

Why the Best Teachers Don't Give Tests

Frankly, I'm baffled by the number of educators who are adamantly opposed to standardized testing yet raise no objection to other practices that share important features with such testing.

For starters, consider those lists of specific, prescriptive curriculum standards to which the tests are yoked. Here we find the same top-down control and one-size-fits-all mentality that animate standardized testing. Yet from the early days of the "accountability" movement right down to current efforts to impose the Gates-funded Common Core from coast to coast, an awful lot of people give the standards (and the whole *idea* of uniform standards) a pass while frowning only at the exams used to enforce them.¹

Example 2: Elaborate rubrics used to judge students' performance represent another form of standardized assessment that's rarely recognized as such. The point is to break down something, such as a piece of writing, into its parts so that teachers, and sometimes the students themselves, can rate each of them, the premise being that it's both possible and desirable for all readers to arrive at the same number for each criterion. Rubrics are borne of a demand to quantify and an impulse to simplify. One result, argues Maja Wilson, is that "the standardization of the rubric produces standardized writers."² But, again, even many teachers who are outraged by standardized tests don't blink when standardization is smuggled in through the back door. Some insist, against all evidence to the contrary, that there's no problem as long as one uses a *good* rubric.

¹"Why the Best Teachers Don't Give Tests" was originally published as a blog post on October 31, 2014.

It's my third example, though, on which I'd like to linger. When teachers test their students, the details of those tests will differ from one classroom to the next, which means these assessments by definition are not standardized and can't be used to compare students across schools or states. But they're still tests, and as a result they're still limited and limiting.

As with rubrics (and grades), there's a reflexive tendency to insist that we just need *better* tests, or that we ought to just modify the way they're administered (for example, by allowing students to retake them). And, yes, it's certainly true that some are worse than others. Multiple-choice tests are uniquely flawed as assessments for exactly the same reason that multiple-choice standardized tests are: They're meant to trick students who understand the concepts into picking the wrong answer, and they don't allow kids to generate, or even explain, their responses. Multiple-choice exams can be clever, but as test designer Roger Farr of Indiana University ultimately concluded, there is no way "to build a multiple choice question that allows students to show what they can do with what they know."

We can also concede that some *reasons* for giving tests are more problematic than others. There's a difference between using them to figure out who needs help—or, for more thoughtful teachers, what aspects of their own instruction may have been ineffective—and using them to compel students to pay attention and complete their assignments. In the latter case, a test is employed to pressure kids to do what they have little interest in doing. Rather than address possible deficiencies in one's curriculum or pedagogy (say, the exclusion of students from any role in making decisions about what they'll learn), one need only sound a warning about an upcoming test—or, in an even more blatant exercise of power, surprise students with a pop quiz—to elicit compliance.

Even allowing for variation in the design of the tests and the motives of the testers, however, the bottom line is that these instruments are typically more about measuring the number of facts that have been crammed into students' short-term memories than they are about assessing understanding.³ Tests, including those that involve essays, are part of a traditional model of instruction in which information is transmitted to students (by means of lectures and textbooks) so that it can be disgorged later on command. That's why it's so disconcerting to find teachers who are proud of their student-centered approach to instruction, who embrace active and interactive forms of learning, yet continue to rely on tests as the primary, or even sole, form of assessment in their classrooms.

While some of their questions may require problem-solving skills, tests, per se, are artificial pencil-and-paper exercises that measure how much students remember and how good they are at the discrete skill of taking tests. That's how it's possible for a student to be a talented thinker and yet score poorly. Most teachers can, without hesitation, name several such students in their classes when the exams are designed by Pearson or ETS but may fail to see that the same thing applies in the case of performance on tests they design themselves.

Not only do tests assess the intellectual proficiencies that matter least, however—they also have the potential to alter students' goals and the way they approach learning. The more you're led to focus on what you're going to have to know for a test, the less likely you are to plunge into a story or engage fully with the design of a project or experiment. And intellectual immersion can be all but smothered if those tests are given, or even talked about, frequently. Learning in order to pass a test is qualitatively different from learning for its own sake.⁴

Many years ago, the eminent University of Chicago educator Philip Jackson interviewed fifty teachers who had been identified as exceptional at their craft. Among his findings was a consistent lack of emphasis on testing, if not a deliberate decision to minimize the practice, on the part of these teachers.⁵ The first reason for this, I think, is that exemplary educators understand that tests are not a particularly useful form of assessment. Second, though, these teachers learned at some point that they didn't *need* tests. The most impressive classrooms and curricula are designed to help the teacher know as much as possible about how students are making sense of things. When kids are engaged in meaningful, active learning—for example, designing extended, interdisciplinary projects—teachers who watch and listen as those projects are being planned and carried out have access to, and actively interpret, a continuous stream of information about what each student is able to do and where he or she requires help. It would be superfluous to give students a test after the learning is done. We might even say that the more a teacher is inclined to use a test to gauge student progress, the more that tells us something is wrong—perhaps with the extent of the teacher's informal and informed observation, perhaps with the quality of the tasks, perhaps with the whole model of learning. If, for example, the teacher favors direct instruction, he or she probably won't have much idea what's going on in the

students' minds. That will lead naturally to the conclusion that a test is "necessary" to gauge how they're doing.⁶

Assessment literally means to sit beside, and that's just what our most thoughtful educators urge us to do. Yetta Goodman coined the compound noun "kidwatching" to describe reading with each child to gauge his or her proficiency. Marilyn Burns insists that one-on-one conversations tell us far more about students' mathematical understanding than a test ever could—since all wrong answers aren't alike. Of course this assumes that we're really interested in kids' understanding, not merely their level of phonemic awareness or ability to apply an algorithm. The less ambitious one's educational goals, the more likely that a test will suffice—and that the words *testing* and *assessing* will be used interchangeably.

One can fill a bookshelf with accounts of other forms of authentic assessment: portfolios, culminating projects, performance assessments, and what the late Ted Sizer called "exhibitions of mastery": opportunities for students to demonstrate their proficiency not by recalling facts on demand but by *doing* something: constructing and conducting (and explaining the results of) an experiment, creating a restaurant menu in a foreign language, turning a story into a play. In other words, when some form of evaluation is desired after, rather than during, the learning, tests *still* aren't necessary or even particularly helpful. They needn't be used for "summative," let alone for "formative," assessment.

Many of us rail against standardized tests not only because of the harmful uses to which they're put but because they're imposed on us. It's more unsettling to acknowledge that the tests we come up with ourselves can also be damaging. The good news is that far superior alternatives are available.

NOTES

1. See my essay "Beware of the Standards, Not Just the Tests," *Education Week*, September 26, 2001—available at www.alfiekohn.org/article/beware-standards-just-tests.

This phenomenon is even more pronounced in Canada. Its education system is completely decentralized; each province controls its own policies. Despite the considerable variation in the amount of testing from one to the next, however, all of the provinces have very specific grade-by-grade curricula that every teacher is