

who might otherwise be lost in the shuffle. The SAT, this argument goes, puts all students on an equal footing, giving them a uniform, color-blind test on which to demonstrate their ability. "Unlike the personal interview, the classroom test, or the teacher's subjective evaluation," wrote Henry Chauncey in 1961, "the objective test is a common touchstone. It gives all students who take it the same chance, asks them to run the same race — even though they have had different economic backgrounds, different educational, cultural, and social opportunities."³⁹

This is a cheerful thought, but it has no basis in reality. As ETS and the College Board finally acknowledged with published statistics in 1982, there is a considerable gap between the average SAT performance of whites and that of blacks, Mexican-Americans, and Puerto Ricans, and between people whose families have a lot of money and people whose families don't.⁴⁰ In 1983, 7,263 whites scored above 700 on the verbal SAT; so did 66 blacks. In the same year, 54,277 whites scored 600 or above; so did 939 blacks. (The test population included 685,219 whites and 71,490 blacks.) The mean score for whites on the verbal SAT was 443; the mean for blacks was 339.^{*41}

If colleges actually used the SAT as a "color-blind" indicator of academic ability, you wouldn't find many minority students enrolled in selective schools. Most minority applicants who are admitted to selective colleges are admitted *in spite of* their SAT scores, not because of them. Admission to college in these circumstances carries with it a built-in slap in the face: you can come to our school, but you're not really entitled.

Whether the SAT is culturally biased against minorities is another hardly perennial controversy in which ETS takes the kettle position: the tests were never biased; they've now been fixed; but the changes have had no effect. ETS naturally says that it has proved statistically that its tests aren't biased. Just to make sure, for the last few years it has used "an actual member of a minority" (as one ETS employee told me) to read every test before it's published. According to an ETS flier, "Each test is reviewed to ensure

*ETS wouldn't send me the publication in which these 1983 statistics appear. It also turned down my request for the comparable publication from 1982. I ended up getting copies from the College Board.

that the questions reflect the multicultural nature of our society and that appropriate, positive references are made to minorities and women. Each test item is reviewed to ensure that any word, phrase, or description that may be regarded as biased, sexist, or racist is removed."⁴²

But the actual "sensitivity review" process is much more cursory and superficial than this description implies. The minority reviewer simply counts the number of items that refer to each of five "population subgroups" and enters these numbers on a Test Sensitivity Review Report Form. On the verbal SAT administered in May 1982, actual minority member Beverly Whittington found seven items that mentioned women, one that mentioned black Americans, two that mentioned Hispanic Americans, none that mentioned native Americans, and four that mentioned Asian Americans (actually, she was stretching here; these particular Asian Americans were Shang Dynasty Chinese, 1766–1122 B.C.). Two items overlapped, so Whittington put a "12" in the box for Total Representational Items. She also commented "OK" on the exam's test specifications, "OK" on the subgroup reference items, and "OK" on item review. She made no other remarks. If she had discovered the word *nigger* in one of the questions, presumably she would have scratched it out. ETS made Whittington take a three-day training program in "test sensitivity" before permitting her to do all of this. When her report was finished, it was stamped E.T.S. CONFIDENTIAL and SECURE. Then it was filed and forgotten.⁴³

SAT items don't have to be very representational in order to be counted as Representational. The pronoun *she* is enough to make an item "woman-oriented." The SAT I took in December 1983 contained a black-oriented analogy:

18. DASHIKI: GARMENT :: (A) spoon: utensil (B) hat: coat
(C) podium: conductor (D) foot: shoe (E) plate: table

Dashiki is a Yoruban word for a loose-fitting tunic worn in certain parts of Africa; it enjoyed a vogue among politically aware black American males during the late sixties and early seventies. ETS put this item on the SAT to symbolize the company's profound new

218 regard for the black experience. But are black seventeen-year-olds any more likely than white seventeen-year-olds to know what dashikis are? Would ETS ever have asked a white-oriented SAT question about Nehru jackets? Does familiarity with either garment have anything whatsoever to do with scholastic aptitude?

Fifteen years ago, ETS Test Sensitivity Reviewers, had they existed, would have found few dashikis to count in the SAT. But in 1970 the company began adding one or two "minority-oriented" reading passages to each edition. Now, because of the way the SAT is equated, ETS could only add minority-oriented reading passages that performed exactly like non-minority-oriented reading passages, since non-minority-oriented reading passages were what the SAT had always had *before*. In the same sense, it couldn't snip the word *nigger* out of one item unless it could find some way to sneak it into another. But the company had been accused of being biased against minorities and, although it denied the accusations, it decided it ought to make a public gesture.

In 1974, Steven Brill, writing for *New York* magazine, asked ETS executive Marion Epstein about these new reading passages.

Q. If the tests weren't culturally biased in the first place, why did you make the change?

A. Because minorities feel at ease reading this kind of passage.

Q. If they feel at ease reading this one, does that mean they don't feel at ease reading the six or seven others in the text?

A. No. It just means they feel more comfortable with this one.

Q. Well, if they feel more comfortable, does that mean their scores will be higher?

A. No, I don't think there will be any difference in "scores."

Q. Well, if there won't be any difference in scores, why would you make the change? Was it just so you could look like you were doing something?

A. No, it's because when people are more comfortable, they'll do better on the test. They feel less threatened.⁴⁴

Is the SAT biased against blacks? A senior research scientist at ETS, who asked not to be identified, told me that black students tend to do better in college than their SAT scores predict they will.

219 If you have a black student and a white student with identical scores, he said, you can expect the black student to earn a higher grade point average than the white student. No doubt the motive for this assertion is high-minded: to rebut accusations that unqualified blacks are being admitted to selective schools because of favored treatment. But if the assertion is true, the SAT is literally racist: it systematically gives blacks lower SAT scores than they deserve in terms of the sole criterion by which the test's validity is judged.

But maybe the research scientist was wrong.* When I asked Arthur Kroll if the SAT penalized blacks, he said, "If you mean, Does the SAT predict as well how minority students are going to do in college as majority students, then the SAT has done as effective a job for blacks as for whites." If I understand Kroll correctly (why do so many people at ETS have so much trouble with syntax?), this means that the SAT is not biased either for or against blacks. A different story. So I asked Ernest Kimmel. "I guess I'll disagree slightly with my boss," he said. "The scores do not work exactly the same with whites and blacks. If an admissions officer uses a single admissions equation based on a mix of white and black students with the same scores and the same high school rank, he's going to predict the same grade averages. But in actuality the black students in about 80 percent of the studies seem to do a bit worse." In other words, blacks get worse grades than their SATs would predict; the test is biased against whites.

I now had three apparently contradictory explanations. I took them to Richard Noeth, the ETS official who had told me (in Chapter 2) that essay exams are different from multiple-choice exams but are neither better nor worse.

"The thing is," Noeth said, "to my knowledge, there's support-

*Sometime after my *Harper's* article was published, the research scientist called me to say I had been wrong and that "blacks tend to do *worse* in college than their SAT scores predict they will." I checked my tape of our conversation and have no doubt that I had accurately reported what he told me. Perhaps he changed his mind. Whether he did or not, many admissions officers would say he was right the first time. James McMenamin of Columbia told me that "educationally disadvantaged" students (meaning, for the most part, blacks) generally did better at Columbia than their SAT scores had predicted. Admissions officers at several other schools told me the same thing and said they often adjusted such students' SAT scores upward to correct for the disparity.

ing evidence for each of the three positions that you mentioned. I tend to — I believe them all. I'm sure they're all true."

Well, we've certainly cleared up this bias business. But we're left with the disparity in scores. Several hypotheses have been offered to explain this:

1. *Whites are simply more intelligent than blacks; the rich are more intelligent than the poor.*

This explanation is favored not only by dyed-in-the-wool racists but also by meritocratic purists. Daniel Seligman, associate managing editor of *Fortune*, put it this way in his publication's May 5, 1980, issue: "E.T.S. tests persist in showing some people to be smarter than others. And if some people *are* smarter than others, there might actually be some justification for an economic system in which some people have more money and authority than others. . . . The really interesting question is not whether rich people are smarter, but why they are. Is it because of their superior environments or their superior genes? The answer — also attested to in endless studies — is 'both, obviously.'" Seligman's article is entitled "The Rich Are Different." It is illustrated with a drawing of a Rolls-Royce whose license plate reads "IQ 200."⁴⁵

This is pure, unreconstructed Carl Campbell Brigham and the product of an almost unbelievable naïveté about American society. Once, on a reporting assignment, I sat in a bar at an exclusive country club and listened to a silver-haired millionaire tell me that the reason there are so few black professional golfers is that blacks aren't smart enough to grasp the intricacies of the game. The backdrop for our conversation was a golf course where blacks were not allowed to set foot even as guests. The idea that there might be a connection between all such prohibitions and the dearth of black golf pros had never occurred to this man.

To recast Daniel Seligman's argument: If the Professional Golfers Association persists in showing some people to be smarter than others, might there not be some justification for a society in which some people are allowed to join country clubs and others are not?

Racial prejudice is so pervasive and insidious that it can contaminate even the most "scientific" experiments. The history of psychometrics has been, in many ways, the history of self-

deception. Terman, Goddard, Yerkes, Brigham — their private hatreds largely determined their findings. The fact that "endless studies" of comparative intelligence — from Army Alpha to Arthur Jensen's experiments with reaction time⁴⁶ — have shown similar results doesn't strengthen the case: If new findings correlate highly with discredited old findings, shouldn't that cast doubt on the quality of the new findings? Since no one knows what "intelligence" is (a high score on a multiple-choice test? a low score on a golf course?) or how to measure it directly, intelligence testing is extremely susceptible to the conscious and unconscious predispositions of the testers. Measures that seem to confirm these biases are held to be measures of intelligence; measures that don't confirm them are presumed to be measures of something else. When studies show that the wealthy in America generally score higher on aptitude tests than the poor, people begin to talk about genetically determined differences in "mental capacity"; when studies show that Japanese and Taiwanese first-graders score higher on aptitude tests than their American counterparts, people talk about "low teacher salaries and curriculum content."⁴⁷

2. *The test is not a fair measure for blacks and members of various other disadvantaged social groups.*

This argument holds that questions on the SAT systematically penalize students from outside the American economic elite. The poor tend to do worse than the wealthy on the SAT, this argument states, because the questions on the test reflect only a white upper-middle-class point of view; if the questions were written differently, racial and economic differences would disappear.

ETS and the College Board are very sensitive about this issue, as their adoption of "minority-oriented" reading passages proves. But both ETS and the Board believe firmly that this was a solution without a problem; the passages were added to improve public relations, not to correct an injustice. ETS says that wealthy whites don't have an undeserved advantage on the SAT, and it says it has proven this statistically.

Thomas F. Donlon, an ETS research scientist, addressed this issue several years ago in the *College Board Review*, a quarterly magazine for Board members and others. "The SAT," he wrote, "by virtue of the breadth of its coverage, and the careful editing of

its content, is a balanced instrument with relevance for a variety of candidates. Although it is now only two hours long, it covers a range of topics and tasks that tends not to favor any one subgroup."⁴⁸ To prove his point, Donlon analyzed an SAT item of the sort that critics often point to when they claim that wealthy whites have an unfair advantage on the test. The item was an analogy from a verbal SAT:

RUNNER: MARATHON::

- (A) envoy: embassy
- (B) martyr: massacre
- (C) referee: tournament
- (D) oarsman: regatta
- (E) horse: stable

Regattas, Donlon conceded, "are less frequently associated with the minority experience." But he said that most white people probably don't know much about them either. And although he acknowledged that only 22 percent of black students — compared with 53 percent of whites — had selected the correct answer, he said that this did not make the question unfair. To prove it, he compared the performance of blacks with the performance of whites of "matched ability" (as determined by SAT scores). Looked at this way, the results were closer. Only 26 percent of the "matched" whites selected the right answer, compared with 22 percent of the blacks.⁴⁹

"It is clear," Donlon concluded, "that these matched whites perform very similarly to blacks. The clear implication of this is that it is not a black subculture which influences these patterns (for whites are not raised in that subculture), but some sort of general lack of knowledge or understanding of this item. . . . When the responses are examined in depth, the patterns of success and error for racially defined groups of comparable ability are not different."⁵⁰

* Actually, the way I figure it, 22 percent and 26 percent *are* different, as are the "patterns of error" Donlon gives for other responses. If these differences are meaningless, Donlon should explain why.

But all of this begs the real question: What exactly can ETS mean by "ability" if one of the differences between having it and not having it is the difference between knowing and not knowing about oarsmen and regattas (or about polo and mallets, or golf and clubs — as other SAT questions have asked)? Surely the fact that rowing, polo, and golf are sports of the wealthy accounts in some measure for the mysterious "general lack of knowledge" that Donlon detects among the huddled masses. Donlon's wholly circular statistical argument does not prove that the SAT "tends not to favor any one subgroup." Obviously the SAT favors at least *one* subgroup: the subgroup that does well on the SAT. And if this particular subgroup has certain consistent characteristics (familiarity with polo equipment, for example) that have little to do with "ability" but much to do with success on the test, then the SAT is unfair.*

I can think of many subgroups for whom the SAT is unfair: the subgroup that believes coaching doesn't work; the subgroup that thinks guessing is wrong; the subgroup that takes the test under the supervision of incompetent proctors; the subgroup that takes the test under the supervision of incompetent observers from "Princeton"; the subgroup that doesn't cheat; the subgroup that believes the SAT measures what the College Board once called "the future worth of candidates";⁵¹ the subgroup that attends schools where ETS tests are not a way of life; the subgroup that doesn't understand the way the test is constructed; the subgroup that uses the Arco strategy for answering analogy items; the subgroup that applies to colleges with uninformed admissions officers.

* In the mid-1970s, the Golden Rule Insurance Company of Lawrenceville, Illinois, sued ETS on behalf of five blacks who had failed an ETS exam used by the state to license insurance agents. Golden Rule argued that the test was biased against blacks, since white candidates as a group earned substantially higher scores on it. ETS fought the suit for years, dragging it through various courts in Illinois. Finally, in November 1984, when it seemed clear that the case would actually go to trial, ETS settled with Golden Rule. The testing company agreed to pretest questions for the Illinois exam (something it had never done before) and to eliminate those items on which there is a marked disparity between performance by whites and performance by members of minorities. The settlement applies only to the Illinois insurance exam, but in time the principle will undoubtedly be applied to other standardized tests, including the SAT.