

The Pennsylvania State University

The Graduate School

College of Education

**FROM IQ TO THE SAT AND ITS IMPACT ON AFRICAN
AMERICANS: PROMISE FOR THE SAT-OPTIONAL
MOVEMENT**

A Thesis in

Educational Theory and Policy

by Horatio W. Blackman

© 2012 Horatio W. Blackman

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

August 2012

The thesis of Horatio W. Blackman was reviewed and approved* by the following:

Mindy L. Kornhaber

Associate Professor of Education

Thesis Advisor

Katerina Bodovski

Assistant Professor of Education

Gerald LeTendre

Professor of Education and International Affairs

Head of the Department of Education Policy Studies

*Signatures are on file in the Graduate School.

Abstract: This work traces the rise of the SAT, from IQ testing to the tests' prominence in college admissions. Issues concerning the gap in SAT scores and the subsequent barrier to access to higher education for African-American students are also considered. This thesis highlights an increasingly used policy, that of SAT-optional admissions, at many of our nation's institutions of higher education. The basis for the policy being employed is the persistent score-gap on the SAT between minority students and their white peers.

Table of Contents

List of Tables.....	v
Introduction.....	1
Chapter 1. The Intelligence Quotient.....	3
Testing IQ.....	5
Validity: Army Alpha and Beta.....	8
Chapter 2. The Scholastic Aptitude Test.....	11
Illustrating the SAT Score Gap.....	12
Explanations for the Score Gap.....	13
Importance of Attending a Selective Institution.....	24
The SAT as a Barrier to Access to Higher Education.....	25
Chapter 3. SAT-Optional Admissions.....	28
Conclusion.....	33
References.....	35

List of Tables

Table 1. “SAT mean scores of college-bound seniors by race: 1999-2000 through 2010-2011” p. 12

Table 2. “SAT mean scores by income, 2011” p. 19

Introduction

Modern forms of human intelligence testing began around the turn of the 20th Century. The rise of this form of testing coincided with the expansion of mass education in developing European countries and in the United States. It was during this time that the Intelligence Quotient, “IQ”, came to be defined and subsequently used to classify individuals and groups.

In the United States, IQ would become intricately intertwined with access to higher education and the socioeconomic benefits that obtaining a post-secondary degree confer. The SAT exam, derived from early IQ tests, has become the most influential instrument in college admissions. However, the SAT, much like its IQ predecessors, places the average African-American at the bottom of the distribution of scores, thus limiting access to higher education for our nation’s black youths.

Currently, over 750 colleges and universities across the country have adopted SAT-optional admissions policies that show benefits to the college-going African-American population. However, the majority of post-secondary institutions across the country, including some of our most elite, still place considerable weight on SAT scores, thus restricting opportunities for blacks to attend these schools; particularly our nation’s more selective universities. The movement towards SAT-optional admissions policies needs to be emulated by a greater proportion of our nation’s colleges and universities if we are to achieve a more equitable higher-education system.

This paper first discusses the rise of IQ testing and the basis on which it was founded. This discussion will lay the foundation for problematic issues of the SAT itself, and consequent detrimental impact on access to higher education for African-Americans in the United States.

The final piece of this paper focuses on the path many universities have taken towards adopting SAT-optional admissions practices and the increased equity it yields concerning admission to our nation's colleges and universities. It is the goal of this paper to highlight the shortcomings of the SAT, its harmful effect on black college-aspirant, and to shed light on a recent practice: the SAT-optional admissions policy that has proven beneficial to this traditionally underserved minority population.

Chapter 1. The Intelligence Quotient

In the late 19th Century, the French government underwent a major shift from religious education to secularized, mass education. Recognizing that mass education generated a heterogeneous student body, including students with learning difficulties, the government recruited Alfred Binet, a leading French psychologist of his time who had worked extensively on methods to measure human intelligence. In 1904, Binet was tasked by his government to develop a measurement, a test, to identify students who performed poorly in school due to intellectual difficulties in order to provide them with special education services. In fulfillment of this request, he created a test that consisted of a series of short tasks designed to measure a student's ability to comprehend, reason, and carry out other basic mental activities such as counting and ordering (Gould, 1996). By assessing a student using a variety of tasks, Binet believed that he would be able to capture the general intellectual potential of the child, thus uncovering whether or not their poor school performance was the result of a learning disability that required them to receive some type of special education.

The scale that was used to score the exams was designed using what Binet coined "mental age". The tests were administered with the tasks in order of increasing difficulty. An age level was tied to the tasks, and once a student could not progress past the tasks associated with a given age, the student would be assigned that mental age (Boake, 2002). That number was then compared to their chronological age, and if the computation resulted in a number of -1 or more (negative), the student could be identified as low-functioning (Binet & Simon, 2007). In addition, Binet stated that a child, no matter how little he can answer on the test, should not be considered "defective in intelligence" unless the computation resulted in a number lower than -2 (Binet & Simon, 2007).

Binet did not intend for a number to be affixed as a permanent evaluation of a child. The number was merely a diagnostic tool. Both he and the French government sought to identify those in need of special education. Binet designed the test for this purpose, and this purpose alone. Although Binet was interested in discovering a method to correctly measure human intelligence, he understood the limitations of the test he had designed (Binet & Simon 2007). Throughout the design of this test, and continuing after its rise in importance concerning human intelligence, Binet cautioned against the many interpretations that would eventually be drawn from this test and its descendants. For example, Binet warned that labeling individuals based on a single diagnostic, could create situations in education where teachers would see a child as incapable of learning and put little effort into actually educating the student (Binet & Simon, 2010).

Binet sought to separate learned knowledge from natural intelligence, but he warned that intelligence could not be fully captured by a single piece of information; and certainly not from the test he had created. In *A Method of Measuring the Development of Intelligence of Young Children* (2007), Binet states that "...we [Simon] think it chimerical and absurd to judge the intelligence of a child by one test alone." If intelligence were to be reified (to define something as tangible and real), Binet argued, it would lead to a host of social and educational issues (Binet, Simon & Kite, 2010). He postulated that the reification of human intelligence would lead to the classification of certain individuals and groups as biologically inferior and thus unable to achieve at the same level as other, higher scoring people and groups (Gould, 1996). Despite these warnings, Binet's work would prove to be the first step in the development of the Intelligence Quotient and would have a major influence regarding the realization of the problems he warned might come.

Testing IQ

The Intelligence Quotient was officially created by psychologist William Stern. He used the same mental age from Binet's work but instead divided it by the individual's chronological age to yield an IQ score. Binet and Stern can be credited for their early work in the measurement of intelligence, but it would be several other psychologists, including Lewis Terman and Robert Yerkes who would popularize intelligence testing in the United States and greatly influence the nation's hereditarian beliefs.

Lewis Terman, a Stanford professor, had also been studying human intelligence at the time of Binet's work in France. When Binet's tests were published, it gave Terman a template through which to extend this testing to all age groups, not just school-age children. In doing this, Terman's goal was to prove the innateness of intelligence in order to "properly" classify, and subsequently, sort, people into their appropriate positions in society. Terman believed that he could create a technocracy, where the most intelligent would occupy the positions of most importance in our society and the less intelligent would accept their positions in life based on this testing. IQ testing in primary school would allow the nation to sort its students for training early on (Terman, 1920). In addition, Terman argued that this sorting would create a more efficient educational system.

Terman's Stanford-Binet test, the re-designed version of Binet's, but extended to a wider age range, would ultimately stand as the premier IQ test in the United States. Terman would use this test to measure the IQ of thousands of individuals across races and classes. From a statistical standpoint, Terman would find low-to-moderate correlations of .30 between arithmetic fundamentals and .35 on vocabulary – measures he used for IQ – and social class (quantified by

salary; Terman et al., 1917). Despite this relatively modest finding, Terman would move forward with his testing by studying average differences between races.

Before entering the discussion of racial differences in average IQ, two prominent camps of thought with regard to human intelligence need to be defined. The creators of the IQ tests (with the exception of Binet) were hereditarians. They believed that human intelligence was a product of one's biology. On the other side of this debate were those who believed that intelligence was malleable and was largely influenced by one's environment. They argued that proper training and education were overriding factors that influenced an individual's intellectual ability. While it is my personal belief that the roots of intelligence lie somewhere between biology and the impact of one's environment, hereditarians, like Terman, dismissed the claims of environmentalists despite less than adequate evidence to support their own claims.¹

In fact, Terman and his fellow biological determinists committed what has been termed the hereditarian fallacy. This fallacy pertains to two main claims, or assumptions, made by hereditarians (Gould, 1997). The first is that heritability has become synonymous with inevitability, or the static nature of a trait. In other words, intelligence, if we follow this fallacy, would be considered an innate trait that cannot be modified, positively or negatively, through one's environment (Gould, 1997). Simply because a trait is heritable however, it does not equate to being unalterable (Bailey, 1997; Gould, 1997). The second claim deals directly with issues of race and intelligence. The claim refers to comparisons of heredity both within-group and between-groups (Gould, 1997). Those grounded in hereditary theory apply their explanations for within-group differences (i.e., within a specific race) in intelligence to findings of between-group

¹ The two camps of thought described here are historical ideologies that have influenced policy over time. Modern science, on the other hand, has come to a better understanding about nature and the impact environment has on outcomes, broadly speaking.

differences (i.e., across races). However, this is a practice that is imprudent because two groups, for example White-Americans and African-Americans, operate under vastly different environmental conditions, on average.

In 2009, the U.S. Census Bureau reported that only 37 percent of black children lived in a household with two parents, whereas three-quarters of their white peers had both parents present in the home (Kreider & Ellis, 2011). The poverty statistics across races also highlight stark differences in environmental conditions. In the United States, 38.2 percent of blacks under the age of 18 were living in poverty. That number sat at a comparably low 12.4 percent for non-Hispanic white youths (DeNavas, Proctor & Smith, 2011). The relatively high level of poverty for blacks in the U.S. contributes to their higher rates of mental health issues and negatively impacts children's cognitive development (Aber, Bennet, Conley, & Li, 1997). That intelligence **may** have a high level of heritability, does not mean that the average difference between two groups' IQ scores is a result of innate biological differences. Environmental causes cannot be ruled out even in the event of high heritability of a trait (Gould, 1996).

Hereditarian views about intelligence had been widely circulated near the end of the 19th Century, and early decades of the 20th Century. These views influenced efforts at intelligence testing. Mass IQ testing gave these beliefs more solid ground on which to further hereditarian views. World War I would present itself as a prime opportunity for IQ testing to enter mainstream consciousness. Due to WWI, the government needed millions of young American men for the war. Robert Yerkes enlisted the help of fellow hereditarians, including Lewis Terman, to design the IQ tests that would, in their estimation, correctly sort soldiers into the position (rank) that would be most effective for the military. Yerkes' WWI Army Alpha test was a compilation of many of the same types of questions seen on the Stanford-Binet: analogies,

ordering, and questions evaluating vocabulary/grammar. The Army Beta test included mazes, geometrical construction, and picture completion, among others (Sticht, & Armstrong, 1994). Yerkes and other hereditarians used this large pool of tested individuals to compile data that would then be used to advance their position regarding human intelligence. The scores were disaggregated by race, and, as Yerkes expected, White-Americans scored the highest on average with blacks having the lowest average scores. Believing that this test was a sound measure of raw human intelligence, Yerkes would go on to state that the differences noted between the races (and even by ethnicity within-race) must be due to innate differences in intelligence. However, the Army Alpha and Beta were highly correlated with time spent in the U.S. – especially pertinent when looking at European immigrants – indicating that the test was more a measure of acculturation than intelligence. Despite this (and other) evidence to the counterfactual, Yerkes and his fellow hereditarians would continue to influence the nation’s beliefs and eventually how the nation would come to use these tests to decide the futures of its youth.

Validity: Army Alpha and Beta

The Army Alpha content and administration procedures, alongside many potential, alternate explanations for the results, indicate that the test was not accurately measuring innate intelligence. As a result, claims were made against the test. For example, some argued that the test measured how immersed in American culture the individual or group of test takers was, rather than intelligence (Lippmann, 1976). Within the testing field, there are a number of criteria used to determine how well designed, and administered, the test is. For the purposes of the Army Alpha, a key criterion is validity. When assessing validity, there are several factors that determine how valid the results are, but only the factors most pertinent to the Army Alpha will be brought to attention in this paper.

In testing, validity does not, and cannot, refer to the test itself. A test is simply a compilation of a series of questions or tasks. The conclusions drawn from the test results are what are judged to be valid or invalid. For a valid inference to be drawn from the results of an exam, the test must adequately measure what the test states it should measure. The questions on the exam must be proven to measure only what is in the domain (the range of knowledge/skills) that is sought to be measured and to capture as little extraneous information as possible that may, positively or negatively, influence the scores. What has been described here is referred to, statistically, as construct validity (or lack thereof). The Army Alpha and Beta have long been known to be highly correlated with measures of acculturation, thus reducing the validity of claims that it accurately measures innate intelligence.

Conclusions drawn from the Army Alpha and Beta also provide a perfect example of validity being negatively affected by the administration of a test. When the tests were administered to nearly 2 million soldiers in 1917, issues that would limit the validity of Yerkes' claims that blacks were innately inferior in intelligence were apparent. One main component of validity is having standardization in the administration of a test (Koretz, 2008). This, however, was something that the army tests severely lacked. Testing conditions varied greatly from base to base, and standards for who was to take which test (which was to be determined by literacy) were not routinely followed. Blacks, who overwhelmingly took the Army Beta, were presented with the least efficient (comparably) conditions for taking an examination (Gould, 1996). Lacking equitable conditions, inferences made from the tests about innate intelligence differences should not have been heralded as such.

Without delving too deeply into racist undertones evident in this work, Gould, writing on the Army Alpha, mentions that at one military base, black examinees were further separated by

skin tone. With the knowledge that tests were already not being administered in a standardized fashion, it is no far stretch to argue that those creating and administering the test had beliefs that mirrored their practices about intelligence and the color of one's skin. Many of the validity issues that the Army Alpha and Beta faced resurface in the SAT. It really is no surprise that this is the case, as the SAT was modeled on the Army Alpha.

Chapter 2. The Scholastic Aptitude Test

Carl Brigham, a colleague and mentee of Yerkes', saw the Army Alpha as a step towards creating an impartial, "natural aristocracy" (Lemann, 1999). His beliefs led him to author the Scholastic Aptitude Test in hopes of finding and selecting those college applicants with the greatest intellectual potential. In doing so, Brigham believed, the nation's elite institutions of higher education could further educate and train these bright youths to occupy the most important and influential positions in our country. In effect, the SAT would create a fair and efficient society where the most intelligent – as defined by the SAT – would "naturally" find themselves atop the social hierarchy (Lemann, 1999).

In principle, the natural aristocracy that Brigham and others sought to create may appear to many as an extremely efficient way to organize society. Selecting the most intelligent people in our country for the best training to lead our nation could only result in a better, more structured nation, with little dissent from those further down the hierarchical structure. However, Brigham, like his colleagues who helped design the Army Alpha, was an ardent eugenicist, believing in the intellectual inferiority of blacks. In one of his primary works, *A Study of American Intelligence*, Brigham stated that "The decline of American intelligence will be more rapid than the decline of the intelligence of European groups, owing to the presence here of the Negro" (Brigham, 1923, p. 210).

Brigham's SAT was a direct descendant of the Army Alpha. The test would be accepted as the primary college entrance exam by the nation's universities for over a half-century, continuing to this day. However, much like the Army Alpha, the SAT, due to both design and to environmental issues beyond the test, has had the effect of creating a large gap in scores that

directly impacts blacks' access to higher education in the U.S., especially at the nation's more selective universities.

Illustrating the SAT Score Gap

Before discussing key theories as to the cause of the persistent score gap on the SAT between White-Americans and African-Americans, it is pertinent to highlight the large difference in the scores that has endured since large numbers of blacks began taking the college entrance exam. In 1976, the first year that the College Board released a report on the state of racial differences in scores on the SAT, the average black score was nearly 240 points below that of their white peers (Nettles et. al, 2003). The subsequent decade saw a statistically significant decline in the gap between scores with the aforementioned gap being reduced by approximately 25 percent, down to 189 in 1988. However, the continued gains were short lived, with the reduction in score gap stagnating, and even showing a slight increase, to the present day (Anonymous, 2006).

Table 1. SAT mean scores of college-bound seniors by race: 1999-2000 through 2010-2011

	99-00	00- 01	01- 02	02- 03	03- 04	04- 05	05-06*	06-07	07-08	08-09	09-10	10-11
SAT Composite Score												
White	1,058	1,060	1,060	1,063	1,059	1,068	1,582	1,579	1,583	1,581	1,580	1,579
Black	860	859	857	857	857	864	1,291	1,287	1,280	1,276	1,277	1,272

* Re-designed SAT with Writing Component (For all years starting 2005-2006)

Source: College Entrance Examination Board, College-Bound Seniors: Total Group Profile [National] Report, selected years, 1986-87 through 2010-11

Table 1 illustrates that the gap in combined scores from 1999 to 2010 remained relatively stable. When disaggregating the scores, the math, verbal, and writing sections all show significant differences in average scores, with each section having a consistent gap of approximately 100 points (*SAT Trends*, 2011).

In 2006, the SAT underwent a change in design. Analogies were removed and more critical reading passages were added, as well as a new writing component, with the intent of increasing the validity of SAT scores for predicting college success. Additionally, the new writing component (and the removal of analogies) was believed by many educational researchers to have the potential to improve scores for minority students. Their reasoning for this lay in the fact that the racial score gap on the SAT II writing exam – which has since been discontinued due to the addition of the essay component in the SAT I – was nearly half that of the verbal portion of the SAT I (Anonymous, 2001). However, the data present a different picture from what was hoped would be a movement towards a reduction in the SAT score gap. Although there are only six years of data, the new writing section closely mirrors the gap that is present in the other two sections. Black students fare significantly worse than their white counterparts on the essay section of the SAT, with a sizeable and statistically significant difference in average scores that follows the trend of the rest of the exam.

Explanations for the Score Gap

Researchers analyzing the SAT have put forth many explanations for the black-white score gap, ranging from issues with the design of the test, to a lack of equitable preparation for the exam. Even the Educational Testing Service (ETS), the company that produces the SAT, has offered evidence indicating that the black-white score gap is due to reasons not tied to innate

intelligence. The following section brings to light many of these explanations and illustrates how the SAT creates an uneven playing field for African-American students with regards to the admissions process at many of our nation's colleges and universities.

Test Bias

Among testing professionals, the results of a test are considered biased when problems with the test itself, or its administration, create differences in test scores between groups of test takers that are otherwise equivalent (AERA, APA & NCME, 1999). The results of a test would be considered biased if members of a specific race had statistically different scores when compared to other groups when the groups are equivalent with regard to the overall construct being measured. Take a trigonometry test, for example. If the test overestimated white students' ability to correctly answer trigonometry question and/or underestimated black students' ability, the test results would be racially biased. The Army Alpha contained flaws in both its content and administration that led to significantly lower scores for the black test-taking population.

Hereditarians of the time used these findings to further their agenda regarding innate differences in IQ. Much like the Army Alpha, the SAT, as recently as the early 1990's contained within it, similar, culturally biased questions that directly and negatively impacted the scores of black students. The most commonly referenced, culturally-biased question, asked test-takers to find the words analogous to "runner: marathon" (read: runner is to marathon as...). The answer to the question was "oarsman: regatta". The question was culturally loaded in favor of those students who live near bodies of water and are relatively wealthier. The statistics support this assertion: 53 percent of white test-takers answered this question, compared to only 22 percent of black test-takers (Chideya, 1995). This question, therefore, represented an item on the test that was biased against students who may not have lived near bodies of water (e.g., inner-city) or did not have

the financial capital to participate in these events.² This problematic content, often presented in analogies – which are no longer a component of the SAT – have since been removed from the SAT. However, the SAT remains biased in favor of white test-takers (as will be discussed below) thus resulting in a disproportionate disadvantage for the black population.

Whereas Yerkes, Brigham, and their colleagues overlooked many of the flaws regarding the Army Alpha, it must be made clear here, that this is not an argument that current psychometricians who are in charge of the design of the SAT intend to create a test that results in lower scores for African-Americans. Rather, their intent is to create the most reliable test whose scores validly predict college success (in terms of first-year GPA) for the population they serve. The argument being made is that, due to current practices in test construction, blacks are placed at a distinct disadvantage in the test even though there do exist practices that can lessen this disadvantage. In terms of admissions, African-Americans are considerably underrepresented at our nation's most selective universities. While blacks represent approximately 13 percent of the population in the United States, only 5.3 percent of college students at the nation's 146 most selective universities³ are black. By 2007, the percentages had barely shifted, with only just 5.5 percent of the students at these colleges being black (Mills, 2010). The data become especially salient when compared to Asian students. These students, while also being minorities, comprise a much larger percentage of seats at selective universities relative to the group's proportion of the entire U.S. population. Furthermore, 74 percent of students who attend these universities come from the highest SES quartile (Carnevale & Rose, 2003). This statistic highlights a disproportionate negative impact for African-American students owing to the fact that, as of

² As of 2010, 27.4 percent of Blacks lived in poverty, whereas only 9.9 percent of non-Hispanic Whites were in the same category. Additionally, as of 2010, 83 percent of the black, under-18 population lived in the nation's 100 largest metropolitan areas. (DeNavas-Walt, Proctor, & Smith, 2011).

³ Identified by Barron's Profiles of American Colleges" (Barron's Educational Series, 2006).

2010, nearly 40 percent of blacks under the age of 18 were living in poverty (DeNavas-Walt, Proctor & Smith, 2011).

In the construction of a standardized exam such as the SAT, an individual question is said to have item-bias if one group answers the question correctly at a significantly higher rate than that of another population when both populations are equivalent on the overall construct being tested. Referring to the Verbal section of the October 1998 SAT, the whites' average advantage of 13.4 percent, meaning that across all Verbal questions, whites answered each question correctly at a 13.4 percent higher rate than blacks (Kidder & Rosner, 2002). More than a third of the verbal questions gave whites an advantage of 15 percent or more. On the Math section, the advantage was 16.4 percent with more than 28 percent of the questions having an impact of 20 percent or more.

Clearly, these numbers indicate that many of the SAT questions provide a disproportionate advantage that “favors” white students over their black peers. On the surface, one might argue that this may simply reflect what the original IQ testers believed, that there exist innate differences between certain groups' intellectual capabilities. However, the method in which test questions are chosen accounts for a large amount of the differences in scores. When designing a test for a population, pre-test items that are answered correctly and have a high, positive correlation with higher overall scores will be selected for inclusion on the actual exam (as opposed to questions answered correctly but having low[er] correlations with high overall scores). As we know, white students, on average, score higher than black students on the SAT. Therefore, questions that whites answer correctly at higher rates than blacks will be selected for inclusion on future exams. The questions assessed were of equal difficulty (Kidder & Rosner, 2002). Questions that black students answer correctly at higher rates than white students end up

being excluded from the SAT in order to retain item-test reliability and validity of inference. Speaking in laymen's terms, the majority – white test-takers – win out if this method is followed.

While, statistically speaking, the exam remains unbiased, the actual consequence of eliminating questions that the black population answers correctly at higher rates results in a situation where blacks are less likely to answer questions correctly and therefore score lower on the test as a whole. Reducing black-white item-bias on test questions, however, does not have to weaken the reliability of the test or the validity of inferences drawn from the test. One such method, referred to as the Golden Rule method (Kidder & Rosner, 2002), which requires selecting questions of similar difficulty but with the smallest racial disparities possible, would obviously reduce the proportion of disadvantage on the test.

A central argument of those who oppose this method assumes that the test would then likely contain only questions of lesser difficulty (based on the premise that minorities score lower on average) and, as such, reduce the ability for the test to differentiate between students with greater ability and those with less ability to succeed in college. However, empirical evidence indicates that this is an incorrect assumption. In a study conducted by Freedle (2003), it was confirmed that black students (as well as other minority groups) fared better on more difficult questions on the verbal section of the SAT than their white peers, whereas white students outperformed blacks by a significant margin on easier questions. The Cultural Familiarity Hypothesis attempts to explain this phenomenon. The hypothesis argues that – on easier test items – more common vocabulary is used. Due to the common nature of the language, it is subject to interpretations and uses that are culturally bound. Because the predominant culture in the United States is that of the white majority, these easier questions will be understood and answered correctly at a higher rate by white students as opposed to black students (Freedle,

2003). On more difficult questions, however, rarer vocabulary is used. That vocabulary is less culturally bound, having fewer interpretations deemed acceptable by the SAT.

To quantify this, easy vocabulary terms used on the SAT had an average number of dictionary definitions of 5.2, while difficult terms had an average entry value of 2.0 (Freedle & Kostin, 1990). In summation, according to Freedle and his colleagues, black students fare better on more difficult questions due to the reduction in potential interpretations of the rarer vocabulary. Reducing the racial disparities on individual items would not weaken the ability of the test to differentiate between students. ETS even acknowledges this fact but chooses not to undertake practices mirroring the Golden Rule method: “First, such techniques can reduce impact [and] the average difficulty level of the resulting tests can be maintained without changing current test development procedures for adhering to average difficulty specifications” (Freedle, 2003). With this knowledge at hand, it is imprudent not to rectify the issue.

Income

Not to be lost in explanations for score gaps in the SAT is the association between test-takers’ scores and their socio-economic status. Within the educational world it is an unfortunate fact that with greater income also comes increased access to better educational services. Increased income also yields greater amounts of both human and social capital which impact individuals’ development and learning. These, in turn, influence scores on standardized tests. Quantifying this, students in the United States with the lowest SES stand a 4:1 chance of scoring in the lowest category on the Program for International Assessment (PISA). This ratio ranks the U.S. near the top of all developed countries with regards to the disadvantage that SES creates (OECD, 2006). Table 2 presents an analogous picture with scores on the SAT.

Table 2. SAT mean scores by income, 2011

Family Income	Critical Reading	Mathematics	Writing
\$0 - \$20,000	434	460	429
\$20,000–\$40,000	464	480	454
\$40,000–\$60,000	487	499	475
\$60,000–\$80,000	502	512	489
\$80,000–\$100,000	515	527	503
\$100,000–\$120,000	526	539	515
\$120,000–\$140,000	530	544	520
\$140,000–\$160,000	538	552	529
\$160,000–\$200,000	543	557	536
More than \$200,000	568	586	567

Source: College Entrance Examination Board, (2011). College-Bound Seniors: Total Group Profile [National].

When we look at black-white differences in income across the country, the negative impact of socio-economic status on black test takers becomes evident. As of 2010, large differences in income between black and white families continued to persist. The median income of black families in 2010 was \$32,068. The median income of their white counterparts was \$54,620 (DeNavas-Walt, Proctor, & Smith, 2011). A comparison of average SAT scores of black and white students from their respective median income groups reveals a score gap of 54 points.⁴ A gap of 54 points between students from median wage households across the two races already amounts to a substantial, half a standard deviation on the SAT. When we look across the income distribution, the effect of family income becomes even more pronounced between blacks and whites. In 2006, over half of the white, non-Hispanic, population in the United States earned over \$50,000 with 20 percent of the total population earning over \$100,000. Only 32 percent of African-Americans earned over 50,000 with less than 10 percent earning over \$100,000 (DeNavas-Walt, Proctor, & Smith, 2007). With these large and enduring differences in income, the black-white score gap will, likewise, persist.

⁴ This information underscores the fact, that holding constant income, black students, on average, have lower SAT scores. (Anonymous, 2009)

Academic Preparation

The SAT is designed to measure students' ability to succeed in college, quantified by first-year GPA. As such, the compilation of students' educational experiences leading up to the examination should play an instrumental role in the resulting score. Differences between black and white students in coursework and the correlations with SAT scores, verify this assumption. The new SAT format, noted earlier, illustrates this problem. Perhaps most disconcerting for the new exam with regard to the black-white score gap was the addition of more advanced mathematics.

In the current era, there is a push for students to be better prepared to compete in a global economy, with the roots for that preparation residing in education. Coursework is asked to be more rigorous with more advanced work in mathematics, and that is now mirrored in the newly designed SAT. However, black students still lag behind in advanced coursework in the subjects tested on the SAT. Among SAT test-takers in 2009, only 23 percent of black students had completed coursework in trigonometry (Anonymous, 2009). The number for their white peers stood at 37 percent. White students had also completed calculus coursework at just over twice the rate of black students, 30 percent to 14 percent, respectively.

The gaps present on the Math section of the SAT are no doubt due to lack of preparation in coursework that is specifically tested on the SAT (trigonometry) and also indirectly, by teaching higher-order thinking (calculus). When controlling for whether or not a student took trigonometry (after SES had been used as a control) prior to taking the SAT, the black-white score gap decreases by over 5.7 percentage points. Further controlling for calculus coursework indicates a reduction in the gap of an additional 5.7 percent (Nettles, Millet, & Ready, 2003).

The intersection between SAT scores and English coursework follows a similar trend. After having controlled for SES, the black-white score gap decreases by an additional 7.6 percent among those who have taken Honors English. These statistics clearly indicate that academic preparation impacts SAT scores. For a wide variety of reasons, including, but not limited to, disproportionate levels of blacks living in poverty (VanTassel-Baska & Gordon, 1987), the development of an oppositional subculture (Ogbu & Simons, 1998), and teacher quality (Hanushek, 2005), black students are less prepared academically to perform comparably to whites. Inequitable academic preparation negatively affects blacks' scores on all sections of the SAT.

Oppositional Subculture and Stereotype Threat

Culminating the many theories on the low performance of black students as a group on the SAT is the theory of the development of an oppositional subculture by the African-American population in the United States. The theory of an oppositional subculture is, perhaps, the most widely known cultural explanation for black underperformance in educational settings. Hypothesized by the late anthropologist and educational theorist John Ogbu, the development of an oppositional subculture by blacks in the U.S. links the history of black mistreatment, generally speaking and in educational contexts, to the subsequent outcome of their low academic performance.

Ogbu hypothesized that the history of African-Americans in the U.S., from their status as enslaved, involuntary minorities, to facing both *de jure* and *de facto* discrimination, has led to a black subculture that de-values all things “white” (Ogbu & Simon, 1998). Blacks in the U.S. have long faced discrimination in this country since before its inception as a sovereign nation.

From slavery, to segregation, to the long battle in the mid-20th century for civil rights, African-Americans have always faced barriers to equality. Ogbu argued that this historical mistreatment of the black population in the U.S. has led a large portion of the black population to develop a culture that opposes acceptance of any facet of the white-majority culture, and this extends to behaviors expected within classrooms and schools.

Ogbu argues that development of an oppositional subculture arises out of a need for a specific group to ensure its survival. This survival manifests itself in the creation of cultural bonds that directly oppose mainstream culture. Focusing on education, Ogbu claims that low black performance in school is a result of black culture rejecting high academic achievement due to high achievement being a cultural expectation of white culture (Ogbu, 2003). When high achievement is not fully rejected by black students, cultural pressures to conform (or else be alienated) further limit black student achievement (Fordham & Ogbu, 1986).

Involuntary minorities also hold very different views about the adult opportunity structure. In the United States, mainstream culture believes that, if individuals work hard enough, they can achieve their desired socioeconomic status. Blacks however, being involuntary minorities, perceive significant barriers to success that are external to their own attempts to improve their socioeconomic standing (Ogbu & Simons, 1998). Because education is integrally tied to improved social mobility, Black-Americans come to regard investment in education as unimportant and ineffectual in their own lives. High-effort and significant personal investment in education then become associated with white culture, and the opposite, low-effort and disengagement, is normalized by black students.

When looking specifically at the SAT, low-black performance can be viewed through the oppositional subculture lens. A major piece of oppositional subculture argues that beyond personal (and group) disinvestment, involuntary minorities internalize attitudes and perceptions from the majority about their lack of intellectual competence. Researchers argue that this internalization leads to ‘stereotype threat’, a psychological barrier to high academic performance present in educational testing (Steele & Aronson, 1995). Stereotype threat is an assertion that African-Americans have internalized beliefs of inferior intellectual ability impressed upon them by the culture of the majority through both personal and non-personal interactions (media, structure of institutions they must interact with, etc.).⁵ This internalization causes “inferiority anxiety”, a state in which students actually experience test anxiety due to negative perceptions about their ability to perform. Steele & Aronson (1995) carried out a study among Stanford students to confirm test this hypothesis. When cued for race by being told that they were going to be given a test measuring cognitive ability, these highly capable black Stanford students significantly underperformed when compared to their white classmates who were primed in the same manner. On the other hand, when told that the test was a “laboratory problem-solving task that was non-diagnostic of intellectual ability”, performance by black students did not have a statistically significant difference from that of white participants.

Qualitative evidence also supports claims that stereotype threat is salient to minority test-takers. Based on both group and individual interviews, it has been shown that black students are aware of racial issues surrounding the test. Students know that their white counterparts generally outperform them on the SAT, and are also aware of claims about the test having a white-

⁵ Stereotype threat is not exclusive to the African-American population. In a study assessing stereotype threat, researchers were able to induce the same phenomenon among white male test-takers when primed for race using Asian males as a comparison group. Aronson, J., Lustina, M.J., Good, C., Keough, K., Steele, C.M. & Brown, J. (1999). "When White Men Can't Do Math: Necessary and Sufficient Factors in Stereotype Threat". *Journal of Experimental Social Psychology*. 35(1), 29-46.

preference bias (Walpole et al., 2005). With the internalization of negative perceptions about test performance (and educational performance in general), theories about environmental and societal influences have merit and no doubt account for a portion of the persistent black-white score gap present on the SAT.

Importance of Attending a Selective Institution

There are a number of important reasons why attending a selective university is of great benefit to students. Selective universities, be it their reputation, financial resources, or collection of human and social capital, impart significant, and lasting, socioeconomic benefits to those fortunate enough to pass through their doors. Gaining admittance to a selective university also positively affects students' chance of graduating. It also positively impacts their future earnings and thus their social mobility (Hoxby, 2001).

Graduation Rate

It has long been known that students are more likely to complete their undergraduate degrees at selective colleges and universities when compared to non-selective institutions. At non-selective institutions 59.3 percent of students graduate within six years. The graduation rate is significantly higher at selective institutions with 80 percent of students graduating within the same time frame (Alon & Tienda, 2005). Some researchers claim that black students, who, on average, have lower SAT scores, would be more likely to graduate at non-selective institutions (Thernstrom & Thernstrom, 1997). Data, however, indicate that they graduate at much higher rates at selective institutions; 47.6 percent compared to 71.7 percent, respectively (Alon & Tienda, 2005). Therefore, attending a selective institution, through increased graduation rates,

operates as a mechanism for improved socioeconomic standing and mobility (Berg & Krueger, 2002).

Future Earnings

One of the driving forces behind high application rates at selective colleges and universities is the expectation that a degree from a more prestigious institution will yield significant economic benefits beyond that of a degree from a non-selective institution. Data indicate that this expectation is rooted in fact; obtaining a degree from a selective college or university does, on average, yield higher future earnings in the labor market. Using data on calculated hourly wage rates, Brewer, Ride and Ehrenberg (1999), found a significant wage premium for attending the nation's most selective institutions, with additional information indicating that the premium increases over time within a cohort.

When looking at the intersection between the selectivity of college attended and prior family income, the importance of attending a selective institution with regards to future earnings is also significant. Students from lower-income backgrounds reap greater economic benefits from attending more selective institutions (Berg & Krueger, 2003). This becomes especially salient for African-American students who, on average, come from more impoverished backgrounds than their white peers.

The SAT as a Barrier to Access to Higher Education

The prior discussion of explanations for the black-white gap in scores on the SAT, and the benefits of attending a selective institution, would be of little importance if the SAT were not still an integral component in admissions decisions at many of our nation's colleges and universities. Despite calls from ETS to limit the weight placed on the use of their test in

admissions decisions, and amidst claims from many admissions boards that the SAT isn't facing over-reliance issues, evidence that the SAT is still an extremely important factor in admissions speaks to the contrary.

The SAT stands as a single, standardized examination that allows admissions officials to “impartially” evaluate applicants. As such, it makes it easier to argue in favor of using the SAT as an integral component in admissions decisions (Durlauf, 2008). However, information indicates that the SAT does present a significant barrier to access to higher education for black college aspirants. When looking at the nation's most selective colleges and universities – those that admit fewer than 50 percent of its applicants – the SAT carries significant weight in the admissions process (Schmidt, 2008). Furthermore, the more selective a university (i.e., the lower the acceptance rate), the greater the weight placed on the SAT. Although other criteria – such as class rank – also gain in importance as the acceptance rate of a university decreases, the value placed on SAT scores increases at a faster rate than that of other factors in admissions as the acceptance rate as a university decreases. In other words, there is a larger correlation between weight placed on SAT scores and selectivity of an institution than with other admissions criteria (Alon & Tienda, 2007).

As of 2006, the percent of colleges attributing “considerable importance” to admissions test scores at our nation's selective universities sat at 57.1 percent (NACAC, 2008). Additionally, a representative sample of admissions officers at these universities who considered class rank an important factor declined from 42 percent in 1993 to 31 percent in 2005. These statistics become increasingly troublesome because as of 2006, approximately 70 percent of high schools in the country utilized GPA in their class ranks (Alon & Tienda, 2007). When coupled,

these factors create a perilous situation for African-American students aspiring to gain access to our institutions of higher education and have far reaching implications beyond college.

Chapter 3. SAT-Optional Admissions

To this point, the origins of the SAT, problems in its design with regard to bias, its (mis)uses, and the effect the exam has had on the African-American population regarding access to higher education have all been discussed in order to present a complete picture of how the SAT has created an unequal playing field. Education is the most reliable means towards upward social mobility, with higher education increasingly being regarded as a necessity for achieving middle-class status. As of 2011, there were 2,870 degree-granting institutions of higher education (NCES, 2010), and the majority of these continue to place weight on standardized test scores in their admissions process. In that year, approximately 70 percent of colleges and universities required the SAT or ACT as a prerequisite for admittance with over 650 of these organizations continuing to rank the SAT as the second most important factor in admissions just below grades achieved in college preparatory courses (NACAC, 2008).

In light of what has been discussed regarding the inequality that the SAT creates, postsecondary institutions' continued reliance on this test presents a bleak outlook for African-American students aspiring to acquire a post-secondary degree. There is, however, a movement towards SAT-optional admissions processes that has shown positive results for black college aspirants. The SAT-optional movement had its beginnings almost three decades ago when Bates College, a small, selective liberal arts college in Maine, decided to design an admissions procedure without the use of the SAT. Since 1984, Bates has used a more holistic approach to evaluating its applicants, placing more weight on Advanced Placement exam scores, class rank and GPA, as well as personal qualities discerned from the application essay (Hiss & Neupane, 2004).

The results from a 20-year study conducted at Bates College regarding its admissions process confirm that the SAT does not need to carry a significant portion of weight (if any) in admissions. Additionally, and central to this paper's concern, Bates College was able to significantly increase the African-American admittance rate without reducing the overall quality of its admitted students as measured by students' cumulative college GPA. Those who did not submit an SAT score had nearly identical grade point averages compared to submitters with a small and statistically insignificant GPA difference of .05; 3.06 to 3.11, respectively. Non-submitters also graduated at practically the same rate over the course of the 20 year study, with non-submitters having a graduation rate of 86.6 percent and submitters graduating at 86.7 percent.

The effect that discontinuing the SAT has had on African-American admittance rates is striking. Since the introduction of the SAT-optional admissions protocol, Bates College has seen the black application pool double with the corresponding admittance rate increasing by a comparable amount. The data clearly indicate that operating under an SAT-optional procedure affords black students a greater opportunity to attend the college. If the SAT had continued as a factor in admission this would likely not be the case. To evaluate this claim, the Bates study required that admitted students release their scores after acceptance. The data show that the SAT gap in scores between submitters and non-submitters (who did take the SAT) was, on average, 160 points; a significant score gap that would have eliminated many black students from gaining acceptance to the college and may even have deterred them from applying in the first place (Hiss & Neupane, 2004).

A similar case exists at Providence College, another selective, highly ranked institution. As of 2007, the college has employed an SAT-optional admissions process that has yielded

similar results. The total applicant pool at Providence College increased by 12 percent after the first year of the new policy. Accordingly, they saw a 17 percent increase in applicants of color and a 19 percent rise in enrollment from that subgroup (Robinson & Monks, 2004).

The SAT-optional movement is not without its critics; especially with the affirmative action challenges that have continually re-surfaced over the past several decades. In 1978, the Supreme Court, in the case of the *Regents of the University of California v. Bakke*, ruled out the use of racial quotas in admissions policies. The decision, however, did not outlaw the consideration of race in admissions. That particular decision came in 1996 when the court ruled, in the *Hopwood v. Texas* case appeal at the Fifth Circuit, that the University of Texas Law School could not use race as a factor in admissions. This decision helped set the precedent for other lawsuits to strike down affirmative action policies at institutions around the nation. The decision was detrimental to black enrollment rates at the law school. In the years following the Hopwood decision, black enrollment at the University of Texas Law School fell from 7.4 percent in 1995 to 0.9 percent in 1997 (MALDEF et al., 2005).

Proponents of affirmative action won a major victory in 2003 when the Supreme Court upheld the affirmative action admissions policy of the University of Michigan Law School in the case of *Grutter v. Bollinger*. In that case, a major factor in the decision to uphold the school's affirmative action policy was the presentation of a significant amount of social science research that highlighted the benefits of having a diverse student body (Buckner, 2004). This victory for affirmative action proponents might be short-lived, however, as the case of *Fisher v. Texas* is to be decided at the convening of the next Supreme Court (Docket no. 11-345). This case could potentially have wide reach. If it overturns the Grutter decision, *Fisher* could signal the end of affirmative action at our nation's colleges and universities. While that remains to be seen, to the

extent that race is considered in SAT-optional admissions policies, is whether SAT-optional policies could be endangered by the *Fisher v. Texas* ruling.

There have also been arguments made against the SAT-optional movement regarding the type of schools that can successfully implement this type of admissions policy. Bates and Providence College are similarly small, liberal arts colleges, located in one section of the country. Opponents argue that these schools, for a variety of reasons, including the relatively low ratio of applicant to admissions officer, allows for more holistic approaches with which to evaluate applicants. For larger institutions that have much larger applicant pools, opponents of the SAT-optional movement state that a more holistic approach cannot be effectively carried out (Epstein, 2009). For example, Towson's former Director of Admissions, Louise Shulack, used this as her office's stance for continuing to use the SAT in admissions, stated that "Because of the number of applicants and the variety of school systems from which we receive applications, the SAT is our best option at this point as far as standardized tests go" (Stone, 2002).

The traditional argument, that the SAT remains one of the best single predictors of college success continues to be made as well. This is made evident by the words of leaders of some of our nation's most prestigious universities such as John's Hopkins University (Stone, 2002). John Latting, serving as Director of Admissions at John's Hopkins (now working at Emory University under the same title), stated that "It's [the SAT] a good basic measurement in wide ranges...and the SAT can serve as a way to compare students who are otherwise so different."

Despite the continuing battle over the SAT-optional movement from some larger selective schools, there has been movement among others to adopt this process. In 2009, Wake

Forest became the first top 30 university to adopt an SAT-optional admissions policy with stated intentions to remove any unequal barriers to access to its institution for minority applicants. New York University, which enrolled 3,595 freshmen in 2009, soon followed suit in 2010 by not requiring the SAT if students submit a compilation of SAT II and AP exams. There are now 850 colleges and universities that employ an SAT-optional admissions policy (Grove, 2012). The majority of these institutions are smaller liberal arts colleges, but there are larger public universities that are joining the mix such as the University of Arizona, the University of Mississippi, and Kansas State University (De Vise, 2011). Time will tell if these institutions can succeed in increasing access for black students, but if the data from the smaller liberal arts colleges are any indication, the SAT-optional movement may be another step towards removing institutional barriers to higher education for African-Americans for those universities that have the ability to adopt such a policy.

Conclusion

For over half a century, the SAT has been used as a tool to select and sort students for places at our nation's colleges and universities. At its inception, the SAT's stated purpose was to identify students with exceptional intellectual ability in order to prepare them to lead a nation organized as a natural meritocracy (Lemann, 1999). The most intelligent, defined by the SAT, would gain admission to the nation's elite institutions of higher education in order to be trained for the most important occupations that would most benefit society. However, the creators of the SAT were strong believers in genetic differences intelligence across races; many of them eugenicists. The tests that were created were designed in a manner that reflected their beliefs, with the SAT – the grandchild of the Army Alpha – leading the way.

The SAT, like its predecessors, has a large and persistent gap in scores between black and white test takers. Factors such as test bias and unequal academic preparation are a few, among the myriad, causes for the sizeable gap in scores. The SAT is problematic with regard to validity, and it creates a disproportionate disadvantage for black test-takers. Combined with the considerable weight that the majority of our country's selective universities still place on SAT scores, access to higher education for African-American youths is considerably more difficult than it is for their white peers. Without equitable access to these institutions, the opportunities for upward social mobility for African-Americans are considerably diminished.

Looking forward, the current movement towards SAT-optional admissions policies being practiced by an ever-increasing number of colleges and universities shows promise. Case studies from a few of these universities indicate that this practice has proven beneficial for African-American college aspirants. The black applicant pool has increased dramatically at these schools

as well as their enrollment numbers. All this has occurred without admitting students who were under-qualified, as measured by their cumulative college GPA's and graduation rates, which have been on par with students who chose to submit their SAT scores. The success of the SAT-optional policies at these schools provides a framework for how to increase access to higher education for African-Americans. While this new practice is likely not a final solution to creating a truly equitable system of higher education, it presents itself as a step in the right direction towards providing greater access to higher education for African-American students as long as the score gap persists on the SAT.

References

- Anonymous. (2001). The shrinking black-white gap on SAT II achievement tests. *The Journal of Blacks in Higher Education*, 34, 57-61.
- Anonymous. (2006). A large black-white scoring gap persists on the SAT. *The Journal of Blacks in Higher Education*, 53, 72-76.
- Anonymous. (2009) The persisting racial chasm in scores on the SAT college entrance examination. *The Journal of Blacks in Higher Education*, 65, 84-89.
- Anonymous. (2009). Why family income differences don't explain the racial gap in SAT scores. *The Journal of Blacks in Higher Education*, 62, 10-12.
- Aber, J.L., Bennett, N.G., Conley, D.C., & Li, J. (1997). The effects of poverty on child health and development. *Annual Review of Public Health*, 18, 463-483.
- Alon, S., & Tienda, M. (2005). Assessing the "mismatch" hypothesis: Differences in college graduation rates by institutional selectivity. *Sociology of Education*, 78(4), 294-315.
- Alon, S., & Tienda, M. (2007). Diversity, opportunity, and the shifting meritocracy in higher education. *American Sociological Review*, 72(4), 487-511.
- Aronson, J., Lustina, M.J., Good, C., Keough, K., Steele, C.M., & Brown, J. (1999). When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology*. 35(1), 29-46.
- Bailey, R.C. (1997). Hereditarian scientific fallacies. *Genetica*, 99(2-3), 125-133.
- Barron's Educational Series, Inc (2006). *Barron's profiles of American colleges* (27 ed.). Hauppauge, NY: Barron's Educational Series.
- Berg, S.D., & Krueger, A. B. (2002). Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *Quarterly Journal of Economics*, 117(4), 1491-1527.
- Binet, A., & Simon, T., (2007) *A method of measuring the development of intelligence of young children*. (C.H. Town, Trans.). Whitefish, MT: Kessinger Publishing. (Original work published 1915).
- Binet, A., & Simon, T. (2010). *The development of intelligence in children: (the Binet - Simon scale)*. (E.S. Kite, Trans.). Whitefish, MT: Kessinger Publishing. (Original work published 1916).

- Boake, C. (2002). From the Binet-Simon to the Wechsler-Bellevue: Tracing the history of intelligence testing. *Journal of Clinical and Experimental Neuropsychology*, 24(3), 383.
- Brigham, C. (1923). *A study of American intelligence*. Princeton, NJ: Princeton University Press.
- Brewer, D.J., Eide, E.R., & Ehrenberg, R.G. (1999). Does it pay to attend an elite private college? Cross-cohort evidence on the effects of college type on earnings. *The Journal of Human Resources*, 34(1), 104-123.
- Buckner, C.J. (2004). Realizing Grutter v. Bollinger's "compelling educational benefits of diversity" – transforming aspirational rhetoric into experience. *University of Missouri-Kansas City Law Review*, 72(4), 877-947.
- Chideya, F. (1995). *Don't believe the hype: Fighting cultural misinformation about African-Americans*. New York, New York: Plume.
- Carnevale, A.P. & Rose, S.J. (2003). *Socioeconomic status, race/ethnicity, and selective college admissions*. New York, NY: The Century Foundation.
- The College Board. (2011). SAT Trends: Background on the SAT Takers in the class of 2011. Retrieved June 10, 2012, from:
http://professionals.collegeboard.com/profdownload/SAT_Trends_Report_9_12_2011.pdf
- College Entrance Examination Board. (2011). College-bound seniors: Total group profile [national] report. Retrieved June 23, 2012 from
http://professionals.collegeboard.com/profdownload/cbs2011_total_group_report.pdf
- De Vise, D. (2011, June 21). The most selective test-optional schools. *The Washington Post*. Retrieved June 5, 2012, from http://www.washingtonpost.com/blogs/college-inc/post/the-most-selective-test-optional-schools/2011/06/21/AGBB3heH_blog.html
- DeNavas-Walt, C., Proctor, B.D., Smith, J.C. (2007). *Income, Poverty, and Health Insurance Coverage in the United States: 2006*. Washington, DC: U.S. Government Printing Office. Retrieved June 15, 2012, from <http://www.census.gov/prod/2007pubs/p60-233.pdf>
- DeNavas-Walt, C., Proctor, B.D., Smith, J.C. (2011). *Income, Poverty, and Health Insurance Coverage in the United States: 2010*. Washington, DC: U.S. Government Printing Office. Retrieved June 15, 2012 from <http://www.census.gov/prod/2011pubs/p60-239.pdf>
- Durlauf, S. N. (2008). Affirmative action, meritocracy, and efficiency. *Politics, Philosophy & Economics*, 7(2), 131-158.
- Education at a Glance: OECD Indicators (2006). Paris: Organisation for Economic Co-operation and Development.

- Epstein, J. P. (2009). Behind the SAT-optional movement: Context and controversy. *Journal of College Admission*, (204), 8-19.
- Fordham, S., & Ogbu, J.U., (1986). Black students' school success: Coping with the "Burden of 'Acting White'". *The Urban Review*, 18(3), 176-206.
- Freedle, R. O. (2003). Correcting the SATs ethnic and social-class bias: A method for reestimating SAT scores. *Harvard Educational Review*, 73(1), 1-43.
- Gould, S.J. (1996). *The mismeasure of man*. New York, NY 10110: W. W. Norton & Company, Inc.
- Grove, A. (2012). 850 colleges don't require SAT or ACT scores. Retrieved July 5, 2012 from <http://collegeapps.about.com/b/2012/02/17/now-850-test-optional-colleges.htm>
- Hanushek, E. A. (2005). The economics of school quality. *German Economic Review*, 6, 269–286.
- Hiss, W. C., & Neupane, P. R. (2004, October). 20 years of optional SATs at Bates. Paper presented at the annual meeting of the National Association for College Admission Counseling, Milwaukee, WI.
- Hoxby, C.M. (2001). The return to attending a more selective college: 1960 to the Present. *Forum Futures 2000*, 3, 12-42.
- Joint Committee on Standards for Educational and Psychological Testing of the AERA, APA, and NCME. (1999). *Standards for educational and psychological testing*. Washington DC: American Educational Research Association.
- Kidder, W.C., & Rosner, J. (2002). How the SAT creates "built-in headwinds": An educational and legal analysis of disparate impact. *Santa Clara Law Review*, 43, 131-211.
- Kreider, Rose M. and Renee Ellis. (2011). Living Arrangements of Children: 2009. *Current population reports*, (P70-126), U.S. Census Bureau, Washington, DC.
- Lemann, N. (1999). *The big test: The secret history of the American meritocracy*. New York, NY: Farrar, Straus and Giroux.
- Lippmann, W. (1976) Readings from the Lippmann-Terman debate. Block, N.J., & Dworkin, G. (Eds.). *The IQ controversy: Critical readings*. New York: Pantheon.
- Mexican-American Legal Defense and Educational Fund (MALDEF), Americans for a Fair Chance, Equal Justice Society, and The Society of American Law Teachers. (2005). Blend it, don't end it: Affirmative action and the Texas ten percent plan after Grutter and Gratz. *Harvard Latino Law Review*, 8, 33-92.

- Mills, M. (2010). Five years post-Grutter, little progress in Black and Latino/a enrollments at selective colleges and universities. *Journal of College Admission*, (208), 6-7.
- National Association for College Admission Counseling. (2008). NACAC 2008 annual report. Arlington: VA.
- National Institute for Literacy. (1994). *Adult literacy in the United States: A compendium of qualitative data and interpretive comments*. Washington, D.C.: Sticht, T.G., & Armstrong, W.B.
- Nettles, M.T., Millet, C.M., Ready, D.D., Ludwig, J., & Foreman, J. (2003). Attacking the African American-White achievement gap on college admissions tests. *Brookings Papers on Education Policy*, 6, 215-252.
- Ogbu, J.U., (2003). *Black American students in an affluent suburb: A study of academic disengagement*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Ogbu, J.U., & Simons, H.D. (1998). Voluntary and involuntary minorities: A cultural-ecological theory of school performance with some implications for education. *Anthropology & Education Quarterly*, 29(2), 155-188.
- Schmidt, P. (2008). Researchers accuse selective colleges of giving admissions tests too much weight. *The Chronicle of Higher Education*, 54(35), 20-21.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African-Americans. *Journal of Personality and Social Psychology*, 69(5), 797.
- Stone, A. (2002, Feb 08). SAT controversy: Despite on-going criticism, local colleges continue to use this standard admissions test. *Baltimore Jewish Times*, 264(6), p. 44.
- Terman, L.M. (1920). The use of intelligence tests in the grading of school children. *The Journal of Educational Research*, 1(1), 20-32.
- Terman, L.M., Otis, A.S., Dickson, V., Hubbard, O. S., Norton, J. K., Howard, L., Flanders, J. K., & Cassingham, C. C. (1917). A trial of mental and pedagogical tests in a civil service examination for policemen and firemen. *Journal of Applied Psychology*, 1(1), 17-29.
- Thernstrom, S., & Thernstrom, A. (1997). *America in black and white: One nation indivisible*. New York: Simon & Schuster.
- U.S. Department of Education, National Center for Education Statistics. (2010). Education directory, colleges and universities: "Institutional characteristics survey"(IPEDS-IC: 86-99). Washington, DC: U.S. Government Printing Office. Retrieved from http://nces.ed.gov/pubs2011/2011015_3a.pdf

VanTassel-Baska, J., & Gordon, W. (1987). A three year study of the effects of low income on SAT scores among the academically able. *The Gifted Child Quarterly*, 31(4), 169.

Walpole, M., McDonough, P.M., Bauer, C.J., Gibson, C., Kanyi, K., & Toliver, R. (2005). This test is unfair: Urban African American and Latino high school students' perceptions of standardized college admission tests. *Urban Education*, 40(3), 321-349.