Whoever Said There's No Such Thing As a Stupid Question Never Looked Carefully at a Standardized Test (##)

Whoever Said There's No Such Thing As a Stupid Question Never Looked Carefully at a Standardized Test

By Alfie Kohn

It can't be repeated often enough: Standardized tests are very poor measures of the intellectual capabilities that matter most, and that's true because of how they're designed, not just because of how they're used. Like other writers, I've relied on arguments and research to make this point. But sometimes a telling example can be more effective. So here's an item that appeared on the state high school math exam in Massachusetts:

n 1 2 3 4 5 6

t_n 3 5 __ __ __ __

The first two terms of a sequence, t_1 and t_2, are shown

above as 3 and 5. Using the rule: $t_n = (t_{n-1})$ plus (t_{n-2}), where n is greater than or equal to 3, complete the table.

If (a) your reaction to this question was "Huh??" (or "Uh-oh. What's with the teeny little n's?") and (b) you lead a reasonably successful and satisfying life, it may be worth pausing to ask why we deny diplomas to high school students just because they, too, struggle with such questions. Hence [Deborah] Meier's Mandate: "No student should be expected to meet an academic requirement that a cross-section of successful adults in the community cannot."

But perhaps you figured out that the test designers are just asking you to add 3 and 5 to get 8, then add 5 and 8 to get 13, then add 8 to 13 to get 21, and so on. If so, congratulations. But what is the question really testing? A pair of math educators, Al Cuoco and Faye Ruopp, pointed out how much less is going on here than meets the eye:

> *The problem simply requires the ability to follow a rule; there is no mathematics in it at all. And many 10th-grade students will get it wrong, not because they lack the mathematical thinking necessary to fill in the table, but simply because they haven't had experience with the notation. Next year, however, teachers will prep students on how to use formulas like $t_n = t_{n-1} + t_{n-2}$, more students will get it right, and state education officials will tell us that we are increasing mathematical literacy.[1]*

In contrast to most criticisms of standardized testing, which look at tests in the aggregate and their effects on entire populations, this is a bottom-up critique. Its impact is to challenge not only the view that such tests provide "objective" data about learning but to jolt us into realizing that high scores are not necessarily good news and low scores are not necessarily bad news.

If the questions on a test measure little more than the ability to apply an algorithm mindlessly, then you can't use the results of that test to make pronouncements about this kid's (or this school's, or this state's, or this country's) proficiency at mathematical thinking. Similarly, if the questions on a science or social studies test mostly gauge the number of dates or definitions that have been committed to memory — and, perhaps, a generic skill at taking tests — it would be foolish to draw conclusions about students' understanding of those fields.

A parallel bottom-up critique emerges from interviewing children about why they picked the answers they did on multiple-choice exams — answers for which they received no credit — and discovering that some of their reasons are actually quite sophisticated, which of course one would never know just by counting the number of their "correct" answers.[2]

No newspaper, no politician, no parent or school administrator should ever assume that a test score is a valid and meaningful indicator without looking carefully at the questions on that test to ascertain that they're designed to measure something of importance and that they do so effectively. Moreover, as Cuoco and Ruopp remind us, rising scores over time are often nothing to cheer about because the kind of instruction intended to prepare kids for the test — even when it does so successfully — may be instruction that's not particularly valuable. Indeed, teaching designed to raise test scores typically reduces the time available for real learning. And it's naïve to tell teachers they should "just teach well and let the tests take care of themselves." Indeed, if the questions on the tests are sufficiently stupid, bad teaching may produce better scores than good teaching does.

_____

Notes

1. Cuoco and Ruopp, "Math Exam Rationale Doesn't Add Up," *Boston Globe*, May 24, 1998, p. D3.

2. For examples (and analysis) of this kind of discrepancy, see Banesh Hoffmann, *The Tyranny of Testing* (New York: Crowell-Collier, 1962); Deborah Meier, "Why Reading Tests Don't Test Reading," *Dissent*, Fall 1981: 457-66; Walt Haney and Laurie Scott, "Talking with Children About Tests: An Exploratory Study of Test Item Ambiguity," in Roy O. Freedle and Richard P. Duran, eds., *Cognitive and Linguistic Analyses of Test Performance* (Norwood, NJ: Ablex, 1987); and Clifford Hill and Eric Larsen, *Children and Reading Tests* (Stamford, CT: Ablex, 2000).