
[This oped summarizes *Collateral Damage: How High-Stakes Testing Corrupts America's Schools*, by Nichols & Berliner, just published by Harvard Education Press. I haven't finished the book yet, but based on the early chapters, would highly recommend it.]

Published in Education Week: March 12, 2007
Commentary

High-Stakes Testing is Putting the Nation At Risk

By David C. Berliner & Sharon L. Nichols

In his 2007 State of the Union address, President Bush claimed success for the federal No Child Left Behind Act. "Students are performing better in reading and math, and minority students are closing the achievement gap," he said, calling on Congress to reauthorize this "good law." Apparently, the president sees in No Child Left Behind what he sees in Iraq: evidence that his programs are working. But, as with Iraq, a substantial body of evidence challenges his claim.

We believe that this federal law, now in its sixth year, puts American public school students in serious jeopardy. Extensive reviews of empirical and theoretical work, along with conversations with hundreds of educators across the country, have convinced us that if Congress does not act in this session to fundamentally transform the law's accountability provision, young people and their educators will suffer serious and long-term consequences. If the title were not already taken, our thoughts on this subject could be headlined "A Nation at Risk."

We note in passing that only people who have no contact with children could write legislation demanding that *every* child reach a high level of performance in three subjects, thereby denying that individual differences exist. Only those same people could also believe that *all* children would reach high levels of proficiency at precisely the same rate of speed.

Validity problems in the testing of English-language learners and special education students also abound, but we limit our concerns in this essay to the No Child Left Behind law's reliance on high-stakes testing. The stakes are high when students' standardized-test performance results in grade retention or failure to graduate from high school. The stakes are high when teachers and administrators can lose their jobs or, conversely, receive large bonuses for student scores, or when humiliation or praise for teachers and schools occurs in the press as a result of test scores. This federal law requires such high-stakes testing in all states.

More than 30 years ago, the eminent social scientist Donald T. Campbell warned about the perils of measuring effectiveness via a single, highly consequential indicator: “The more any quantitative social indicator is used for social decisionmaking,” he said, “the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor.” High-stakes testing is exactly the kind of process Campbell worried about, since important judgments about student, teacher, and school effectiveness often are based on a single test score. This exaggerated reliance on scores for making judgments creates conditions that promote corruption and distortion. In fact, the overvaluation of this single indicator of school success often compromises the validity of the test scores themselves. Thus, the scores we end up praising and condemning in the press and our legislatures are actually untrustworthy, perhaps even worthless.

Campbell’s law is ubiquitous, and shows up in many human endeavors. Businesses, for example, regularly become corrupt as particular indicators are deemed important in judging success or failure. If stock prices are the indicator of a company’s success, for example, then companies like Enron, Qwest, Adelphia, and WorldCom manipulate that indicator to make sure they look good. Lives and companies are destroyed as a result. That particular indicator of business success became untrustworthy as both it and the people who worked with it were corrupted.

Similarly, when the number of criminal cases closed is the indicator chosen to judge the success of a police department, two things generally happen: More trials are brought against people who may be innocent or, with a promise of lighter sentences, deals are made with accused criminals to get them to confess to crimes they didn’t commit.

When the indicators of success and failure in a profession take on too much value, they invariably are corrupted. Those of us in the academic world know that when researchers are judged primarily by their publication records, they have occasionally fabricated or manipulated data. This is just another instance of Campbell’s law in action.

We have documented hundreds of examples of the ways in which high-stakes testing corrupts American education in a new book, *Collateral Damage*. Using Campbell’s law as a framework, we found examples of administrators and teachers who have cheated on standardized tests. Educators, acting just like other humans do, manipulate the indicators used to judge their success or failure when their reputations, employment, or significant salary bonuses are related to those indicators.

We found examples of administrators who would falsify school test data or force low-scoring students out of school in their quest to avoid public humiliation. We documented the distortion of instructional values when teachers focused on “bubble” kids—those on the cusp of passing the test—at the expense of the education of very low or very high scorers. We found instances where callous disregard for student welfare had replaced compassion and humanity, as when special education students were forced to take a test they had failed five times, or when a student who had recently suffered a death in the family was forced to take the test anyway.

Because so much depends on how students perform on tests, it should not be surprising that, as one Florida superintendent noted, “When a low-performing child walks into a classroom, instead of being seen as a challenge, or an opportunity for improvement, for the first time since I’ve been in education, teachers are seeing [that child] as a liability.” Shouldn’t we be concerned about a law that turns too many of the country’s most morally admired citizens into morally compromised individuals?

We also documented the narrowing of the curriculum to just what is tested, and found a huge increase in time spent in test preparation instead of genuine instruction. We found teachers concerned about their loss of morale, the undercutting of their professionalism, and the problem of disillusionment among their students. Teachers and administrators told us repeatedly how they were not against accountability, but that they were being held responsible for their students’ performance regardless of other factors that may affect it. Dentists aren’t held responsible for cavities and physicians for the onset of diabetes when youngsters don’t brush their teeth, or eat too much junk food, they argue.

Teachers know they stand a better chance of being successful where neighborhoods and families are healthy and communicate a sense of efficacy, where incomes are both steady and adequate, and where health-care and child-care programs exist. So the best of them soon move to schools with easier-to-teach students. This is no way to close the achievement gap.

Dozens of assessment experts have argued eloquently and vehemently that the high-stakes tests accompanying the implementation of the No Child Left Behind Act are psychometrically inadequate for the decisions that must be made about students, teachers, and schools. Furthermore, the testing standards of the American Educational Research Association are being violated in numerous ways by the use of high-stakes tests to comply with the law. The law, therefore, makes all who engage in compliance activities traitors to their own profession. It forces education professionals to ignore the testing standards that they have worked so hard to develop.

We wonder, would the federal government treat members of the American Medical Association or the National Academy of Sciences with such disdain?

In reauthorization hearings for the law, members of Congress should abandon high-stakes testing and replace it with an accountability system that is more reasonable and fair.

What might such a system look like?

A move to more “formative” assessments and an abandonment of our heavy commitment to “summative” assessments would be welcome. Assessment *for* learning, as opposed to assessment *of* learning, has produced some impressive gains in student achievement in other countries, and ought to be tried here. Likewise, the use of an inspectorate—an agency that sends expert observers into schools—has proved itself useful in other countries, and could also help improve schools in the United States.

End-of-course exams designed by teachers, as some states are now offering, increase teachers’ commitment to the testing program and, if the teachers get to score the tests, can also be a great professional-development opportunity. There are other alternatives to high-stakes testing, as well.

Our research informs us that high-stakes testing is hurting students, teachers, and schools. It is putting the nation at risk. By restricting the education of our young people and substituting for it training for performing well on high-stakes examinations, we are turning America into a nation of test-takers, abandoning our heritage as a nation of thinkers, dreamers, and doers.

David C. Berliner is the Regents’ professor of education at Arizona State University, in Tempe, and a past president of the American Educational Research Association. Sharon L. Nichols is an assistant professor of educational psychology at the University of Texas at San Antonio. They are the co-authors of *Collateral Damage: How High-Stakes Testing Corrupts America’s Schools*, published this month by Harvard Education Press.