

The New York Times
Opinionator

SEPTEMBER 21, 2011, 8:15 PM

What Do Test Scores Tell Us?

By [GARY GUTTING](#)

[The Stone](#) is a forum for contemporary philosophers on issues both timely and timeless.

Tags:

[Education](#), [Philosophy](#), [testing](#)

The Stone is featuring occasional posts by Gary Gutting, a professor of philosophy at the University of Notre Dame, that apply critical thinking to information and events that have appeared in the news.

Tests used to be just for evaluating students, but now the testing of students is used to evaluate teachers and, in fact, the entire educational system. On an individual level, some students and parents have noticed a change — more standardized tests and more classroom and homework time devoted to preparation for them.

So what exactly do test scores tell us?

Poor test scores are the initial premises in most current arguments for educational reform. At the end of last year, [reading scores](#) that showed American 15-year-olds in the middle of an international pack, led by Asian countries, prompted calls from researchers and educators for immediate action. This year two sociologists, [Richard Arum and Josipa Roksa](#), showed that 45 percent of students, after two years of college, have made no significant gains on a test of critical thinking. Last week's report of [falling SAT scores](#) is the latest example.

Given poor test results, many critics conclude that our schools are failing and propose plans for immediate action. For example, when Arum and Raksa published their results, many concluded that college teachers need to raise standards in their courses, requiring more hours of study and assigning longer papers.

It is, however, not immediately obvious what follows from poor test scores. Without taking any position about the state of our schools or how, if at all, they need reform, I want to reflect on what we need to add to the fact of poor scores to construct an argument for changing the way we educate.

The first question is whether a test actually tests for things that we want students to know. We very seldom simply want students to do well on a test for its own sake. Tests must require or demonstrate some valuable knowledge or skill, like knowing how to multiply, understanding the Civil War or being able to think critically. It is entirely possible for students to fail tests on such topics and still have, for example, the mathematical abilities or historical knowledge we want. If, for example, a math test required mentally multiplying 392×654 in five seconds or a history test required knowing the precise date of 50 Civil War battles, we would not be concerned that students did poorly.

This problem becomes especially serious when we are testing a complex skill such as critical thinking. The test of college students' critical thinking that Arum and Raksa used involved a simulated "performance task"; for example, weighing evidence to determine if a fictional company should buy a type of airplane that had recently crashed. Do we in fact care whether students can think critically about artificial cases on topics that may be outside their experience and interest? Perhaps so, but then we need to show that doing well with such cases corresponds to an ability to think critically in the real world.

In my own teaching, I've noticed that students who have great difficulty arguing cogently about philosophical questions I raise in class nevertheless develop very sophisticated cases for being allowed to turn a paper in late. I've learned that if I can get them interested in those philosophical questions, the critical skills follow rather quickly. This simple example is not meant to refute the validity of the test of critical thinking Arum and Ruska used. It merely illustrates the need for thorough reflection on just what doing poorly on a test means.

There is also the question of whether any standardized test is adequate or needed to evaluate certain sorts of student learning. There was a time when we were happy with Miss Goodteacher's judgment that her class knew how to read. There are doubtless cases where we can't trust instructors' judgments. But is there reason, especially in college-level work, to think that this is generally the case?

When a test does show that students lack some knowledge or skill we want them to have, there remains the question of what to do about it. Here many of us have an unfortunate tendency to think there's a panacea that could cure all our educational ills: well-paid teachers, small classes, better methods of instruction, following the Japanese model, etc. It's much more likely that, as in cancer treatment, we need to proceed piecemeal, with different solutions for different problems.

Even then, there's no knowing what an adequate solution might cost. So we also have to think hard about whether we are able or willing to provide the resources needed to meet all our many educational goals. We may also eventually need to think the unthinkable: that some students may simply be incapable of learning at the level we would like.

Negative test results should never be presented as reasons for immediate and drastic action. When our students fail standardized tests, we need to work through many further difficult steps before concluding what, if anything, needs to be done. There is a long argumentative path from bad tests results to a justified call to arms.

Copyright 2011 The New York Times Company | [Privacy Policy](#) | [NYTimes.com](#) 620 Eighth Avenue New York, NY 10018