459	8784	511	9616	531
3.0	666	490	748	503	4532	459	3006	477	3561	487	3130	500	685	464	783	460	10131	514	10550	531
3.1	667	493	847	514	4682	466	2961	485	3680	493	3027	509	323	466	354	484	10994	519	10952	538
3.2	723	497	811	519	4446	472	2540	493	3640	499	2951	518	418	482	405	490	11253	525	10528	545
3.3	811	506	951	523	4613	477	2384	495	3872	504	2939	520	422	481	421	478	12500	530	11512	551
3.4	838	510	868	527	4310	486	2065	504	3839	510	2824	530	430	492	382	502	12641	538	11018	557
3.5	906	522	887	535	4010	489	1907	512	3793	518	2737	535	406	498	351	495	12871	545	10803	561
3.6	977	521	950	537	3758	498	1625	516	3814	524	2496	538	442	521	332	518	13001	551	10398	569
3.7	1018	530	866	541	3172	508	1363	524	3496	530	2164	545	343	503	247	498	12248	558	9548	577
3.8	1072	537	941	549	2903	516	1150	535	3406	538	2136	550	408	522	292	525	12830	565	9310	580
3.9	1042	550	852	560	2625	525	995	537	3249	546	1948	557	281	511	192	507	12364	574	8911	590
4.0	1125	554	800	572	2080	537	800	551	2719	552	1593	565	430	548	337	532	11251	583	7555	597
4.1	1013	564	730	572	1875	548	657	556	2377	562	1269	580	170	549	125	529	10609	592	6870	605
4.2	908	578	660	588	1333	551	447	571	1906	574	976	588	156	554	82	564	8713	603	5504	613
4.3	929	592	599	599	1116	564	374	570	1578	586	802	599	113	552	63	570	8483	610	5089	624
4.4	779	606	521	626	763	579	232	581	1199	598	654	607	100	576	45	565	7099	622	4029	635
4.5	583	625	424	636	539	590	149	590	879	605	431	621	56	575	32	594	4951	629	2802	643
ACT Composite																				
2.5	96	19	118	20	2290	17	1827	17	696	18	591	19	42	17	70	18	2317	20	2275	21

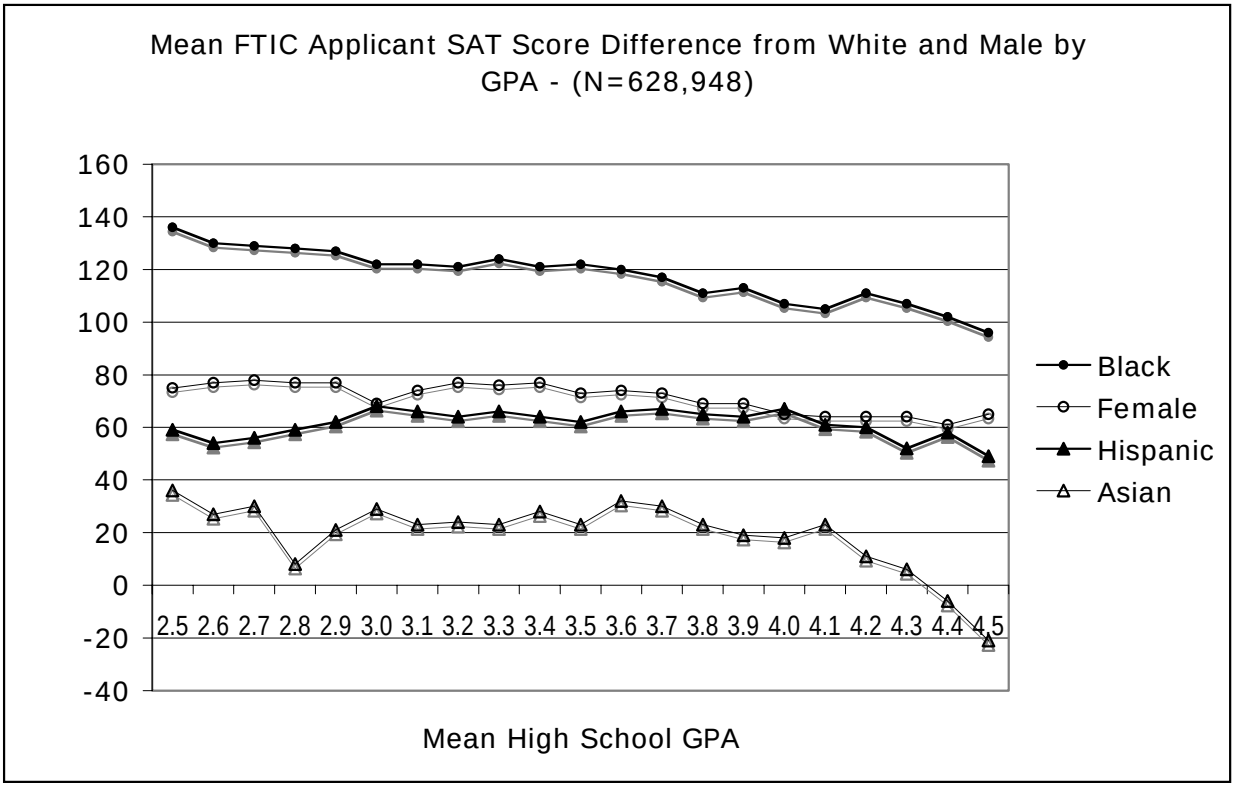
HS	Asian				Black				Hispanic				Other				White			
GPA	Female		Male		Female		Male		Female		Male		Female		Male		Female		Male	
	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
2.6	117	19	123	21	2676	17	1860	18	806	18	773	19	49	17	57	19	3018	20	2661	21
2.7	178	19	150	20	3192	17	1936	18	1061	19	803	20	44	18	76	18	3356	20	2785	21
2.8	180	19	197	21	3567	18	1968	18	1250	19	842	20	85	18	65	19	4366	21	3476	22
2.9	260	19	201	21	4006	18	1901	18	1515	19	947	20	99	19	84	19	4922	21	3809	22
3.0	342	20	243	21	4368	18	1988	19	2024	19	1097	20	137	18	76	19	6258	21	4179	22
3.1	387	20	305	21	4519	18	1829	19	2089	20	1136	21	124	19	77	20	6699	21	4275	22
3.2	383	20	300	21	4261	19	1576	19	1981	20	1041	21	104	19	80	20	6937	22	4353	23
3.3	442	21	319	22	4264	19	1514	20	2089	20	1127	21	109	19	69	20	7859	22	4649	23
3.4	541	21	306	22	3993	20	1294	20	2130	21	1038	22	111	19	82	21	8005	22	4527	23
3.5	583	21	309	23	3714	20	1157	20	2205	21	1026	22	128	20	76	22	8219	23	4572	24
3.6	570	22	325	23	3390	20	1024	21	2191	21	945	22	89	20	66	22	8635	23	4486	24
3.7	593	22	317	23	2974	21	891	21	1978	22	892	23	131	21	56	21	8228	23	4253	25
3.8	586	22	312	24	2616	21	705	22	1954	22	863	23	100	21	59	22	8569	24	4319	25
3.9	605	23	310	24	2230	22	638	22	1756	23	776	24	103	21	54	22	8469	24	4181	25
4.0	655	24	285	25	1853	22	480	23	1546	23	623	24	100	22	55	23	7918	25	3540	26
4.1	517	24	303	25	1651	23	399	23	1395	23	558	25	62	23	48	23	7432	25	3256	26
4.2	538	25	235	25	1172	23	269	24	1117	24	452	25	56	24	28	24	5976	26	2626	27
4.3	524	25	227	26	987	24	222	24	955	25	326	25	47	23	20	25	5720	26	2432	27
4.4	432	26	167	27	677	24	156	25	741	25	284	26	39	25	17	24	4676	27	1943	27
4.5	304	26	138	27	453	25	101	25	519	25	190	26	24	25	5	25	3315	27	1408	28

GAPS Become Smaller as GPA Values Increase

Figure 2 shows that as GPA values increase, the advantage favoring whites and males tends to become smaller. In fact, at the 4.3 GPA level, Asian students begin scoring higher than white students. However, given the 2001 statement by ETS about how a difference of 60 points represents a "true" difference in ability, one must realize looking at Figure 2, that if that statement is in fact accurate, Black and Female students always have "less" ability than whites and males, and that only at score points of 4.3 to 4.5 do Hispanic students have no "true" disadvantage in ability.

Figure 2

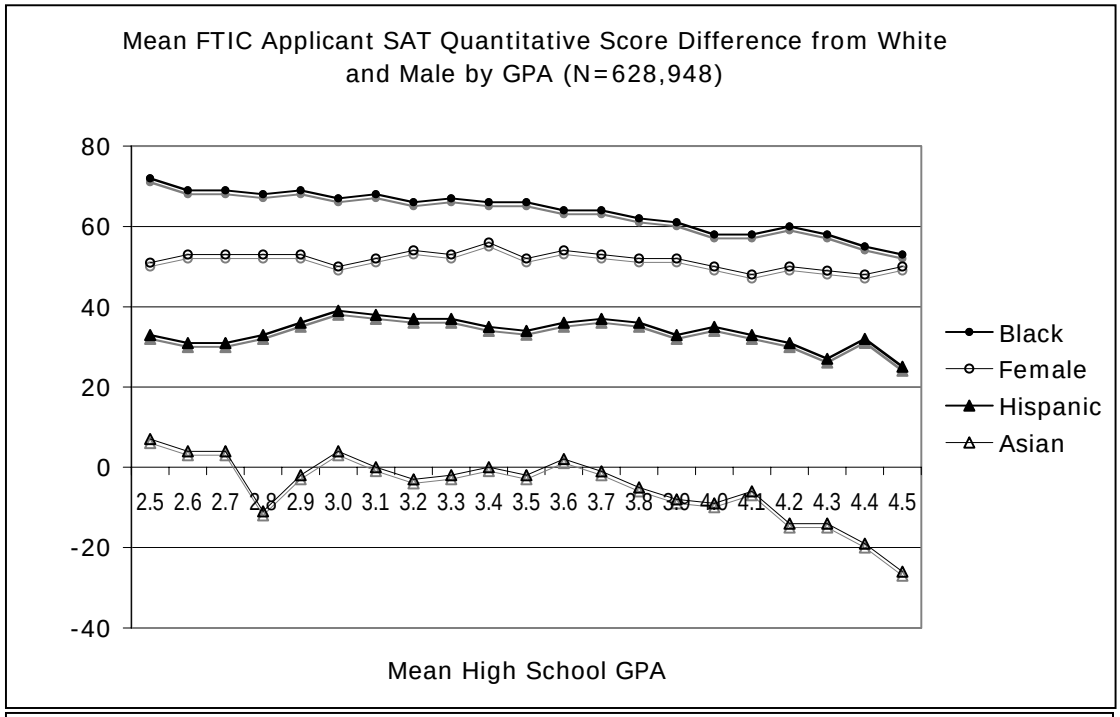
Racial/Ethnic and Sex Advantages of Whites and Males by GPA Category – Mean of SUS Fall Annual FTIC Cohorts 1997-98 through 2005-06

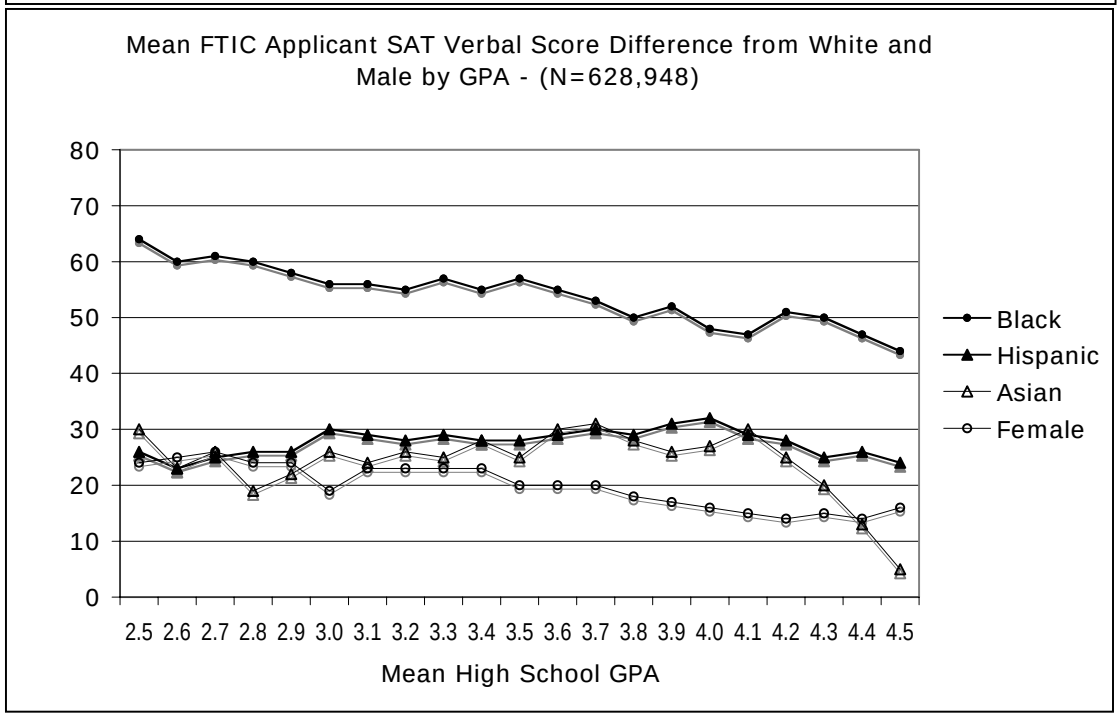


Quantitative and Verbal Subscore Differences

Figure 3 displays SAT quantitative and verbal subscore differences between whites and minorities and males and females. ETS (2001) notes that differences of 30 points on the subscores, although less reliable than the total score, represent "real" differences in ability. A few rather interesting differences occur between these two subscores. First, on the Quantitative portion (top panel), Asians almost always score as high as or higher than whites. Second, regarding verbal differences (bottom panel), three groups usually score near or below the 30 point difference criterion of "true" differences: Females, Asians and Hispanics. Finally, females always show only a 20 point or lesser disadvantage on the verbal subtest, but a 50 to 60 point disadvantage on the quantitative subtest.

Figure 3
Quantitative and Verbal Differences





Summary, Conclusion and Implications

Detailed analysis in this discussion was limited to SAT test scores, although similar consistent differences occurred for ACT tests as well. The following points appear to be consistent in these analyses:

- All groups and subgroups show a monotonic relationship between HSGPA values to the 10th of a point and test score values.
- Given that one should expect a random distribution of differences at the same 10th of a GPA level, these analyses indicate that consistent biases favoring whites and males result from the use of standardized tests as admission requirements.
- Usually the gaps favoring whites or males are well above the 60 points that ETS (2001) states reflects a "true" difference in ability. This "true" difference occurs where minority and female students are exhibiting exactly the same level of academic performance as measured by GPA (actual performance).
- The gaps favoring whites and males decrease somewhat as GPA values increase.

One vital point regarding all of these analyses is that the GPA is the criterion of success in school, whether high school or college. Standardized test scores, no matter how high, do not earn a degree, the certificate verifying academic success or failure. Thus, GPA must be considered the "true" score, with test scores merely a theoretical proxy. Secondly, it is important to realize that GPA tends to be a very reliable estimate of academic performance and certainly, next to affluence, the strongest predictor of success in college (Florida SUS, 1995; Mortenson, 1999, 2000). A students high school or college GPA derives from multiple observers (teachers), in multiple disciplines, over an extended period of time (three to six years for college). As a result, this tends to produce a considerably more valid and reliable estimate of academic performance than the point-in-time, strictly abstract estimate that tests provide.
In summary, these results are consistent and appear clear. This is particularly true when viewed within the context of the literature, in that a specific bias favoring whites and males results from the use of standardized tests as admissions criteria. Further, this appears to be a factor, if not a major factor in the recent movement of minority students from entry into SUS institutions as FTIC students to entry as CCT students because the Florida Community Colleges require only a high school diploma as a criterion for entry as a degree-seeking student.

The apparently increasing use of test scores as a criterion for entry to higher education appears to produce both an unjust and socially unwise bias favoring males, whites and the affluent (Heller & Rasmussen, 2006). This reflects poorly on the integrity of Florida's SUS in that institutions continue to perpetuate discriminatory techniques that reduce the access of minorities and females to the more prestigious academic pipeline. The problem is exacerbated as one moves up the ladder of prestige and its effects are well stated by Haycock & Gerald (2006, p. 7): "...the 50 flagship universities now look less and less like America—and more and more like 'gated communities of higher education.'" Further (p. 5): "...the highest achieving students from high-income families-those who earned top grades, completed the full battery of college prep courses, and took AP courses as well-are nearly four times more likely than low-income students with exactly the same level of academic accomplishment to end up in a highly selective university." These recent findings extend the research reported by such as Mortenson (1990). The current study indicates that one reason this occurs is the use of test scores that are biased against minorities and the less affluent. The use of tests is influenced by rankings like U.S. News and World Reports 'America's Best Colleges," with was specifically designed to sell magazines. Haycock & Gerald note (p. 3), the result is: "Rated less for what they accomplish

with the students they let in than by how many students they keep out, many of these flagship institutions have become more and more enclaves for the most privileged of their state's young people." Of course, most faculty come from the elite colleges, thus perpetuating the historic academic status quo of white male faculty.

A Call for Change

Study after study has shown that neither for graduates nor undergraduates do standardized tests provide any useful information beyond that provided by GPA regarding a student's likelihood of success. For undergraduate students, two factors have shown to be the best predictors of academic success, HSGPA and the rigor of high school coursework a student takes (Adelman, 2004). Standardized tests should probably be used as a criterion for college entry in one and only one situation: when a student has not taken school seriously through high school, a high test score may indicates good skills in standard English, which is a vital factor in higher education success, and therefore suggest taking a chance on students with low performance in high school (e.g. 2.0 GPA).

References

Adelman, C. (2006). *The Toolbox Revisited: Paths to Degree Completion From High School Through College.* Washington, D.C.: U.S. Department of Education, 2006. Downloaded from the WWW on August 1, 2006: http://www.ed.gov/rschstat/research/pubs/toolboxrevisit/toolbox.pdf

Adelman, C. (2004). *The Empirical Curriculum: Changes in Postsecondary Course-Taking, 1972-2000*. Washington, DC: U.S. Department of Education. Downloaded from the WWW on April 6, 2005: http://www.ed.gov/rschstat/research/pubs/empircurr/empircurric.pdf

Astin, A.W. (1993). *How Good Is Your Institution's Retention Rate?* Unpublished manuscript, University of California, Los Angeles.

Bangura, A.K. (1995). *Hold Everything!: Emerging Problems in Institutional Accountability for Retention and Graduation*. ERIC_NO: ED410845

Baum, S. & Payea, K. (2004). *Education pays 2004: The benefits of higher education for individuals and society*. The College Board. Downloaded from the WWW, March 2005. http://www.collegeboard.com/prod_downloads/press/cost04/EducationPays2004.pdf

Bieker, Richard F. (1996). Factors Affecting Academic Achievement in Graduate Management Education. *Journal of Education for Business*; v72 n1 p42-46 Sep-Oct 1996.

Borman, K., Workman, C., Miller, J. & Micceri, T. (2006). *The Role of the Community College in Contemporary Society: The Florida Context*. Paper presented at the Annual American Anthropological Association Conference, San Jose, CA, November 15-20.

Bracey, G.W. (1997). *Setting the Record Straight: Responses to Misconceptions about Public Education in the United State*s. Association for Supervision and Curriculum Development. Alexandria, Va. Downloaded from the WWW in August, 1999: http://www.ascd.org/readingroom/books/bracey97book.html..

Breland, H.M. & Minsky, S. (1978). Population Validity and College Entrance Measures. Princeton, NJ: Educational Testing Service, 1978.

CDC (2006). U.S. Census Bureau Census Download Center. Data downloaded from the WWW on August 28, 2006: http://tinyurl.com/guad6.

Chernyshenko, Oleksandr S.; Ones, Deniz S. (1999). How Selective Are Psychology Graduate Programs? The Effect of the Selection Ratio on GRE Score Validity. *Educational and Psychological Measurement*; v59 n6 p951-61 Dec 1999.

Commonfund (2006). *2006 HEPI – Higher Education Price Index*. Commonfund Institute. Downloaded from the WWW on September 1, 2006 (http://tinyurl.com/jqhch).

Day, J.C. & Newburger, E.C. (2002). *The Big Payoff: Educational Attainment and Synthetic Estimates of Work-Life Earnings*. Special Studies, U.S. Census Bureau, Downloaded from the WWW on January 5, 2007: http://tinyurl.com/46s3.

Elert G. (1992). The SAT, Aptitude or Demographics? *E-World*, Downloaded from the WWW in December, 1999. http://hypertextbook.com/eworld/sat.shtml

ETS (2001). *Test Characteristics of the SAT I: Reliability, Difficulty Levels, Completion Rates*. Downloaded from the WWW in February, 2002: http://www.collegeboard.org/sat/cbsenior/stats/stat002.html or http://tinyurl.com/38xu4r.

FairTest (2006). *Gender Bias in College Admissions Tests*. The National Center for Fair and Open Testing. Cambridge, MA. Downloaded form the WWW in June, 2006: http://fairtest.org/facts/genderbias.htm..

Florida SUS (1995). *A Study of Admissions Criteria*. Technical Report, Florida Board of Regents, Tallahassee, FL.

French Jr., John R. P. "Quantification of Organizational Stress." *Managing Organizational Stress: Proceedings of the Executive Study Conference*: November 29-30, 1967. Princeton, NJ: Educational Testing Service, 1968: 5-41.

Gibson, R. (2001). *The Fascist Origins of the SAT Test*. BRC-NEWS, April, 2001. Downloaded from the WWW on April 8, 2001: http://www-rohan.sdsu.edu/~rgibson/SATFascistOrigins.htm or http://tinyurl.com/3ajnhx

Goldin, C. & Katz, L.F. (1999). The shaping of higher education: The formative years in the United States, 1890 to 1940. *The Journal of Economic Perspectives*, 13:1, pp. 37-62.

Grapevine (2006). *An Annual Compilation of Data on State Tax Appropriations for the General Operation of Higher Education*. Center for the Study of Education Policy, Illinois State University. Downloaded from the WWW on September 1, 2006 (http://www.grapevine.ilstu.edu/)

Haycock, K. & Gerald, D. (2006). *Engines of Inequality: Diminishing Equity in the Nation's Premier Public Universities*. Technical Report of the Education Trust, downloaded from the WWW on November 20, 2006: http://www2.edtrust.org/NR/rdonlyres/F755E80E-9431-45AF-B28E-653C612D503D/0/EnginesofInequality.pdf, or http://tinyurl.com/ydlcru.

Heller, D.E. & Rasmussen, C.J. (2006). Merit Scholarship and College Access: Evidence from Florida and Michigan. A Selection from: *Who Should We Help? The Negative Social Consequences of social of Merit Scholarships*, Downloaded from the WWW on February 20, 2007: http://tinyurl.com/2ffw44 or http://www.civilrightsproject.harvard.edu/research/meritaid/4MIFLHellerch2.pdf

Hodgkinson, H. (1999) *CONNECTION: New England's Journal of Higher Education and Economic Development*. September 22, 1999 - Downloaded from the WWW in October, 2000. http://www.nebhe.org/bkissues/Sum99.html

House, J. D. (1998). *Gender Differences in Prediction of Graduate Course Performance from Admissions Test Scores: An Empirical Example of Statistical Methods for Investigating Prediction Bias*. AIR 1998 Annual Forum Paper. ERIC_NO: ED424810

Illig, D.C. (1996). *Reducing Class Size: A Review of the Literature and Options for Consideration*. http://www.library.ca.gov/CRB/clssz/clssiz.html. Downloaded from the WWW in May, 1999.

Lemann, N. (1999). *The big test: the secret history of the American meritocracy*. Farrar, Straus and Giroux, NY.

Lemann, N. (1999a). Behind the SAT. *Newsweek*, 9/6/1999, pp. 52-56.

Lindle, Jane Clark; Rinehart, James S. (1998). *Emerging Issues with the Predictive Applications of the GRE in Educational Administration Programs: One Doctoral Program's Experience*. ERIC_NO: ED424632.

Micceri, T. (2001). *Why Do We Waste So Many Resources on Incredibly Costly and Practically Worthless Standardized Tests?*. Internal Technical Report, Office of Budget and Policy Analysis, USF, Downloaded from the WWW on September 9, 2005: (http://tinyurl.com/oml8g).

Micceri. T. (2005). *An Attempt to Identify Comparatively Supportive and non-Supportive Environments for Underrepresented Minorities and Females in SUS Colleges of Engineering*. Internal Technical Report. AAREA. University of South Florida, Tampa, FL., December, 2005.

Morrison, T. & Morrison, M. (1995). A Meta-analytic Assessment of the Predictive Validity of the Quantitative and Verbal Components of the Graduate Record Examination with Graduate Grade Point Average Representing the Criterion of Graduate Success. *Educational and Psychological Measurement*; v55 n2 p309-16 Apr 1995

Mortenson, T.G. (2000). *Opportunity for Higher Education in Florida in the Human Capital Economy*. Presentation to the Florida Postsecondary Education Planning Commission.

Mortenson, T.G. (1999). The Impact of Financial Aid on College Retention for Minority Students. *Postsecondary Education Opportunity*, Oskaloosa, IA, 52577.

Mortenson, T.G. (1990). The Reallocation of Financial Aid from Poor to Middle Income and Affluent Students, 1978 to 1990. *ACT Student Financial Aid Research Report Series 90-2*.

Mosteller, F. (1995). The Tennessee Study of Class Size in the Early School Grades. *The Future of Children*. 5:2,

MSCHE (1997). *Relationship between High School and College Performance by Maryland Students. Student Outcome and Achievement Report*. Maryland State Commission on Higher Education, 16 Francis Avenue, Annapolis, MD 21401; phone: 410-974.2971.

Murphy, M. (2000). *Predicting Graduation: Are Test Score and High School Performance Adequate*. Paper presented at the AIR Annual Forum, Cincinnati, OH, May 21-24, 2000.

Onasch, C. (1994) – *Undergraduate Grade Point Average and Graduate Record Exam Scores as Predictors of Length of Enrollment in Completing a Master of Science Degree.* (ERIC No: ED375739)

Palmer, J.D. (2004). *GRAPEVINE Survey of State Higher Education Tax Appropriations for Fiscal Year 2004*. Center for the Study of Education Policy, Illinois State University, Normal, IL.

Paszczyk, S. L. (1994). *A Comparative Analysis of ACT Scores and Final GPAs of Chicago State University Undergraduate Students*. ERIC_NO: ED370519

Raven, J.C. (1939). The R.E.C.I. Series of Perceptual Tests: An Experimental Survey. *British Journal of Medical Psychology*. 18, pp. 16-34.

Robbins, J. (1987). *Diet for a New America*. Stillpoint Publishing, Walpole, NH.

Rosser, P. (1989). *The SAT Gender Gap: Identifying the Causes*. Center for Women's Policy Studies, Washington, D.D.

Sahadi, J. (2006). Wealth gap widens: Chasm between wealthiest households and everyone else has grown more than 50% since the early 1960s. CNNMoney.com. August 29 2006. Downloaded from the WWW on December 28, 2006: http://tinyurl.com/hs7u

SHEEO (1997). *State Higher Education Appropriations*. State Higher Education Executive Officers. Boulder, CO. Downloaded from the WWW on September 1, 2006 (http://www.grapevine.ilstu.edu/Appropriations1996-97.pdf)

Sternberg, G. & Williams, J. (1997). Does The Graduate Record Examination Predict Meaningful Success in The Graduate Training of Psychologists? A case study. *The American Psychologist*, 52, 630.

Stricker, Lawrence J.; Rock, Donald A. (1993). *Examinee Background Characteristics and GRE General Test Performance*. GRE Board Research Report No. 89-07R. ERIC_NO: ED385589Tisue, Linda J.; Whitaker, Dianne P. (1999). The GRE: A Barrier to Women Returning to Graduate School? *Journal of Educational Opportunity*; v17 n2 p31-43.

Vogel, M. (2006). The College Admissions Squeeze: Florida Colleges. *Florida Trend*. August 31, 2006.

Waugh, G., Micceri, T. & Takalkar, P. (1994). *Using Ethnicity, SAT/ACT Scores, and High School GPA to Predict Retention and Graduation Rates*. Paper Presented at the Annual FAIR Conference, Orlando, FL, June 15-17, 1994.

Webster, B.H, Jr., & Bishaw, A. (2006). *Income, Earnings, and Poverty Data From the 2005 American Community Survey*. U.S. Census Bureau, American Community Survey Reports, ACS-02, U.S. Government Printing Office, Washington, DC.

Wellman, J. (1999). *Contributing to the civic good: Assessing and accounting for the civic contributions of higher education*. The New Millennium Project on Higher Education Costs, Pricing and Productivity. The Higher Education Policy Institute. Downloaded from the WWW, January, 2002. http://tinyurl.com/vpcb5.

Witte, G. & Henderson, N. (2004). *Wealth Gap Widens For Blacks, Hispanics: Significant Ground Lost After Recession*. Washington Post, October 18, 2004; Page A11. Downloaded from the WWW on December 28, 2006: http://tinyurl.com/yaxokd.

Xiao, B. (1998). *Factors Influencing Master's Degree Attainment in Business, Engineering, Health and Human Sciences and Visual and Performing Arts*. Paper Presented at the AIR Annual Forum, Minneapolis, MN, May 17-20, 1998.