

# Standardized Tests Do Not Effectively Measure Student Achievement

---

*Standardized Testing, 2012*

Phillip Harris is executive director of the Association for Educational Communications & Technology. For twenty-seven years, Bruce M. Smith was a member of the editorial staff of the *Phi Delta Kappan*, the flagship publication of Phi Delta Kappa International, the association for professional educators. He retired as editor-in-chief in 2008. Joan Harris has taught first, second, and third grades for more than twenty-five years. In 1997, she was recognized by the National Association for the Education of Young Children as the outstanding teacher of the year.

Contrary to popular assumptions about standardized testing, the tests do a poor job of measuring student achievement. They fail to measure such important attributes as creativity and critical thinking skills. Studies indicate that standardized tests reward superficial thinking and may discourage more analytical thinking. Additionally, because of the small sample of knowledge that is tested, standardized tests provide a very incomplete picture of student achievement.

Despite what reports in your local newspaper suggest, scores of standardized tests are *not* the same as student achievement. What's more, the scores don't provide very much useful information for evaluating a student's achievement, a teacher's competency, or the success of a particular school or program. To make such judgments, you need to move beyond the scores themselves and make some inferences about what they might mean....

The assumption underlying standardized testing ... is: *When we want to understand student achievement, it is enough to talk about scores on standardized tests.* Accepting this assumption at face value, as nearly all journalists, pundits, and politicians do, is to fall prey to a "dangerous illusion."

"Achievement" means more than a score on a standardized test.

## How to Define "Achievement"

---

Let's start with the question of defining achievement. If someone asked you to say in your own words exactly what is meant by "student achievement," how would you respond? If you said student achievement is what's measured by the state achievement tests, it's time to look a little harder at what these tests actually can and cannot do. More than a decade ago, education economist Richard Rothstein stated the problem directly: "Measurement of student achievement is complex—too complex for the social science methods presently available." And those methods certainly included standardized testing.

That was 1998, but the passage of more than a decade hasn't made it easier to evaluate student achievement in any systematic way, especially in a way that will yield the kind of numbers you can spread out along an axis to make comparisons. If anything, the intervening years—primarily the years of No Child Left Behind (NCLB) and its strict test-driven regimen—have made the problems in this area worse because we've asked test scores to carry ever more weight and we've depended on them to make ever more consequential decisions. Because of NCLB—and the [Barack] Obama administration's "blueprint" places similar weight on test scores—we now use "achievement test" scores to decide whether students

are entitled to tutoring services or whether they can transfer to a different school or whether we should close a school and reconstitute its staff. And many states now have strict rules about who qualifies to receive a high school diploma primarily by the scores on a standardized test of "achievement."

But "achievement" means more than a score on a standardized test. We knew it in 1998, and we know it now. For instance, as part of a larger project to ensure equity in math classrooms, the National Council of Teachers of Mathematics (NCTM), a group whose members are not strangers to the use of numerical data and statistical interpretation, reminded its members of some terms and definitions that would be important in the larger equity project. Rochelle Gutiérrez and her colleagues offered readers of the *NCTM News Bulletin* the following description of an appropriate understanding of "achievement": "Achievement—all the outcomes that students and teachers attain. Achievement is more than test scores but also includes class participation, students' course-taking patterns, and teachers' professional development patterns." The standardized tests we all know so well don't even come close to assessing all the outcomes that students and teachers attain.

## Many Attributes Cannot Be Measured

---

As psychometrician Daniel Koretz puts it, scores on a standardized test "usually do not provide a direct and complete measure of educational achievement." He cites two reasons why this is so, and both are related to our earlier discussion of sampling. First, tests can measure only a portion of the *goals* of education, which are necessarily broader and more inclusive than the test could possibly be.... Here is Gerald Bracey's list of some of the biggies that we generally don't even try to use standardized tests to measure:

creativity

critical thinking

resilience

motivation

persistence

curiosity

endurance

reliability

enthusiasm

empathy

self-awareness

self-discipline

leadership

civic-mindedness

courage

compassion

resourcefulness

sense of beauty

sense of wonder

honesty

integrity

Surely these are attributes we all want our children to acquire in some degree. And while not all learning takes place in classrooms, these are real and valuable "achievements." Shouldn't schools pursue goals such as these for their students, along with the usual academic goals? Of course, a teacher can't really teach all of these things from a textbook. But, as Bracey points out, she can model them or talk with students about people who exemplify them. But she has to have enough time left over to do so after getting the kids ready for the standardized test of "achievement."

Standardized tests inadvertently create incentives for students to become superficial thinkers.

## A Reward for Shallow Thinking

---

In fact, there are more problems associated with the impact of standardized testing on "achievement" than simply the fact that the technology of the testing cannot efficiently and accurately measure some vitally important attributes that we all want our children to "achieve." Alfie Kohn put it this way:

Studies of students of different ages have found *a statistical association between students with high scores on standardized tests and relatively shallow thinking*. One of these studies classified elementary school students as "actively engaged if they went back over things they didn't understand, asked questions of themselves as they read, and tried to connect what they were doing to what they had already learned; and as 'superficially' engaged if they just copied down answers, guessed a lot, and skipped the hard parts. It turned out that the superficial style was positively correlated with high scores on the Comprehensive Test of Basic Skills (CTBS) and Metropolitan Achievement Test (MAT). Similar findings have emerged from studies of middle school and high school students."(Emphasis in original.)

So by ignoring attributes that they can't properly assess, standardized tests inadvertently create incentives for students to become superficial thinkers—to seek the quick, easy, and obvious answer. That's hardly an "achievement" that most parents want for their children. And surely it's not what our policy makers and education officials hope to achieve by incessantly harping on "achievement." In our view, most of these policy makers mean well, but when they say "achievement," they clearly mean test scores and only test scores. But to assume that the test scores can take the place of all the other information we need to know in order to have a good understanding of students' development leads us to some poor conclusions about how our children are growing physically, emotionally, and intellectually. The information provided by test scores is very limited, and consequently we must be very careful in drawing inferences about what the

scores mean.

## Measuring Only Small Samples

---

The second reason Koretz cites for the incompleteness of test scores as a measure of achievement [is as follows]: "Even in assessing the goals that can be measured well, tests are generally very small samples of behavior that we use to make estimates of students' mastery of very large domains of knowledge and skill." So apart from not doing a very good job of measuring achievement in such areas as creativity or persistence, standardized tests have another serious limitation: whenever a small part of a domain is made to stand in for the larger whole, we must be very careful about the inferences we draw from the data we obtain.

---

## Further Readings

---

### Books

- Gerald W. Bracey *Education Hell: Rhetoric vs. Reality*. Alexandria, VA: Educational Research Service, 2009.
- John Cronin et al. *The Proficiency Illusion*. Washington, DC: Thomas B. Fordham Institute and Northwest Evaluation Association, 2007.
- Linda Darling-Hammond *The Flat World and Education: How America's Commitment to Equity Will Determine Our Future*. New York: Teachers College Press, 2010.
- Daniel Koretz *Measuring Up: What Educational Testing Really Tells Us*. Cambridge, MA: Harvard University Press, 2008.
- Jennifer McMurrer *Choices, Changes, and Challenges: Curriculum and Instruction in the NCLB Era*. Washington, DC: Center on Education Policy, 2007.
- Linda Perlstein *Tested: One American School Struggles to Make the Grade*. New York: Henry Holt, 2007.
- Richard Rothstein, Rebecca Jacobsen, and Tamara Wilder *Grading Education: Getting Accountability Right*. Washington, DC: Economic Policy Institute and Teachers College Press, 2008.
- J. Michael Spector et al. *Learning and Instruction in the Digital Age: Making a Difference through Cognitive Approaches*. New York: Springer, 2010.
- Matthew G. Springer, ed. *Performance Incentives: Their Growing Impact on American K-12 Education*. Washington, DC: Brookings Institution Press, 2009.
- Herbert J. Walberg *Tests, Testing, and Genuine School Reform*. Washington, DC: Hoover Institution Press, 2011.

### Periodicals and Internet Sources

- Kam Wing Chan "Despite Recent Test Scores, China Is Not 'Eating Our Lunch,'" *The Seattle Times*, January 2, 2011.
- Linda Darling-Hammond "Recognizing and Enhancing Teacher Effectiveness," *International Journal of Educational and Psychological Assessment*, December 2009.
- Arne Duncan "Despite Cheating Scandals, Testing and Teaching Aren't at Odds," *The Washington Post*, July 19, 2011.

- Susan Engel "Scientifically Tested Tests," *The New York Times*, September 19, 2010.
- James Fallows "On Those 'Stunning' Shanghai Test Scores," *The Atlantic*, December 7, 2010.
- Chester E. Finn Jr. "A Sputnik Moment for U.S. Education," *The Wall Street Journal*, December 8, 2010.
- Dan Fletcher "Standardized Testing," *Time*, December 11, 2009.
- Christopher Harper and Robert J. Vanderbei "Two Professors Retake the SAT: Is It a Good Test?" *The Chronicle of Higher Education*, June 12, 2009.
- Oliver Lewis "The Gaokao Challenge: China's Fearsome University Entrance Exam Shows Us Why British Schools Are Falling Behind," *Spectator*, December 11, 2010.
- National Association for College Admission Counseling "Report of the Commission on the Use of Standardized Tests in Undergraduate Admission," September 2008. [www.nacacnet.org](http://www.nacacnet.org).
- Kristina Rizga "What Standardized Tests Miss," *Mother Jones*, May 19, 2011.
- Peter Z. Schochet and Hanley S. Chiang "Error Rates in Measuring Teacher and School Performance Based on Student Test Score Gains," National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, and US Department of Education, July 2010. <http://ies.ed.gov/ncee>.
- Sue Shellenbarger "Education Reform and Leaving No Child Left Behind," *The Wall Street Journal*, March 25, 2010.
- Bobby Ann Starnes "Twisted NCLB or Twisting NCLB?" *Phi Delta Kappan*, December 2007.
- John R. Tanner "Incomplete Measures: In K-12 Accountability, Are We Answering the Wrong Questions Well and the Right Questions Poorly?" *School Administrator*, February 2010.
- Phyllis Tashlik "Changing the National Conversation on Assessment: A Consortium of New York Public High Schools Serves as a Model for a Multi-Dimensional System That Is Performance-Based," *Phi Delta Kappan*, March 2010.
- Marcus Winters "Costly, but Worth It," *The New York Times*, May 30, 2011.

**Full Text:** COPYRIGHT 2012 Greenhaven Press, a part of Gale, Cengage Learning.

---

### Source Citation

Harris, Phillip, Joan Harris, and Bruce M. Smith. "Standardized Tests Do Not Effectively Measure Student Achievement." *Standardized Testing*. Ed. Dedria Bryfonski. Detroit: Greenhaven Press, 2012. At Issue. Rpt. from "Chapter 3: The Tests Don't Measure Achievement Adequately." *The Myths of Standardized Tests: Why They Don't Tell You What You Think They Do*. 2011. 33-45. *Opposing Viewpoints In Context*. Web. 23 Sept. 2013.

### Document URL

[http://ic.galegroup.com/ic/ovic/ViewpointsDetailsPage/ViewpointsDetailsWindow?faIlOverType=&query=&prodId=OVIC&windowstate=normal&contentModules=&mode=view&displayGroupName=Viewpoints&limiter=&currPage=&disableHighlighting=true&displayGroups=&sortBy=&source=&search\\_wthin\\_results=&zid=&action=e&catId=&activityType=&scanId=&documentId=GALE%7CEJ3010478218&userGroupName=summerchs&jsid=2992b6f68f71141961418b7f1073be07](http://ic.galegroup.com/ic/ovic/ViewpointsDetailsPage/ViewpointsDetailsWindow?faIlOverType=&query=&prodId=OVIC&windowstate=normal&contentModules=&mode=view&displayGroupName=Viewpoints&limiter=&currPage=&disableHighlighting=true&displayGroups=&sortBy=&source=&search_wthin_results=&zid=&action=e&catId=&activityType=&scanId=&documentId=GALE%7CEJ3010478218&userGroupName=summerchs&jsid=2992b6f68f71141961418b7f1073be07)

