

No Exit?

Testing, Tracking, and Students of Color in U.S Public Schools

© Applied Research Center, February 1999

Principal Researchers: Rebecca Gordon & Libero Della Piana

Additional research by: Phyllida Burlingame, Akilah Monifa & Oliver Wang

"Now, I want especially to speak to the notion of accountability. While a number of other states require students to pass a statewide minimal skills exam to graduate from high school, California does not. I believe we need to do even better. I am proposing a rigorous high school graduation exam, second to none in America."

- California Governor Gray Davis, 1999 State of the State Address

The word "accountability" carries considerable weight in debates about public education. In his 1999 State of the Union address, President Clinton outlined a package of "accountability" measures for students, teachers, and schools, measures which would reward the schools whose students get high test scores and punish the others. Following the example of 26 other states, California's Governor has called for deployment of an exit exam—a test high school seniors must pass in order to receive their diplomas. So crucial to California's future is this new accountability measure that in January 1999 the Governor called a special session of the Legislature to take up the issue.

While the President and California's Governor may well hope to hold schools accountable for their students' performance, high-stakes testing policies end up punishing individual students. It is not the school but the student who, after passing four years of high school classes, must either pass a test or enter the job market without a high school diploma. For many students, and disproportionately for students of color, the imposition of exit exams is like adding an unanticipated 100-yard dash to the end of a 26-mile marathon.

The question of fairness aside, there is no evidence that the use of exit exams in particular, or standardized testing in general, actually improves the education students receive. Robert Schaeffer of FairTest, an advocacy group in Cambridge, Massachusetts, says, "Believing we can improve schooling with more tests is like believing you can make yourself grow taller by measuring your height."

Testing, Testing

School systems use standardized tests for two main purposes: either to measure the performance of a group of students in the aggregate—for example in a single school, a school district, a state or the nation—or to measure the performance of individual students.

Generally, a test that has been validated for one purpose should not be used for the other. For example, a test designed to measure the progress of an entire class of third graders is not validated to determine whether an individual third grader should be allowed to take classes for "gifted" students. Nor should tests designed to measure a whole grade level's performance be used to deny students promotion to the next grade, as President Clinton proposed at a 1999 education summit. Unfortunately, all too often standardized tests designed to measure the performance of entire schools or school districts have powerful effects on the lives of individual students.

Tests that can determine the outcome of a student's academic career are known as "high stakes" tests, because their results are so significant for the individuals who take them. There is considerable controversy about the value of using standardized, multiple-choice tests by themselves, either as predictors of a student's future performance or as yardsticks for educational institutions. Many experts now favor examining a combination of grades, open-ended tests, teachers' impressions, and portfolios of student work, to gain a more complete picture of a student's (or a school's) performance.

High Hopes, High Stakes

"Policymakers need to be aware that more rigorous assessments will likely produce higher failure rates, particularly among educationally disadvantaged students."

- National Governors' Association, "High School Exit Exams: Setting High Expectations"

Exit Exams: Twenty-six states currently have high school exit exam programs. Some states, like Florida, have used the exams for many years, but in most states they are relatively new. Arizona's program begins in 1999; programs are slated to start in the next few years in Alaska, Louisiana, Maryland, Massachusetts, New Jersey, Ohio, Virginia, and Washington.

One of the first states to use exit exams was Florida, which first implemented the Florida High School Competency Test in the 1977-78 school year. The next year, a group of African American students challenged this new graduation requirement as racially biased, based on a disproportionate failure rate among African American students, in a case known as *Debra P. v. Turlington* 644 F. 2d 397 (5th Cir. 1981).

The district court initially ruled in favor of the plaintiffs, and the Fifth Circuit Court of Appeals affirmed its ruling, writing that the state had not "made any effort to make certain whether the test covered material actually studied in the classrooms in the state and because the record was insufficient in proof on that issue." In addition, because the State of Florida had created an "expectation" that if students attended school during the required years and passed the required courses, students had a legally recognizable "property interest" in their diploma.

Exit Exam Criteria: The case was reheard in 1983 (*Debra II*), and this time the court reversed its decision, ruling that in fact the students had had an opportunity to learn the material contained in the exit exam. In this ruling, the court set three standards that exit tests must meet to be considered constitutional:

- The test must measure the knowledge and skills that are actually taught in the state's schools; in other words, the test must have "curricular validity."
- Students must receive adequate notice of the test, the requirements for passing the test, and the consequences of not passing the test.
- The test must not intentionally discriminate against a protected group or class.

Even though a disproportionate number of African American students who had completed all other graduation requirements were failing the test, the courts did not consider the exit exam to be discriminatory, because there was no intentional discrimination. But, as we have seen, a test can have an outcome bias without expressing the intentional bias of any individual person or group of people.

But an outcome bias is still a bias; whether or not anyone intended it, a disproportionate number of African Americans were deprived of diplomas they had earned, no matter what anyone meant to happen. Good intentions do not excuse racist results.

There is a recognized method of reducing cultural bias in testing known as the "Golden Rule Bias Reduction" safeguards. This technique makes exams less discriminatory by drawing from pools of equally difficult test items those questions that produce the least difference in correct answer rates between whites and people of color.

In virtually every state that has implemented high school exit exams, a disproportionate number of those who have passed all other requirements but fail to graduate, or who must go into the world with a "certificate of completion" instead of a diploma, are students of color. As critics of exit exams point out, exit exams essentially punish poor students and those of color for attending substandard schools.

The State of Texas is a good example. The Texas Assessment of Academic Skills (TAAS) was first implemented in 1990-91. According to a suit brought last year by the Mexican American Legal Defense and Education Fund (MALDEF), "The TAAS passage rates of Mexican American and African American first-time test takers are significantly lower than those of white students. Whites are almost twice as likely as Mexican Americans and African Americans to pass the TAAS. Although white students have passed the test at a rate of approximately 70%, Mexican Americans and African Americans have passed at rates of around only 40%. About 60% of the minority students in Texas public schools begin their junior years under a cloud of doubt about their futures in the public schools of Texas. They will not be allowed to graduate if they do not pass at least one more part of the test, regardless of their grades and academic record." (Maldef 1997)

Like most states, Texas allows students to take the tests repeatedly, but some students, including students with high grade-point averages, never succeed in passing the TAAS. According to MALDEF's complaint, although Mexican Americans and African Americans make up approximately 40% of all Texas seniors, they

represent 85% of the 7,650 students who fail the final administration of the TAAS each year.

Not only do the Texas exit exams unfairly penalize individual students of color, the state's system of rewards and sanctions also operates to the detriment of the schools that serve these children, according to the June 24, 1997 issue of Time magazine. At one time Texas offered \$5,000 bonuses to principals of schools where most students passed, while sanctioning schools with low passing rates. These policies make it harder for schools that need the most help-often the ones with the least resources to begin with-to attract the best teachers and administrators.

As in most states, there is a huge disparity between the resources available to the wealthiest and poorest of Texas schools. For example, the U.S. Department of Education reported that in 1993-94 per capita public school expenditures ranged from \$3,100 to \$42,000, depending on the school district. It is not surprising that students whose schools can spend \$42,000 per year on their education do better.

The impact: Most states that use exit exams do provide some kind of remediation for students who do not pass on the first round. Unfortunately for students of color, this often means that young people who are otherwise succeeding in school and themselves shunted into remedial classes as early as eighth grade, simply as a result of one bad test score.

In addition, schools have been known to inflate their scores by "adjusting" the pool of students taking exit exams. According to Time, these techniques include shunting students who do not pass into special education programs, forcing students who have passed their courses to repeat a grade, or simply encouraging them to dropout. This is particularly harmful to students of color, considering that Latino and African American high school drop out rates were 30% and 12.6% respectively in 1997, while the white dropout rate is at an all-time low of 7.7%.

Another problem with high school exit exams affects all students, not just students of color: conditioning graduation on passing a test forces schools and teachers to "teach to the test"-which means concentrating on narrow drills, rather than an expansive curriculum. Some Texas teachers report that since implementation of the TAAS, they now spend two out of every four terms each year teaching specifically for the TAAS, instead of their schools' own curriculum. This can be particularly difficult for schools that wish to experiment with more inclusive, multicultural curricula. * The SATs, which are administered by the nonprofit Educational Testing Service (ETS) in Princeton, New Jersey, are probably the best-known standardized academic tests. For most colleges, students' SAT scores, especially on the tests of math and verbal skills, form an important part of the application process.

What do the SATs measure? That is a good question. The expression "SAT" originally stood for "Scholastic Aptitude Test"-indicating that the exam was supposed to measure a student's native intellectual abilities. Later, the SAT came to be the "Scholastic Achievement Test," and was supposed to measure what a student had learned. Today, the initials "SAT" do not stand for anything at all, and the tests are validated (statistically verified) for only one purpose: predicting first-year college grades.

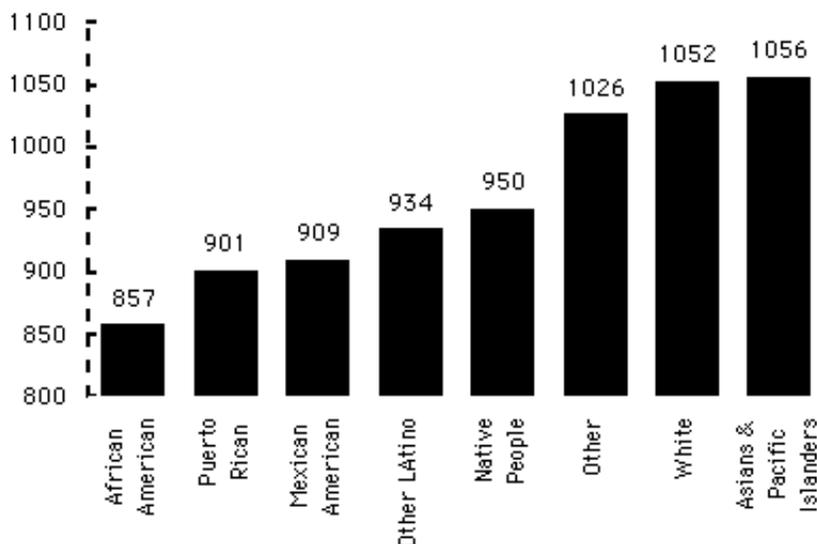
How well does the SAT score on the one thing it is supposed to measure? Not so well. In fact, the ETS concedes that high school grades are a better predictor of first-year college grades than are SAT scores. Furthermore, using the SATs as a factor in college admissions disproportionately excludes qualified students of color. In 1988, James Crouse and Dale Prushein performed a careful academic study (Crouse 1988), comparing two college admissions strategies-one using high school records alone and the other using high school records in combination with SAT scores. More than 90% of the admissions decisions were the same under both strategies, but including SAT scores in the admissions process resulted in the rejection of many students of color who were otherwise academically qualified.

Racial Outcome Bias: Colleges first began considering the SATs in their admission decisions in the 1920s. The original SATs were open-ended exams, modeled on early IQ tests, which are now generally considered to have been culturally biased against African Americans and most immigrants. The designer of the original SAT, psychologist Carl Brigham, was a committed eugenicist, who believed in the natural inferiority of Southern and Eastern Europeans, Jews, and African Americans. Of the latter group, he once wrote, "The decline of American society will be more rapid than the decline of intelligence of European national groups owing to the presence here of the Negro." (Weissglass 1998)

Almost all standardized tests, including IQ tests and the SATs, have what is called a statistical "outcome bias" against African Americans and other people of color. That is, African Americans consistently score measurably lower on these tests than do white test takers. (The fact that a test has an outcome bias does not mean that the people who designed the test were consciously, or even unconsciously, biased. It simply means that there is more than an accidental difference among the scores of different groups of test-takers.)

Although the gap between the scores of whites and students of color has narrowed over time, it is still significant. For example, on the 1997 SAT tests, out of a possible combined score of 1600 (for verbal and math portions together), the average score for African Americans was 857 and for whites, 1052—a difference of 195 points. With the exception of Asians and Pacific Islanders (whose scores were virtually identical to those of whites), scores for all other minority groups were also lower than for whites. See the charts below:

fig. 1: Average 1997 SAT Scores for Minority Groups



Because the SAT does not predict college grades accurately, and because of the SAT's outcome bias, almost 300 four-year colleges have dropped it as an admissions requirement for at least some of their students. Some highly selective schools like Bates and Bowdoin have seen increased diversity in their student bodies. Ironically, just as states are implementing high school exit exams, state university systems in Texas, Oregon and California have either dropped SAT requirements or are considering doing so.

On the Wrong Side of the Tracks?

High-stakes tests create enough barriers for students of color, but when schools use standardized tests to assign students to different "tracks," they raise the walls even higher.

"Tracking" is the practice of placing students in different classes based on perceived differences in their abilities. A related practice, "ability" grouping, refers to the separation of students within the same classes into different groups, again based on perceived—although not necessarily real—differences in their abilities. Tracking takes a variety of forms, such as remedial and special education programs, as well as programs for gifted and talented students. At the high school level, many school systems distinguish between college preparatory and vocational tracks. In general, African American and Latino students are under represented in gifted and college-prep tracks, and over represented in remedial and vocational tracks. Today's programs for students described as "gifted and talented" generally provide enriched curricula, with an emphasis on development of higher cognitive skills. For example, in a gifted English class, the curriculum might include a focus on literary such devices as metaphor or irony and might require students to write analytic papers. Remedial classes by comparison would emphasize basic reading skills using repetition and drills, and might require only one-paragraph "essays." In math, remedial courses focus on basic arithmetic and culminate with basic algebra, whereas gifted programs extend into geometry, trigonometry, and eventually, calculus.

While there is little doubt that students who have been placed in programs for the gifted benefit from enriched curricula, there is significant evidence that students in other tracks would benefit equally from this teaching approach, and from the expectation of high achievement that accompanies it.

Tracking-which can determine a student's entire academic career-can begin as early as second or third grade. Tracking decisions are usually based on some combination of standardized testing, teacher recommendation, and parental intervention. As we have seen, many assessment tests are problematic because of potential racial and cultural bias, but subjective criteria, such as teacher recommendations, also create inequalities. In some cases, students of similar ability end up in different tracks based on little more than a teacher's impressions. In others, a determined parent's petition can win a child's placement in a gifted program. Basing admission to gifted programs on parental intervention can put low-income parents and those of color at a disadvantage, compared to their counterparts who have greater resources and are more used to "working the system." This is especially true for students whose parents do not speak much or any English. Assignment to remedial classes can be equally arbitrary. These classes often become all-purpose catchalls for emotionally troubled students, special needs children, and other students on the margins, regardless of their actual academic abilities. For example, bilingual students are often shunted into remedial classes under the assumption that because they only speak a little English, they would also struggle with the content matter of other subjects such as mathematics and science.

In theory, remedial classes are designed to help students fill in particular knowledge gaps. Once remedial work is completed, students rejoin the mainstream classes. Studies show, however, that it is very difficult for children assigned to a lower track to move into a higher one. In fact, rather than catching up with the mainstream, students in lower tracks generally fall further behind their upper-track peers every year they are in school.

Because each level of schooling builds on earlier prerequisites, students assigned to lower tracks in elementary or middle school have little opportunity to take the advanced courses, especially in math and science, required for acceptance by most major universities. For example, whether or not a student takes algebra in the eighth grade determines whether he or she will qualify to study calculus in the twelfth grade.

- By seventh grade, two-thirds of all schools have so-called ability grouping in some classes.
- By seventh grade, 20% of schools have tracking or grouping in every subject.
- Tracking is more common in schools with significant numbers of African American and/or Latino students than in schools that are predominantly white.

There are few national studies of the racial implications of tracking, although numerous local studies exist. For example, Professor Jeannie Oakes of the University of California at Los Angeles looked at school districts in Rockford, Illinois and San Jose, California. In each case, she found that "African American and Latino students were much less likely than white or Asian students with the same test scores to be placed in accelerated courses." (Oakes 1995)

The most extensive national survey of the racial demographics of tracking was conducted by researchers Jomills Henry Braddock II and Marvin Dawkins. (Braddock and Dawkins 1993) Braddock and Dawkins analyzed data contained in the National Center for Education Statistics' 1998 Longitudinal Study, which follows a cohort of students who were eighth graders in 1988. Their analysis of middle schools shows that in high tracks, all other students (excepting Asian Americans) were under represented compared to white students. In the lower tracks, all students of color were over represented, especially African Americans, Native Americans, and Latinos. Proportionally more than twice as many African American as white students were assigned to lower tracks in English and Math, while less than half as many were placed in higher tracks.

In the same study, Braddock and Dawkins examined mixed-ability classes (i.e., non-tracked classrooms) and found that in general, "students who experience low eighth-grade ability-group placements are significantly less likely to enter high school college preparatory programs than are eighth graders who were placed in mixed-ability classes." In high school, African American, American Indian, and Latino students were highly over represented in the vocational tracks and under represented in the academic tracks. Local studies in Selma, Alabama, Rockford, Illinois, Boston, Massachusetts, Johnston, North Carolina, and San Jose, California, among others, have produced similar results.

Tracking is less universally accepted as a "common sense" teaching method now than it was 20 years ago. Unfortunately, the current vogue for high-stakes tests, especially exit exams, threatens to reverse that trend. Policy Recommendations

Multiple levels, multiple solutions. Just as the systems that govern public schools operate at multiple levels, so must advocates for racial equality be prepared to work at multiple levels.

Classrooms and Schools:

- Never use standardized tests that are known to produce a racial outcome bias to determine the futures of individual students.
- Eliminate tracking based on ability assessments. Mixing students of differing level of achievement helps faster and slower students learn better. Substitute heterogeneous work groups for "ability-based" groups in individual classes.
- Never use tests designed to serve one purpose for another purpose. In particular, do not base admission to "gifted and talented" classes, or promotion, retention and graduation decisions, on tests designed to measure the progress of an entire class or school.
- Base student grades on portfolio content, in addition to tests. Teacher opinions and the body of a student's work should weigh toward that student's chances for academic advancement.

School Districts:

- All students should have access to classes that prepare them for college. Courses that are prerequisites for college admission—such as ninth grade Algebra—should be provided to all students, with assistance and tutoring for those who need it.
- Guarantee that all high schools offer a full range of academic courses, including Advanced Placement courses. In many districts, the schools attended by students of color do not provide these college-preparatory offerings, while schools serving primarily white students do.

State Governments and School Boards:

- Eliminate high-stakes high school exit exams. These tests do nothing to improve the quality of education students receive, and they deny high school diplomas to a disproportionate number of students of color who have completed all other graduation requirements.
- Eliminate SAT scores as an entrance requirement to state universities. Rely instead on high school records, which are more predictive of success in college, and less likely to produce racially biased results.

Federal Government:

- Offer financial incentives to districts to assist them in instituting and evaluating "detracking" programs.

Sources

Braddock, Jomills Henry and Dawson, Marvin, "Ability Grouping, Aspirations and Attainments: Evidence From the National Educational Longitudinal Study of 1988," *Journal of Negro Education*, Vol. 62, No. 3, 1993

Crouse, James and Trushein, Dale, *The Case Against the SAT* (Chicago: Chicago University Press, 1988)

Debra P. v. Turlington, 564 F. Supp. 177 (M.D. Fla. 1983)

Kunen, James S., "The Test of Their Lives: Exit Exams May Whip America's Schools Into Shape, but Some Good Kids Pay a Heavy Price" (*Time*, June 16, 1997, Vol. 149, No. 24)

MALDEF, Complaint brought before the United States District Court for the Western District of Texas, San Antonio Division, 1997

National Governors Association, NGA Reports Online, "High School Exit Exams: Setting High Expectations" (National Governors Association, September 1, 1998)

Oakes, Jeannie, "Two Cities' Tracking and Within-School Segregation," *Teachers College Record*, Vol. 96, No. 4, Summer 1995

Weissglass, Julian "The Sat: Public-Spirited or Preserving Privilege?" (*Education Week*, April 15, 1998)

Wheelock, Anne, *Crossing the Tracks: How "Untracking" Can Save America's Schools* (New York: New Press, 1992)

The ERASE Initiative is a national program which challenges racism in public schools and promotes racial justice and academic excellence for all students. The program is coordinated by the Applied Research Center, a public policy, education, and research institute which emphasizes issues of race and social change.

CARE Takes on Tracking in Alabama

In the late 1980s, Rose Sanders, an African American civil rights attorney in Selma, Alabama, discovered that her daughter had been placed in the bottom of three "tracks" of her public elementary school. She also found that once students were in this track, their educational opportunities were both limited and set in stone. For example, her child and many others would never be permitted to take algebra, a class Sanders considered important for a true education. In addition, without algebra on their transcripts, these students would be denied access to college.

When Sanders asked the school about her daughter's placement, district officials told her that students were "grouped by ability," but they could point to no specific indicators demonstrating that Sanders' daughter—who had tested as "gifted" in a privately administered test—was less able than others who were in the higher track. Nor could they explain why, in a school district that was approximately 60% African American and 40% white, the upper track was almost exclusively populated by white students. Indeed, at Selma High, 90% of the white students were placed in the high-ability track, compared to only 3% of African American students.

Sanders' discovery sparked a movement against tracking that gripped the Selma community, caused a sit-in by students that temporarily shut down the school system, and eventually forced the school board to pass a resolution ending tracking in the Selma schools. In 1993, Sanders co-founded CARE, the Coalition of Alabamians Reforming Education, to fight tracking and to educate communities in Alabama and elsewhere about the issue. The organization is currently working with two other school districts to "de-track" their schools.

The Fight in Selma

To initiate her battle against tracking in the Selma schools, Sanders joined with other parents to form BEST, the Best Education Support Team. At BEST meetings, African American parents described how their children had achieved the same scores as white children but had been placed in lower levels because the teacher said they weren't mature enough. Others spoke of how the most talented teachers were all teaching students in the top track. BEST began pressuring the school district to do away with tracking, while Sanders also started challenging the system by visiting the schools and encouraging students in the lower track to sign up for algebra classes. She worked with a youth group to organize high school students to protest tracking.

After a year of pressure, Superintendent of Schools, Roussell, an African American, acknowledged that the tracking system was flawed. He suggested a program of minor changes to the tracking system, such as opening honors classes and algebra to a greater number of students.

The white community of Selma denounced Superintendent Roussell's plan as one that would lower standards and lead to white flight from the schools. In early 1990, the school board, a body appointed by the city council, fired Roussell over the protests of its African American members. This was the final straw: African Americans in Selma boycotted the businesses of school board members, students conducted a one-day boycott of the schools, and Sanders and other parents paid a visit to the mayor. After the mayor had kept them waiting for hours and called the police, Sanders tried to enter his office, but she was dragged away by the police and later hospitalized for cuts and bruises. Two other people were arrested. That night, more than 2,000 people jammed into a school board meeting, and a few days later, 150 students staged a sit-in at Selma High School. In response, the school board shut down the schools for five days and agreed to keep Roussell on as superintendent until the end of the school year.

Although the school board continued to deny that race or racism played a role in tracking or in its decision to fire Roussell, the city was still acutely aware of its reputation as the site of a vicious police attack on civil

rights marchers on the Edmund Pettus Bridge in 1965. When the 1990 protests led to national attention, including a 60 Minutes segment, the board ultimately passed a referendum that it would no longer use tracking. "They were embarrassed," reflects April England-Albright, the current chair of CARE, which is BEST's successor organization.

The Selma protests led to victory in a number of respects: they brought national attention to the problem of tracking; they created an anti-tracking movement in Selma that could ignite and serve as a model for protests elsewhere; and they won the school board resolution.

CARE educates and agitates

In 1993, Rose Sanders and others formed CARE, the Coalition of Alabamians Reforming Education, to continue the battle that began in Selma. Since its inception, CARE has conducted education summits in more than 22 communities across the state. During trainings, CARE shares its lessons from the Selma experience and discusses tactics that communities can use to educate and organize people on the issue, including conducting door-to-door outreach, speaking at churches, attending school board and city council meetings, and conducting direct action protests like those in Selma.

In addition to its work educating communities, CARE has established a significant presence at the state level, through anti-tracking legislation and through pressuring Alabama's elected State Board of Education. Recognizing the interconnections between tracking and testing, the organization is currently using Alabama's school exit exam to highlight the inequities of the system and pressure the State Board of Education. "While you might think that requiring students to read at an eleventh grade level is a good thing, the problem is that many students haven't been adequately taught, so they're failing the exam," says England-Albright. "They've finished their high school requirements and they can't go back to school to learn more, but they're being denied their diplomas. We're saying that you shouldn't have an exit exam if you don't teach students the skills they need to pass it. They're testing all the students at the same level, but they're not preparing them all at the same level."

For more information, write to: CARE, P.O. Box 2516, Selma, AL 36702.

See also the [Spring 1999 issue of ColorLines](#), featuring a special section on *Race and Education*.

ColorLines: Race, Culture, Action

The nation's leading magazine on race, culture, and organizing.