College Admission Testing in America

Brent J. Evans

Introduction

There is rarely a more anxiety inducing event for a high school student than sitting down to take a college admission test. Scores on the SAT or ACT, the two nationally recognized college admission tests in the United States, are an important component of the admission application to colleges and universities and help the admission officers decide whether or not to accept the student into the incoming class. Most of the selective colleges and universities require high scores for acceptance, and many competitive scholarships use admission test scores as one of several selection criteria. For athletes, obtaining at least the minimum score established by the National Collegiate Athletic Association (NCAA) is essential to be eligible to play inter-collegiate sports. To gain a sense of how important students and their families perceive these tests to be, consider the fact that an estimated 1.5 million students purchase test preparation and tutoring services for the SAT, collectively spending about $530 million annually (Marte, 2011).

This chapter presents an overview of the SAT and ACT. After describing the two exams, it provides a brief history of their development before explaining why college admission offices use testing to make admission decisions. It then summarizes the scholarly research on the predictive validity of the exams in predicting college outcomes. Evidence from interviews with college admission officers explains the extent to which admission test scores factor into the admission decision. Finally, the chapter offers some criticisms aimed at the SAT and ACT and presents alternatives to the exams.

Current Format of the SAT and ACT

The exams are only offered six to seven times each year, and students preparing for college often take either one or both of the exams in the spring of their junior year or in the fall of their senior year. They must register in advance and pay a registration fee of $50 for the SAT and $35 for the ACT, although fee waivers exist for financially disadvantaged students. Students can take either exam multiple times, and colleges typically use the highest scores across multiple exams to evaluate students' admission applications.

Although either exam fulfills the admission requirements of colleges and universities, the SAT and ACT differ in important ways. The SAT contains separate tests in mathematics, critical reading, and writing each of which is scored on a scale of 200-800 such that a perfect score is 2400. The exam's length is three hours and forty-five minutes. In contrast, the ACT has four separate subject tests: English, mathematics, reading, and science and lasts three hours and twenty-five minutes. Each subject test has a minimum score of 1 and maximum score of 36, and the exam provides a composite score which is the average of the four subject tests rounded to the nearest whole number. Both exams rely predominantly on multiple choice questions; however, the SAT penalizes students' scores for incorrect answers to discourage guessing.

Both organizations that administer the exams (the College Board for the SAT and ACT Inc. for the ACT) offer additional, optional exams.[1] ACT Inc. offers an optional writing section which students take at the end of their regular ACT exam. The College Board offers a host of SAT II subject exams which students register for and take separately. They test specific subject knowledge ranging from chemistry to French. Some college admission offices require or recommend that students submit scores from two or three SAT II subject tests to provide a more complete picture of their ability.

A Brief History of College Admission Testing in the United States

The history of standardized college admission tests begins with the development of the IQ test in the early 20th century by Alfred Binet (Fletcher, 2009). The military quickly adapted intelligence testing to assess the abilities of incoming soldiers in an effort to choose officers, and this wide scale use laid the groundwork for applying testing more widely in education. Although the College Board first established the SAT in 1926, it took until the late 1930's before a group of elite northeastern colleges, the members of the College Board, agreed to use the test as a common admission metric (Lemann, 2004). As Lemann describes, the goal of implementing the SAT was to move away from the standard practice of assessment based on traditional knowledge developed at private high schools and replace it with a test that identifies the most intellectually able students. Due to this philosophy, the test is explicitly designed not to be an achievement test that measures knowledge learned in the traditional high school curriculum.

The ACT grew out of a high school academic competition in Iowa meant to highlight the top academic talent in the state (ACT, 2009). Everett Lindquist, a professor of education at the University of Iowa, worked extensively to develop statistical tools and automatic grading devices for use in subsequent iterations of the initial academic competition tests for primary and secondary schools. He collaborated with Ted McCarrel, the director of admissions at the University of Iowa, to apply those concepts in the development of a standardized college admission test. Thus the ACT emerged in 1959 and served two purposes. The first was to establish a common admission test that could be used across the nation as an assessment of a student's preparation for college. It was intended as an achievement test, measuring knowledge of the curriculum, unlike the SAT, and one exam could replace the disparate admission tests that were offered by individual institutions and states. In this effort, the ACT competed with the SAT, but the SAT's focus was predominately on selective institutions in the northeast. The second purpose was to inform students, by asking questions about their interests, about which institution to attend and which program to study.

Throughout the second half of the twentieth century, the exams became firmly established as a critical element of the admission process across all of the four-year institutions of higher education, and both exams expanded dramatically thanks to the massification of higher education which increased college enrollment nationwide. For the graduating high school class of 2011, about 1.65 million students nationwide took the SAT (Adams, 2011) and 1.62 million took the ACT (ACT, 2011). Despite their near equivalence in popularity, their

[1] Although the College Board owns and administers the SAT, Educational Testing Service (ETS) writes the exam questions and scores the exam.

regional nature remains apparent with the SAT maintaining popularity in the Northeast and on the West coast and the ACT dominating much of the Midwest and South.

Why Do Admission Offices Rely on Testing?

To fully answer this question, it is important to understand the purpose of the admission process. In their seminal work on race and the admission process, *The Shape of the River*, Bowen and Bok (1998) describe the goals of admission officers. They first and foremost focus on admitting students that will succeed academically at the institution. It benefits neither the institution nor the students to admit an applicant that is not likely to persist. At selective colleges where there are far more academically capable applicants than can be admitted, the admission office chooses applicants who fall into one of four categories: applicants who are academically exceptional; those that have a diverse background, experience, or talent; students likely to contribute valuably to society or their profession; and applicants that have a strong loyalty to the institution. As a former admission officer, I distinctly remember examples from each of these categories. Whether it was a national science fair winner, exceptional harp player, or the daughter of a faculty member, admitted students usually had a distinguishing characteristic in one of these four categories in their application. Although testing is but one component of the admission application, it serves to ensure that students fulfill the very first requirement of being able to succeed at the institution to which they applied. Admission officers tend to agree with this assessment of the use of testing in evaluation process.

In 2009, I interviewed admission officers at all thirteen of the public universities in Ohio. This group of diverse institutions constitutes a wide range of admission practices from very selective to open enrollment, but all of them require submission of the SAT or ACT.[2] When asked why their institutions rely on test scores from the SAT or ACT as part of the application evaluation process, they provide the same two answers almost without exception. First, the exams provide a common metric upon which to evaluate students with different backgrounds and high school preparation. Due to the high variance of the quality of secondary education, it is extremely difficult for an admission officer to judge the level of rigor of a high school curriculum by examining a student's transcript. The exam score therefore functions as a way to compare students on the same scale not only to other current applicants but to past applicants as well. It serves to identify students that may be underachievers in high school but have high potential for succeeding in college and students that may have received great grades in high school but not be well prepared for the additional rigors of the college curriculum.

The second major reason admission offices rely on test scores is that the tests predict how well students will do in college. This concept is called predictive validity, and the majority of studies on the SAT and ACT focus exactly on this issue. As discussed above, the goal of the admission process is to accept students who will succeed at the institution academically, and there is widespread belief that using test scores will help the admission staff select students who are likely to succeed at their institution in terms of receiving high grades and persisting from year to year. Next, I turn to the large body of research that supports this claim.

[2] The ACT is the dominant admission test for Ohio students, so the majority of applicants to these institutions reported ACT scores instead of SAT scores.

Predictive Validity Research

The general goal of predictive validity is to determine how much a current measure of performance predicts future performance. For the purposes of test scores and college admission, the critical underlying assumption is that there is an overlap in students' ability on test performance and their college outcomes such as grades and retention (Kane, 2006). The predictive validity of an exam is commonly measured by finding the correlation coefficient between the exam score and the college outcome of interest, and most of the research on predictive validity reports findings in this manner.

Although preceded by a small number of SAT studies most of which are unreliable because they rely data drawn from only one college, the first major analysis of the predictive power of the ACT exam is provided by Munday (1967). Using information from 398 colleges across the nation collected by ACT Inc. on test scores and students' self reported most recent high school grades in each of the four subject areas on the ACT at the time (English, math, science, and social studies), he shows the correlation between high school and college grades is improved by using high school grades together with ACT scores to predict college grades in the four subject areas tested by the ACT. Munday's paper sets the general pattern for much of the subsequent research on the predictive power of standardized tests. Variations occur with the outcome variables, which include individual college course grades, freshman college GPA, overall GPA, and graduation probability.

Burton and Ramist (2001) report results from a recent meta-analysis of studies predicting cumulative undergraduate GPA and graduation probabilities using SAT scores and high school grades. They use a weighted average of results from studies consisting of students that graduated from college between 1980 and the mid 1990's. The findings indicate that the average correlation of verbal and math SAT scores with cumulative college GPA is 0.36 compared to a correlation of 0.42 for high school grades with college GPA. A smaller sample of studies use verbal and math SAT scores combined with high school grades to predict college GPA, and their weighted average correlation is 0.52. Because of the higher correlation when using both high school grades and test scores, the results of this meta-analysis support the testing agencies' assertion that standardized test scores should be used together with high school records to predict college success.

The correlation for graduation outcomes is lower because of the many years between taking the exam and potentially graduating from college. Burton and Ramist use eight studies of correlations between SAT scores and college graduation to calculate a 0.33 correlation. Unfortunately, the limited number of studies conducting on graduation rates limits the reliability and generalizability of these findings. Still, it appears that testing does predict college graduation.

To present these outcomes in more interpretable manner, we can examine the findings of Bowen and Bok's (1998) work examining the relationship between SAT scores and college outcomes. They base their analysis on a dataset of students in the mid-1990's who attended 28 selective liberal arts colleges and research universities (the College and Beyond database constructed by the Mellon Foundation). They determine a positive relationship between SAT scores and college class rank based on cumulative GPA and conclude that an increase of 100 points on the combined SAT score is associated with a 5.9 percentile point increase in college class rank. This controls for race, gender, socioeconomic status, high school grades, college major, and college selectivity. Regarding the

predictive validity of the test on graduation outcomes, they note a small positive correlation that disappears when considering students above scores of 1100. This implies standardized tests may not have as much validity once students achieve a certain level of aptitude, but Bowen and Bok's study is hampered by its extremely selective sample of prestigious institutions.

Although the previous discussion concentrates on the SAT, which dominates the literature, comparable findings exist for ACT scores. In the literature review of their study, Noble and Sawyer (2002) describe research sponsored by ACT Inc. indicating the median correlation across 129 colleges between the current four ACT subject scores (English, math, reading, and science) and college freshman GPA is 0.43. By comparison, the correlation between high school grades and freshman GPA is 0.48 and grows to 0.53 when using both high school grades and ACT scores. These results are very similar to the SAT results.

Research consistently demonstrates the predictive validity of the SAT and ACT, but looking at the single total score is not the entire story. In research I have conducted with two colleagues, we find that the ACT subject tests are not equally predictive of college outcomes. Scores on the English and math exams predict college grades and the likelihood of dropping out of college, but the reading and science exams offer no additional predictive validity (Bettinger et al, forthcoming). Instead of using the composite score which simply averages the four subject scores together, admission offices could improve their determination of which students are likely to succeed in higher education by weighting the English and math scores more heavily.

Despite the carefully constructed scholarly work of most predictive validity studies, almost all of them suffer from a flaw. The goal of the admission office is to apply the predictive validity of test scores to the applicant pool in order to gauge how successful each applicant will be at the institution. However, a problem arises because studies estimating the predictive validity must rely solely on observable outcomes (freshman GPA, graduation, etc.) only available for enrolled students, which are likely different than the unobserved outcomes of all applicants. Two stages of selection occur separating the applicants from the eventual enrollees. First, the admission office decides if each applicant is offered admission, and second, the admitted student decides whether to enroll.

The selection problem posed by the first form of selection is attributable to the information differential between the admission office and researcher. Many factors affecting the admission decision for an individual student are unobservable to the researcher. Application materials such as letters of recommendation, extracurricular activities, and essays present a problem of selection on unobservables. The researcher cannot account for all differences in applicants when estimating the correlation between test scores and outcomes, so the predictive validity estimates will be biased.

The second form of selection relates to the student's decision to enroll conditional on receiving an offer of admission. The probability of an individual student choosing to enroll is determined by a number of observable factors such as the financial aid award and legacy status and countless unobservable factors such as the prestige of alternative colleges to which the student has been admitted. These factors are highly likely to produce differences between the students who choose to enroll and those who do not, further biasing the predictive validity conclusions when they are estimated using only matriculated student data.

Rothstein's (2004) study of the two stage selection procedure used in the University of California's (UC) admission system provides a solution to the first selection issue. In the first stage of the UC application, a central office determines eligibility for admission to the system based solely on high school GPA and SAT I scores which are completely observed by the research. This strategy helps to isolate the predictive power of the SAT on first year college GPA while avoiding some of the selection bias. His methods provide more plausibly valid estimates that the SAT's predictive power is approximately twenty five percent lower than the biased correlations typically reported.

Even after correcting for some of the selection issues, it remains clear that the predictive validity studies confirm that college admission testing does predict future college performance. They also support the assertion that predictive validity works best when used in conjunction with high school grades. How much weight test scores should be given in the admission process remains an open question, which is discussed below.

How Do Admission Offices Use Test Scores?

Having seen why admission offices employ test scores, it is important to consider how they use students' performance on the exams. Admission tests serve many distinct purposes. The most universal and widely known purpose is using the scores for general admission as part of the college application, but admission offices can also use the exam scores for admission into a specific program of study. Although outside the realm of this chapter, some institutions use admission test scores to place students into freshman classes and/or as an eligibility criterion for merit-based financial aid.

Most, but not all, institutions requiring exam scores use the exams for the purpose of general admission. There are, however, a few open enrollment institutions that require the submission of test scores but do not deny admission to any applicant. For the majority that do use test scores for admission, an important question is how much weight is placed on tests scores relative to other factors in the application. I present two sources of information in an attempt to answer this question.

The first is taken from a national survey of American four-year public and private colleges conducted by the National Association for College Admission Counseling (NACAC) (Clinedinst et al., 2011). NACAC annually conducts this survey, and the 2010 survey results reported here represent 330 responses, about 17% of the not-for-profit, degree granting four-year institutions in the United States. The survey asks admission offices to rate each of 16 factors used in the admission decision as considerable, moderate, limited, or no importance. Admission test scores finish as the third most important criterion with 59.3% of institutions labeling them as considerably important in the admission decision. Test scores are behind grades in college preparatory high school courses (83.4%) and strength of high school curriculum (65.7%). For comparison, 26.6% of institutions rate essays or writing samples as considerably important, and 19.0% rate teachers' recommendations considerably important.

Breaking down the analysis by institutional control, the survey shows that public institutions rate testing as considerably important more than private institutions: 73.2% compared to 53.8% respectively. Comparing the most selective colleges (as defined by those that accept less than 50% of applicants) to the least selective (those accepting more than 85% of applicants) demonstrates that testing is a slightly less important admission factor at the more selective colleges with 54.0% rating it considerably important relative to the 57.8% of nonselective colleges,

although this can probably be explained by the tremendous amount of self selection that takes place as typically only top students with high test scores apply to selective colleges. NACAC has conducted this survey since 1993, and standardized admission tests have consistently finished third in importance; however, their significance has grown over time as only 46% of surveyed colleges in 1993 claimed testing was considerably important.

The second source of evidence that provides insight into the important test scores in the college admission decision derives from my previously mentioned interviews of public university admission offices in Ohio. Instead of asking admission officials to rate different factors on a four point scale of importance, I simply asked what percentage of the admission decision is based on test scores for an individual applicant. Excluding the open enrollment institutions which accept everyone, the responses varied from 33% to 50% of the admission decision. This means one third to one half of the decision rests on an applicant's SAT or ACT score. The interviews confirm NACAC's findings that the high school curriculum and grades are the other most important components of the application.

A few of the interviewed universities explicitly demand minimum test scores for admission. For example, one institution requires an ACT score of 21 to be further considered for general admission. More common is a policy that requires a minimum ACT score for admission into a specific program such as engineering or architecture, and these minimum scores may only focus on one subject of the exam. For example, several admission offices require minimum math ACT scores for entry into the undergraduate business or engineering majors.

Criticisms of Standardized Admission Tests

Due to their importance in the admission process and to the increasingly competitive admission standards at selective colleges, standardized admission tests have received intense scrutiny and garnered substantial criticism. This section summarizes the arguments against using the SAT and ACT in the admission process by focusing on racial and financial inequalities in test performance.

Racial inequalities in SAT and ACT scores have been acknowledged for several decades. In 2011, the average scores for white students on the critical reading and mathematics components of the SAT were 528 and 535 respectively, but the scores for black students were at least 100 points lower on each test, 428 and 427 (College Board, 2012). A similar pattern holds for the ACT with a 2011 average composite score of 22.4 for white students and 17.0 for black students (ACT, 2012). Given these large differences, it is not surprising that studies have long shown admission policies demanding equal testing performance for white and black students will disadvantage black applicants (Crouse and Trusheim, 1988).

Researchers have offered several potential explanations for racial differences in admission test scores. The well documented educational inequalities in primary and secondary education between black and white students are certainly one cause of the black white test gap. For example, Card and Rothstein (2007) show that racial segregation in cities, neighborhoods, and schools can explain about a quarter of the test score gap. This is the explanation embraced by the makers of the SAT and ACT; however, two other explanations also deserve consideration: test bias and stereotype threat.

Jenks (1998) deconstructs five different types of test bias and applies them to the SAT. For instance, he argues there is a labeling bias associated with the SAT. Because it is a so called aptitude test that is not directly related to the high school curriculum, he suggests that the exam motivates the false perception that black students have lower aptitude for college. Although that labeling bias does not explain the racial test score gap, the content bias as described by Santelices and Wilson (2010) may contribute to it. By examining the differential item functioning of same-ability black and white students on SAT takers from California, the two researchers demonstrate that the racial subgroups perform differently on different items on the verbal section of the SAT.[3] White students perform better than black students on the easier questions while black students outperform white students on the more challenging verbal questions. The authors find neither any difference on math performance nor any effect on Hispanic students. It seems clear that differential item functioning exists; however, it appears equally clear that it cannot explain the black white gap in test scores.

Another potential explanation for the racial test score gap is stereotype threat, a concept initially outlined by Claude Steele. The idea is that if a group is stereotyped to perform poorly, individuals associating themselves with that group and aware of that stereotype will actually perform worse because of it. Stereotype threat has been clearly documented in a variety of circumstances for many groups including gender and racial groups. Using a laboratory experiment, Steele and Aronson (1998) show that when black students answer questions similar to those on the SAT, they perform worse when they are told the test measures their ability. The experimenters also demonstrate that black students perform worse on verbal questions when they are primed to consider their race even when they do not present the test as a measure of ability. Due to these findings, it is feasible that stereotype threat contributes to the black white score gap.

The other systematic inequality in SAT and ACT performance is financial. Although some critics complain that even registering for the test costs money, fee waivers for low-income students resolves that issue in most circumstances. The larger concern is that test scores are highly correlated with family income and wealth. For example, the average composite score in the ACT in 2005 for students from families whose income is $18,000 or less was 17.9 while students from families whose income was above $100,000 scored 23.5 on average (Jaschik, 2005). Similarly large gaps exist for the SAT.

What explains these large differences? Some critics blame test preparation services, often called coaching. These are usually commercial services that charge students for test tutoring, and some companies claim large gains of between 100-200 points for students that make use of their services. Although somewhat dated, the most recent data I can find on the demographic characteristics of students who pay for standardized test coaching show that students with family incomes of $80,000 or more are almost twice as likely as students with family incomes of $40,000 or less to receive coaching (Powers and Rock, 1999). It seems obvious that wealthier students are more likely to purchase coaching services, but the estimated effect of coaching is substantially smaller than what the companies claim. Relying on observation studies and an array of statistical techniques such as matching, Powers and Rock show coaching has a small effect of only 10-20 points on SAT math scores with even smaller effects on

[3] See Santelices and Wilson for a literature review of the extensive prior research on differential item functioning on the SAT.

verbal scores. Briggs (2004) reaches similar conclusions that coaching does not have large effects and explains the difference between the scholarly findings and the companies' assertions. The commercially publicized claims simply report the change of the coached students from their pre-coaching exam scores to their post-coaching exam scores without using any comparison group. However, most students improve their scores when retaking the exam regardless of whether they received coaching, so the real effect of coaching is the much smaller difference in the change in scores between students who did and did not receive coaching.

After examining evidence from a variety of different standardized tests, Zwick (2004) also concludes that formal test preparation is not the root cause of the disparities in test scores between low- and high-income families. Instead, she points to the inequalities in home and school environments as the most likely cause. Parents from low-income households are less likely to read to children than high-income families, and wealthier schools provide greater opportunities for students through higher quality instruction and better resources. Although we may not know the exact extent of each contribution, all of these explanations likely factor into the financial inequality observed in standardized admission test scores.

Alternatives

The criticisms leveled at the SAT and ACT ensure a robust debate about whether alternatives exist. Various scholars, practitioners, and policymakers have all pushed for implementing alternatives to the standard admission test, but consensus has not developed.

Some feel a stronger reliance on achievement tests directly related to the high school curriculum is warranted. Lemann (1999) makes a case that the purpose of an admission test should not be to simply sort students into those that should attend a selective college and those that should not based on their perceived aptitude as defined by the SAT score. Instead, he argues the admission test should be directly aligned with the high school curriculum so that students are judged on their mastery of the material necessary in college courses.

The logical implication of that argument means relying on Advanced Placement (AP) exams and SAT II subject tests as a stronger basis for making college admission decisions. The SAT II subject exams test students on material directly related to a specific subject studied in high school. AP exams follow a yearlong high level course with a nationally standardized curriculum developed by the College Board and offered at many high schools. The International Baccalaureate (IB) program is similar in concept with a college level exam taken at the end of the year. Many colleges grant college credit for high scores on both AP and IB exams taken in high school indicating that they believe those classes are equivalent to college level coursework. The leading problem with using AP and IB exam scores in the admission process is that universities must make an admission decision by the end of March, long before any scores for exams taken at the end of the senior year are available.

The other major alternative to the current reliance on the SAT and ACT is not to use any standardized admission test scores at all. Perez (2004) advocates this position claiming that almost 400 institutions of higher education have dropped the requirement to submit test scores for at least some portion of the applicant pool. For instance, Franklin and Marshall College in Pennsylvania does not consider test scores if an applicant is in the top ten percent of his or her high school graduating class. Bates College in Maine famously eliminated all test score

requirement in 1990. Under its optional test score policy, they found no significant difference between the high school GPAs of applicants who did and did not submit test scores, and they believe the change increased application numbers and minority enrollment (Hiss, 2001).

The obvious downside of this approach is that it increases reliance on other components of the application. This might not be a large concern for liberal arts colleges which have smaller applicant pools and evaluate each applicant holistically by reviewing letters of recommendation, essays, and sometimes interviews with the applicant. It is no surprise that the majority of colleges eliminating testing requirements are smaller liberal arts colleges. Larger universities rely more heavily on admission tests to sort through enormous applicant pools. Reading every applicant holistically requires an investment of resources that many institutions cannot afford, so relying on a GPA and test score cutoff, as the University of California system does, is an efficient, but potentially inexact, way to cull a large number of applicants down to a more practical size.

Conclusion

This chapter has served as a primer for college admission testing in America covering their history, purpose, validity, faults, and alternatives. The ACT and SAT play a significant role in sorting high school graduates into universities across the country. Despite criticisms of the exams receiving a good deal of attention, most admission offices continue to require applicants to submit test scores because of their predictive validity, and society continues to accept their value in helping admission officers select students for college. Even though alternatives may grow in popularity and changes to the current exams may slightly alter their composition, it seems clear that standardized testing will continue to play an important factor in college admission for the foreseeable future.

References

ACT, Inc. 2009. *ACT: The First Fifty Years, 1959-2009*. Iowa City, IA: ACT, Inc.

ACT, Inc. 2011. *The Condition of College & Career Readiness*. Iowa City, IA: ACT, Inc.

ACT, Inc. 2012. *National Score Trends*. http://www.act.org/newsroom/data/2011/trends-text.html#nine

Adams, C.J. 2011. SAT scores fall as number of test-takers rises. *Education Week*. September 14, 2011. http://www.edweek.org/ew/articles/2011/09/14/04sat.h31.html

Bettinger, E.P., Evans, B.J., & Pope, D.G. Forthcoming. Improving college performance and retention the easy way: Unpacking the ACT exam. *American Economic Journal: Economic Policy*.

Bowen, W.G. & Bok, D. 1998. *The Shape of the River: Long-term Consequences of Considering Race in College and University Admissions*. Princeton, NJ: Princeton University Press.

Briggs, D.C. 2004. Evaluating SAT coaching: Gains, effects and self-selection. In *Rethinking the SAT: The Future of Standardized Testing in University Admissions*, (ed. Rebecca Zwick). pp. 217-233. New York: RoutledgeFlamer.

Burton, N. W. and L. Ramist. 2001. *Predicting success in college: SAT studies of classes graduating since 1980.* New York: College Board.

Card, D. & Rothstein, J. 2007. Racial segregation and the black-white test score gap. *Journal of Public Economics* 91:2158-2184.

Clinedinst, M.E., Hurley, S.F., & Hawkins, D.A. 2011. *2011 State of College Admission*. Arlington, VA: National Association for College Admission Counseling.

The College Board. 2012. *2011 SAT Trends*. http://professionals.collegeboard.com/data-reports-research/sat/cb-seniors-2011/tables

Crouse, J. & Trusheim, D. 1988. *The Case Against the SAT*. Chicago, IL: The University of Chicago Press.

Fletcher, D. 2009. Standardized testing. *Time*. Dec 11, 2009. http://www.time.com/time/nation/article/0,8599,1947019,00.html

Hiss, W.C. 2001. Optional SAT's at Bates: 17 Years and Not Counting. *The Chronicle of Higher Education*. October 26, 2001. http://chronicle.com/article/Optional-SATs-at-Bates-17/9274/

Jaschik, S. 2005. ACT scores are level. *Inside Higher Ed*. August 17, 2005. http://www.insidehighered.com/news/2005/08/17/act

Jencks, C. 1998. Racial bias in testing. In *The Black-White Test Score Gap* (eds. Christopher Jencks and Meredith Phillips) pp. 55-85 Washington DC: The Brookings Institution

Kane, M.T. 2006. Validation. In *Educational measurement, 4th edition,* ed. R.L. Brennan, 17-64. Westport CT: American Council on Education.

Lemann, N. 1999. *The Big Test: The Secret History of the American Meritocracy*. New York: Farrar, Straus, and Giroux.

Lemann, N. 2004. A history of admissions testing. In *Rethinking the SAT: The Future of Standardized Testing in University Admissions*, (ed. Rebecca Zwick). pp. 5-14 New York: RoutledgeFlamer.

Marte, J. 2011. 10 things test-prep services won't tell you. *Smart Money*, April 5, 2011.

Munday, L. 1967. Predicting college grades using ACT data. *Educational and Psychological Measurement* 27: 401-406.

Noble, J. and R. Sawyer. 2002. *Predicting Different Levels of Academic Success in College Using High School GPA and ACT Composite Score*. Iowa City: ACT.

Perez, C. 2004. Reassessing college admissions: Examining tests and admitting alternatives. In *Rethinking the SAT: The Future of Standardized Testing in University Admissions*, (ed. Rebecca Zwick). pp. 5-14 New York: RoutledgeFlamer.

Powers, D.E., & Rock, D.A. 1999. Effects of coaching on SAT I: Reasoning test scores. *Journal of Educational Measurement*, 36(2): 93-118.

Rothstein, J.M. 2004. College performance prediction and the SAT. *Journal of Econometrics* 121: 297-317.

Santelices, M.V. & Wilson, M. 2010. Unfair Treatment? The case of Freedle, the SAT, and the standardization approach to differential item functioning. *Harvard Education Review* 80: 106-133.

Steele, C.M. & Aronson, J. 1998. Stereotype threat and the tes performance of academically successful African Americans. In *The Black-White Test Score Gap* (eds. Christopher Jencks and Meredith Phillips) pp. 401-427 Washington DC: The Brookings Institution

Zwick, R. 2004. Is the SAT a "wealth test?" The link between educational achievement and socioeconomic status. In *Rethinking the SAT: The Future of Standardized Testing in University Admissions*, (ed. Rebecca Zwick). pp. 203-216 New York: RoutledgeFlamer.