

Standardized Testing and Its Victims

By Alfie Kohn

Standardized testing has swelled and mutated, like a creature in one of those old horror movies, to the point that it now threatens to swallow our schools whole. (Of course, on “The Late, Late Show,” no one ever insists that the monster is really doing us a favor by making its victims more “accountable.”) But let’s put aside metaphors and even opinions for a moment so that we can review some indisputable facts on the subject.

Fact 1. Our children are tested to an extent that is unprecedented in our history and unparalleled anywhere else in the world. While previous generations of American students have had to sit through tests, never have the tests been given so frequently, and never have they played such a prominent role in schooling. The current situation is also unusual from an international perspective: Few countries use standardized tests for children below high school age—or multiple-choice tests for students of any age.

Fact 2. Noninstructional factors explain most of the variance among test scores when schools or districts are compared. A study of math results on the 1992 National Assessment of Educational Progress found that the combination of four such variables (number of parents living at home, parents’ educational background, type of community, and poverty rate) accounted for a whopping 89 percent of the differences in state scores. To the best of my knowledge, all such analyses of state tests have found comparable results, with the numbers varying only slightly as a function of which socioeconomic variables were considered.

Fact 3. Norm-referenced tests were never intended to measure the quality of learning or teaching. The Stanford, Metropolitan, and California Achievement Tests (SAT, MAT, and CAT), as well as the Iowa and Comprehensive Tests of Basic Skills (ITBS and CTBS), are designed so that only about half the test-takers will respond correctly to most items. The main objective of these tests is to rank, not to rate; to spread out the scores, not to gauge the quality of a given student or school.

Fact 4. Standardized-test scores often measure superficial thinking. In a study published in the *Journal of Educational Psychology*, elementary school students were classified as “actively” engaged in learning if they asked questions of themselves while they read and tried to connect what they were doing to past learning; and as “superficially” engaged if they just copied down answers, guessed a lot, and skipped the hard parts. It turned out that high scores on both the CTBS and the MAT were more likely to be found among students who exhibited the superficial approach to learning. Similar findings have emerged from studies of middle school students (also using the CTBS) and high school students (using the other SAT, the college-admission exam). To be sure, there are plenty of students who think deeply and score well on tests—and plenty of students who do neither. But, as a rule, it appears that standardized-test results are positively correlated with a shallow approach to learning.

Fact 5. Virtually all specialists condemn the practice of giving standardized tests to children younger than 8 or 9 years old. I say “virtually” to cover myself here, but, in fact, I have yet to find a single reputable scholar in the field of early-childhood education who endorses such testing for young children.

Fact 6. Virtually all relevant experts and organizations condemn the practice of basing important decisions, such as graduation or promotion, on the results of a single test. The National Research Council takes this position, as do most other professional groups (such as the American Educational Research Association and the American Psychological Association), the generally pro-testing American Federation of Teachers, and even the companies that manufacture and sell the exams. Yet just such high-stakes testing is currently taking place, or scheduled to be introduced soon, in more than half the states.

Fact 7. The time, energy, and money that are being devoted to preparing students for standardized tests have to come from somewhere. Schools across the country are cutting back or even eliminating programs in the arts, recess for young children, electives for high schoolers, class meetings (and other activities intended to promote social and moral learning), discussions about current events (since that material will not appear on the test), the use of literature in the early grades (if the tests are focused narrowly on decoding skills), and entire subject areas such as science (if the tests cover only language arts and math). Anyone who doubts the scope and significance of what is being sacrificed in the desperate quest to raise scores has not been inside a school lately.

Fact 8. Many educators are leaving the field because of what is being done to schools in the name of “accountability” and “tougher standards.” I have no hard numbers here, but there is more than enough anecdotal evidence—corroborated by administrators, teacher-educators, and other observers across the country, and supported by several state surveys that quantify the extent of disenchantment with testing—to warrant classifying this as a fact. Prospective teachers are rethinking whether they want to begin a career in which high test scores matter most, and in which they will be pressured to produce these scores. Similarly, as the New York Times reported in its lead story of Sept. 3, 2000, “a growing number of schools are rudderless, struggling to replace a graying corps of principals at a time when the pressure to raise test scores and other new demands have made an already difficult job an increasingly thankless one.” It also seems clear that most of the people who are quitting, or seriously thinking about doing so, are not mediocre performers who are afraid of being held accountable. Rather, they are among the very best educators, frustrated by the difficulty of doing high-quality teaching in the current climate.

*Faced with inconvenient facts such as these, the leading fall-back position for defenders of standardized testing runs as follows: Even if it’s true that suburban schools are being dumbed down by the tests, inner-city schools are often horrendous to begin with. There, at least, standards are finally being raised as a result of high-stakes testing.

Let’s assume this argument is made in good faith, rather than as a cover for pursuing a standards-and-testing agenda for other reasons. Moreover, let’s immediately concede the major premise here, that low-income minority students have been badly served for years. The problem is that the cure is in many ways worse than the disease—and not only because of the preceding eight facts, which remain both stubbornly true and painfully relevant to testing in the inner city. As Sen. Paul Wellstone, D-Minn., put it in a speech delivered last spring: “Making students accountable for test scores works well on a bumper sticker, and it allows many politicians to look good by saying that they will not tolerate failure. But it represents a hollow promise. Far from improving education, high-stakes testing marks a major retreat from fairness, from accuracy, from quality, and from equity.” Here’s why.

The tests may be biased. For decades, critics have complained that many standardized tests are unfair because the questions require a set of knowledge and skills more likely to be possessed by children from a privileged background. The discriminatory effect is particularly pronounced with norm-referenced tests, where the imperative to spread out the scores often produces questions that tap knowledge gained outside of school. This, as W. James Popham argues, provides a powerful advantage to students whose parents are affluent and well-educated. It's more than a little ironic to rely on biased tests to "close the gap" between rich and poor.

*Guess who can afford better test preparation. When the stakes rise, people seek help anywhere they can find it, and companies eager to profit from this desperation by selling test-prep materials and services have begun to appear on the scene, most recently tailoring their products to state exams. Naturally, affluent families, schools, and districts are better able to afford such products, and the most effective versions of such products, thereby exacerbating the inequity of such testing. Moreover, when poorer schools do manage to scrape together the money to buy these materials, it's often at the expense of books and other educational resources that they really need.

*The quality of instruction declines most for those who have least. Standardized tests tend to measure the temporary acquisition of facts and skills, including the skill of test-taking itself, more than genuine understanding. To that extent, the fact that such tests are more likely to be used and emphasized in schools with higher percentages of minority students (a fact that has been empirically verified) predictably results in poorer-quality teaching in such schools. The use of a high-stakes strategy only underscores the preoccupation with these tests and, as a result, accelerates a reliance on direct-instruction techniques and endless practice tests. "Skills-based instruction, the type to which most children of color are subjected, tends to foster low-level uniformity and subvert academic potential," as Dorothy Strickland, an African-American professor at Rutgers University, has remarked.

Again, there's no denying that many schools serving low-income children of color were second-rate to begin with. Now, however, some of these schools, in Chicago, Houston, Baltimore, and elsewhere, are arguably becoming third-rate as testing pressures lead to a more systematic use of low-level, drill-and-skill teaching, often in the context of packaged programs purchased by school districts. Thus, when someone emphasizes the importance of "higher expectations" for minority children, we might reply, "Higher expectations to do what? Bubble-in more ovals correctly on a bad test—or pursue engaging projects that promote sophisticated thinking?" The movement driven by "tougher standards," "accountability," and similar slogans arguably lowers meaningful expectations insofar as it relies on standardized testing as the primary measure of achievement. The more that poor children fill in worksheets on command (in an effort to raise their test scores), the further they fall behind affluent kids who are more likely to get lessons that help them understand ideas. If the drilling does result in higher scores, the proper response is not celebration, but outrage: The test results may well have improved at the expense of real learning.

*Standards aren't the main ingredient that's in low supply. Anyone who is serious about addressing the inequities of American education would naturally want to investigate differences in available resources. A good argument could be made that the fairest allocation strategy, which is only common sense in some countries, is to provide not merely equal amounts across schools and districts, but more for the most challenging student populations. This does happen in some states—by no means all—but, even when it does, the money is commonly offered as a short-term grant (hardly sufficient to compensate for years of inadequate funding) and is often earmarked for test preparation rather than for higher-quality teaching. Worse, high-stakes testing systems may provide more money to those already successful (for example, in the form of bonuses for good scores) and less to those whose need is greatest.

Many public officials, along with like-minded journalists and other observers, are apt to minimize the matter of resources and assume that everything deficient about education for poor and minority children can be remedied by more forceful demands that we “raise the bar.” The implication here would seem to be that teachers and students could be doing a better job but have, for some reason, chosen not to do so and need only be bribed or threatened into improvement. (In fact, this is the tacit assumption behind all incentive systems.) The focus among policymakers has been on standards of outcome rather than standards of opportunity.

To make matters worse, some supporters of high-stakes testing have not just ignored, but contemptuously dismissed, the relevance of barriers to achievement in certain neighborhoods. Explanations about very real obstacles such as racism, poverty, fear of crime, low teacher salaries, inadequate facilities, and language barriers are sometimes written off as mere “excuses.” This is at once naive and callous, and, like any other example of minimizing the relevance of structural constraints, ultimately serves the interests of those fortunate enough not to face them.

*Those allegedly being helped will be driven out. When rewards and punishments are applied to educators, those who teach low-scoring populations are the most likely to be branded as failures and may decide to leave the profession. Minority and low-income students are disproportionately affected by the incessant pressure on teachers to raise scores. But when high stakes are applied to the students themselves, there is little doubt about who is most likely to be denied diplomas as a consequence of failing an exit exam—or who will simply give up and drop out in anticipation of such an outcome. If states persist in making a student’s fate rest on a single test, the likely result over the next few years will be nothing short of catastrophic. Unless we act to stop this, we will be facing a scenario that might be described without exaggeration as an educational ethnic cleansing.

Let’s be charitable and assume that the ethnic aspect of this perfectly predictable consequence is unintentional. Still, it is hard to deny that high-stakes testing, even when the tests aren’t norm-referenced, is ultimately about sorting. Someone unfamiliar with the relevant psychological research (and with reality) might insist that raising the bar will “motivate” more students to succeed. But perform the following thought experiment: Imagine that almost all the students in a given state met the standards and passed the tests. What would be the reaction from most politicians, businesspeople, and pundits? Would they now concede that our public schools are terrific—or would they take this result as *prima facie* evidence that the standards were too low and the tests were too easy? As Deborah Meier and others have observed, the phrase “high standards” by definition means standards that everyone won’t be able to meet.

The tests are just the means by which this game is played. It is a game that a lot of kids—predominantly kids of color—simply cannot win. Invoking these very kids to justify a top-down, heavy-handed, corporate-style, test-driven version of school reform requires a stunning degree of audacity. To take the cause of equity seriously is to work for the elimination of tracking, for more equitable funding, and for the universal implementation of more sophisticated approaches to pedagogy (as opposed to heavily scripted direct-instruction programs). But standardized testing, while bad news across the board, is especially hurtful to students who need our help the most.

Copyright © 2000 by Alfie Kohn. This article may be downloaded, reproduced, and distributed without permission as long as each copy includes this notice along with citation information (i.e., name of the periodical in which it originally appeared, date of publication, and author’s name). Permission must be obtained in order to reprint this article in a published work or in order to offer it for sale in any form. Please write to the address indicated on the [Contact Us](#) page.

www.alfiekohn.org — © Alfie Kohn