

Truth, Gender, and the SAT

Joseph M. Horn

The public relations campaign against the Scholastic Aptitude Test (SAT) has depicted the test as a “defective product” that limits the educational opportunities of women. Female high school students score lower than their male classmates on both the verbal and math components of the test, and this “gender gap” is said to disadvantage women unfairly in competition for scholarships and admission to selective colleges. According to the Center for Women Policy Studies (CWPS), the publishers of the SAT are guilty of “consumer fraud” because the test “underpredicts” the college grades of women students. This means that women find the test more difficult than men, but when it comes to getting good grades, the really important criterion according to CWPS, the women do better than the men. And since women get higher grades, they should score higher on the SAT, and would, if the test weren’t biased. If these statements were true, the lower SAT scores of women could reflect some bias in the test and modification would be in order. However, the allegations are false and their framing shows such a determined avoidance of fundamental facts that what now ought to be debated is not the “gender gap” but rather the “truth gap.”

The first erroneous allegation concerns grades. In order to make the claim that women get better grades than men, we have to ignore both the evidence that grades vary widely according to subject matter and that course-taking differs by gender. A’s and B’s are common in art and music, and women take more of these courses. Women also get higher grades in these generally more easily graded courses. The situation is the opposite in mathematics, where A’s and B’s are much harder to attain and men take more courses than women. Men get slightly higher grades in these tougher-graded courses. A more accurate statement of the facts about grades is this: in high school and college, women generally take easier courses than men and their grade point averages are slightly higher as a result.

Each year, students who take the SAT report their high school grades by academic subject. Each year the College Board publishes the results of these self-reported grades and the tabulations are consistent. Girls report higher grades in art, music, English, and languages; and boys report higher grades in math. In spite of these reliable differences, groups assaulting the SAT refuse to acknowledge that grades and grade point averages are influenced markedly by academic subject and choice of courses. If they did, they would lose a major weapon in their propaganda war.

The tendentiousness of feminist groups on this question of grades and SAT scores is also shown by the fact that another, equally plausible explanation is

never considered. Grades might discriminate against boys! Grades are certainly more subjective than objectively scored tests; and if boys score better when tested objectively than when graded subjectively, why not look for bias against boys in grading? As far as I know, Diane McGuinness is the only social scientist to write on how schools “discriminate” against boys.¹

There are other reasons why men might know more than women in a particular subject yet get lower grades. Women are, as a group, more conscientious students. They have better attendance, they more often turn in assignments on time, and their work is generally regarded as better organized. Another reason concerns the pivotal importance of writing skills in a teacher’s determination of grades. The use of essay examinations and written reports systematically favors students who are skillful writers, and women show a definite superiority on the Test of Standard Written English (TSDWE). This test is administered along with the SAT, and results are used in placing students into appropriate levels of freshman English. Women are better at recognizing standard written English than men and this could elevate their grades over those of men who might know more but fail to present their knowledge as effectively. For these and other reasons it shouldn’t surprise anyone that men might score higher on an objective test but get lower grades. It is certainly not wise to jump immediately from sex differences in grades to the assumption of a flaw in the more objectively scored SAT.

Given all that we know about factors beyond aptitude and knowledge that influence grades, it is somewhat surprising that grade differences would be used to challenge the results of objective tests. Surprising, that is, until you realize that the critics really had nowhere else to turn. Achievement tests should have been used since they provide the most logical check on the results of aptitude tests, but they don’t yield the “correct” result.

Achievement tests are most important in this context because they are designed to measure directly what students actually learn in their classes; so if women actually know more than men, even though their SAT scores are lower, it would show up in higher achievement test scores. It doesn’t happen! Women score lower on the SAT, and likewise on achievement tests. For example, women who take the biology, chemistry, math, and physics achievement tests have lower SAT math scores than men who take the same tests. These women also do worse than their male counterparts on each of the achievement tests. Similar results are obtained on the College Board Advanced Placement exams: when males have higher SAT scores, they also do better than females on tests that are designed to test knowledge more directly than does the SAT.

This is a catastrophic result for those who want to argue against the SAT. To show how decisive it is, imagine the following proposal: In any competition for scholarships or admission to a selective college, an SAT score could be used only if it were supported by a battery of achievement tests showing that the high SAT student actually knew more. If this proposal were adopted nation-

wide, the result would be scholarship awards and college admissions that are free of the taint of *unfair* gender bias.² The problem for the critics is that under this system of double-checks men would still prevail more frequently than women. As a matter of fact, there is no evidence that the ratios of men to women selected for any academic honor or program would change significantly. The plain and simple fact is that achievement test results fully corroborate the SAT gender differences. Of course, all of this means that an expensive battery of achievement tests is unnecessary. The much less expensive SAT is just as good.³

So where is the defect in the SAT? It is solely in the eyes of its attackers. Something has to explain the gender differences on aptitude and achievement tests, and it is just too painful for them to contemplate that women could be different from men in a way that is likely to affect their educational and economic futures. Fortunately, less ideologically driven people are open to data suggesting that the academic achievements of men and women develop out of different motivations, interests, and aptitudes. The attack on the SAT is really an intellectually impoverished way of dealing with these complexities.

The transparent disingenuousness of the attacks on the SAT leave one to marvel at such tactics. How can they win? That was my initial response; but, after a long second thought, I realized that a good case is not necessary if you can mobilize political forces for an administrative, legislative, or judicial suppression of the facts. The goal of these groups is not to investigate the validity of the SAT but to eliminate or change it in the direction they want by whatever methods present themselves. For example, can a judge be convinced, in spite of the evidence, that females score lower on the SAT because the test itself is biased? If so, he could stop the use of the SAT. The tragic order of Judge Peckham forbidding the use of intelligence tests in the assessment of minority children in California shows what can be achieved with propaganda and a sympathetic judge.⁴

Fortunately, the most recent court ruling involving the SAT preserves the freedom of scholarship committees to utilize the SAT in their awards process. Judge John Walker did rule, however, that sole reliance on the SAT for scholarship awards has a disparate impact on women, and he directed the state of New York to use some procedure where the SAT is not the only consideration. New York had previously used grades—plus the SAT—to make scholarship awards, but dropped grades because this criterion encouraged grade inflation and placed students from tougher grading schools at a disadvantage. This ruling presents a perfect opportunity to put into practice the proposal to require achievement tests along with the SAT. The sobering fact is, however, that since this won't equalize the number of awards given to men and women, there will be great pressure to reintroduce the use of grades, even though we know that this criterion is contaminated and has the serious problems noted by the state of New York.

The best hope for avoiding rigged criteria that produce the effect of quotas lies in the fact that the “disparate impact” theory used by Judge Walker is now being modified by the United States Supreme Court. Until recently the doctrine of disparate impact required that a selection procedure producing statistical discrepancies from proportional (quota) representation had to be shown to be necessary by the college using the procedure. That is to say, the burden of proof, after disparate impact had been established, lay with the defendant, not with the plaintiff. This is no small matter since the party with the burden of proof usually has expensive validity studies to carry out and greater legal expenses. This distribution of the burden of proof also motivated plaintiffs to sue and defendants to settle. However, the Supreme Court reversed itself on June 5, 1989, ruling that the burden of proof belongs with the plaintiff.⁵ The struggle is far from over, however. Pressure is building on Congress to legislate equality and return the burden to the defendant.

The troublesome picture that emerges from all this conflict over the SAT is that questions of learning, achievement, and development of talent have clearly been subordinated to those of equality. The critics of the SAT really want to find some way to force the educational institutions to reduce, if not eliminate, their legitimate recognition of individual and group differences. This result is already close to being achieved in the primary and secondary schools, where the most pressing concern is that all groups of students graduate at the same rate. If this also becomes true for higher education, we can expect declines in college achievement similar to those we have observed in high school.

The obsession with equity in education has had a debilitating effect on school curricula, and able students are very hard pressed today to find real intellectual challenges in most high schools and, alas, many colleges. Dramatic confirmation of this trend and its consequences was recently provided by the Educational Testing Service in its report “Crossroads in American Education.” This study shows that the recent small gains in overall student achievement are due to improvements at the lower levels of performance. The number of students performing at higher levels continues to decline. This outcome closely mirrors the current concentration of teaching resources and efforts. Today, the focus is on the “at-risk” student. In this environment capable students are relatively neglected and their achievement sags.

What is really at risk today in American education is the idea of excellence. The attacks on the SAT show this quite clearly. Because all groups don’t perform equally well on the test—forget the similar differences in actual knowledge—it must be eliminated or neutered. The SAT reminds us, year after year, that some of our children have future educational prospects that vastly exceed those of other children. We might hate this, but we have to resist the temptation to slay the messenger. It is no improvement to fake equality by bringing down those at the top.

Notes

1. Diane McGuinness, "How Schools Discriminate Against Boys," *Human Nature* (February, 1979): 82-88.
2. If a valid criterion measure contains a bias, e.g., valid achievement tests show differences by socioeconomic class, it would be foolish to remove the same bias from any test developed specifically to predict performance on that biased criterion. This removal of predictor bias has been tried with IQ tests and the socioeconomic bias-free IQ tests that lose their ability to predict biased criteria of school performance. Since males and females differ on achievement test scores, it would be reckless to take gender differences entirely out of the SAT. The question of whether all performance criteria that show biases ought to be abandoned is moot. Our schools cannot stop caring about achievement in reading, writing, and arithmetic.
3. Christopher Jencks and James Crouse report that "for those concerned with equality of opportunity the choice between so-called aptitude tests and conventional achievement tests is a toss-up." See "Should We Relabel the SAT . . . or Replace It?" in *Measurement, Guidance, and Program Improvement: Proceedings of the 1981 ETS Invitational Conference*, ed. William Benton Schrader (San Francisco: Jossey-Bass Inc., 1982), 44.
4. See Roger Elliott, *Litigating Intelligence: IQ Tests, Special Education, and Social Science in the Courtroom* (Dover, Mass.: Auburn House, 1987).
5. *Wards Cove Packing Company, Inc. V. Atonio*, 109 S. Ct. 2115 (1989).