

Editorial

Biased Tests

You might have seen a *New Yorker* cartoon about a man who is trying to get into Heaven. As St. Peter thumbs through the relevant documents, the man says, "You're kidding! You count S.A.T.s?" (Twohy 1991).

The SAT and other standardized tests have enormous power, often determining how far people can advance in the world. And the tests create special roadblocks for low-income students and students of color. Because these students, compared to white middle class students, generally receive lower test scores, they are disproportionately excluded from our top colleges and universities. In this issue of *Encounter*, Tasha Prosper provides an emotional and sharply observed account of the experiences of African American women in today's test-dominated society. As Prosper points out, African Americans are coming to believe that standardized tests have replaced the dogs, fire hoses, and separate-but-equal laws that used to keep her people "in their place."

It's my impression that most educators, and much of the public, know that there are racial, ethnic, and income disparities in standardized test scores. But people rarely question the tests. Instead, they see the problem as having to do with the test-takers—if not with their native intelligence, then with deficits in their educational preparation.

New Evidence of Bias

Educational preparation does matter. But new research suggests that the tests themselves are quite biased. As a psychologist, a profession with expertise in testing, I must confess that one source of bias went unnoticed for a very long time. The bias has to do with a technical goal in test construction.

When they create a new test, test makers strive for a form of *reliability* called *internal consistency*. The test

makers first examine individual pretest items to see how well the items are associated with the overall test scores. Then, when they decide on the final version of the test, they retain only the items that are consistent with the overall scores.

But Jay Rosner, executive director of the Princeton Review Foundation, points out that this seemingly neutral goal hurts people of color. In an interview with the *Chronicle of Higher Education* (Young 2003), Rosner described his findings with respect to the October 2000 SAT. He discovered that the test makers retained all the pretest items on which the white students outperformed the African Americans and Latino students, but they discarded all the pretest items on which the African American and Latinos outperformed the whites. These items weren't associated with most of the other high scores—those achieved by the larger numbers of whites—so the items were thrown out.

The test makers' decisions were not motivated by racial prejudice. The people who develop each version of the SAT—the Educational Testing Service (ETS) researchers—didn't know the students' racial identities. They simply cast aside the items on which the students of color did better because those items weakened the test's internal consistency. It was all done on purely technical grounds.

Freedle's Research on Difficult Items

Another critique has been leveled by Roy O. Freedle (2003), a recently retired ETS scientist. Freedle found that among white and African American students with identical scores on the SAT Verbal test, the white students perform slightly better on the *easy* items, while the African American students perform slightly better on the *hard* items. These differences are small but consistent. At whatever level one looks (e.g., a score of 290, 300, or 640 on the SAT Verbal test), the same pattern appears.

Note: Readers may access a discussion guide for this editorial at www.great-ideas.org/Edit173Guide.htm.

Freedle reports that this pattern also characterizes white and African American students' performances on the SAT's quantitative test, and that the pattern holds for other ethnic minorities and low-income students as well.

Freedle is so impressed by the African-Americans' performances that he wants ETS to use a separate SAT score, an R-SAT score, which would consist only of the *hard* items. He estimates that the R-SAT would reduce general white/African American disparities by a third. What's more, Freedle says, a small number of African Americans would gain enormously. For example, some students with SAT Verbal scores under 300 would soar to over 600, qualifying them for admission to the country's elite colleges and universities.

Why do African Americans do relatively better on the difficult SAT items? Freedle thinks the explanation has to do with language. The easy items typically use everyday vocabulary that is actually the vernacular of white, middle class students. These students are therefore more familiar with its cultural nuances. The difficult items, in contrast, use more abstract terms that are usually found only in textbooks and classroom lectures. This abstract language is more precise and culturally neutral. It provides a more level playing field on which African Americans do relatively better.

Freedle's research doesn't explain all the SAT differences between whites and African Americans—only about a third. So other variables, such as academic preparation and special tutoring, are still important considerations. But Freedle's findings, together with those of Rosner, point to biases that are so substantial that they call the entire SAT into question.

High School "Push Out"

The SAT is not, of course, the student's first experience with standardized tests. Since the 1983 publication of *A Nation at Risk*, almost every state in the country has installed rigorous standardized tests at almost every grade level. What's more, the states are increasingly attaching high stakes to the tests; they are making grade promotion (e.g., advancement from 3rd grade to 4th grade) and high school graduation contingent upon specific test scores. It will be important to see if the same biases that Rosner and

Freedle have uncovered for college-entry tests also plague the tests for younger students.

In the meantime, information on the harmful effects of standardized tests keeps growing. In two recent reports, Gary Orfield at Harvard University (Orfield et al. 2004) and Walt Haney at Boston University (Haney et al. 2004) describe how high-stakes tests are contributing to a crisis in graduation rates.

The SAT and other standardized tests have enormous power, often determining how far people can advance in the world. And the tests create special roadblocks for low-income students and students of color.

Nationwide, today's four-year high school graduation rate is only 68%; among African American and Latino youth, the graduation rate barely exceeds 50% (Orfield et al. 2004). At least part of the explanation, Orfield and Haney contend, is the practice of informally pushing out students who are likely to earn low test scores. By dropping these students from their rolls, schools can raise their overall scores. The victims are commonly low-income students and students of color.

Early Discouragement

High school "push out" has attracted media attention. But high stakes testing also begins turning students off to school in the early grades. Under the pressure to prepare children for the tests, schools have little time for the activities children find exciting and meaningful—activities such as building things, gardening, producing plays, and conducting research projects. Instead, children must spend hours on test-prep drills and exercises that they find extremely boring.

The tests also produce considerable anxiety. Students dread the tests throughout the school year, and as the testing dates approach, their anxiety intensifies. Many children cannot sleep at night and

develop stomach aches and head aches. If they face the humiliation of grade retention, their fear is even greater.

These feelings of boredom and fear ruin the entire educational enterprise. As Dewey, Piaget, Montessori, and others have argued, intellectual development occurs when children are enthusiastic about activities. When children become engrossed in tasks, they think deeply and imaginatively, and their minds expand. By replacing the child's enthusiasm with boredom and fear, test-driven education stifles the urge to learn.

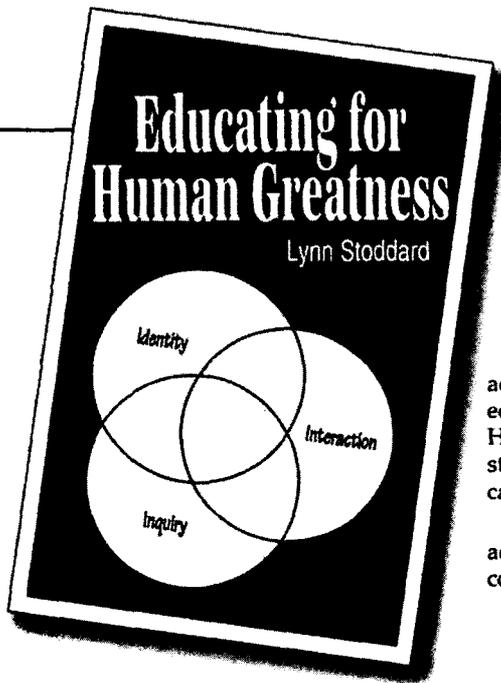
Since low-income children and children of color generally have the most difficulty with the tests, they are given the largest doses of monotonous test-prep drills and they experience the greatest fear of failing. For them, school is an even more unpleasant than it is

for others. It is little wonder that when they are old enough, they seriously consider dropping out. Given a nudge by school officials, many do.

—William Crain, *Editor*

References

- Freedle, R. O. 2003. Correcting the SAT's ethnic and social-class bias: A method for reestimating SAT Scores. *Harvard Educational Review* 73: 1-38.
- Haney, W., G. Madaus, L. Abrams, A. Wheelock, J. Miao, and I. Gruia, I. 2004. *The education pipeline in the United States 1970-2000*. Chestnut Hill, MA: National Board on Educational Testing and Public Policy.
- Orfield, G., D. Losen, J. Wald, and C. B. Swanson. 2004. *Losing our future: How minority youth are being left behind by the graduation rate crisis*. Cambridge, MA: The Civil Rights Project at Harvard University.
- Twohy, M. 1991, September 9. "You're kidding! You count S.A.T.s?" Cartoon. *The New Yorker*.
- Young, J. R. 2003, October 10. Researchers charge racial bias on the SAT. *The Chronicle of Higher Education* 50: A34.



Now Available from
Holistic Education Press
ISBN 1-885580-16-9 \$18.95

Order Your Copy Today!
1-800-639-4122
www.great-ideas.org/Stoddard.htm

Educating for Human Greatness

In this wise and perceptive book, veteran public school teacher/administrator Lynn Stoddard surveys the current state of public education in America and concludes that things have gone terribly wrong. His solution is to have parents and educators start by realizing that standardization in education is neither possible or effective. Only then can they focus on creating schools that truly educate for human greatness.

To create such schools Stoddard proposes that parents, teachers, administrators and school board members keep six cardinal principles constantly in mind:

- Value Positive Human Diversity and Cherish Every Student's Uniqueness
- Draw Out and Develop Each Child's Latent Talents
- Respect the Autonomy of the Individual by Restoring Freedom and Responsibility
- Invite Inquiry, Curiosity, and Hunger for Knowledge in the Classroom
- Support Professionalism as Teachers Live by these Principles
- Parents and Teachers Unite to Help Children Grow in Human Greatness

Educating for Human Greatness deserves an honored place on the reading list of every parent who really cares about the future of their children, every teacher and administrator who puts students first in their professional lives, and every school board member who wants schools to be places where student development is a reality, not just a slogan.

Copyright of Encounter is the property of Psychology Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.