

Beware of the Standards, Not Just the Tests

By Alfie Kohn

A number of prominent educators are finally raising their voices against standardized testing—particularly multiple-choice, norm-referenced tests; particularly tests with “high stakes” (read: bribes and threats) attached; and particularly in the context of a federal mandate to force every state to test every student in grades 3-8 every year. Yet even as more opinion leaders come to understand the damage attributable to testing mania, it is still rare to hear objections to the standards movement as a whole.

The Learning First Alliance, a coalition of leading education groups, cautiously raised concerns about the tests not long ago, but mostly out of fear that the burgeoning grassroots opposition might bring down the state standards, too. Education Week’s 2001 edition of *Quality Counts* likewise worried that tests “are overshadowing” and “do not adequately reflect” the standards. Major conferences carry titles such as “Standards: From Theory to Practice” and “Will Standards Survive the Classroom?” (You will look in vain for conferences called “Will Classrooms Survive the Standards?” or “Standards: From Capitulation to Resistance.”)

A list of boat-rocking books on the subject begins and pretty much ends with Susan Ohanian’s *One Size Fits Few* and Deborah Meier’s *Will Standards Save Public Education?* Alarms have been quietly raised by Nel Noddings, Elliot Eisner, James Beane, and a few other eminent educators in the pages of *Phi Delta Kappan*. Otherwise, the field seems to have closed ranks around the idea that it is permissible to criticize the tests, but not the standards. Indeed, test opponents are sternly reminded to avoid confusing the two, as though they were in fact unrelated. I want to argue not only that they are inextricably connected—the tests serving, at least in theory, as the enforcement mechanism of the standards—but also that the latter may be every bit as problematic as the former.

Of course, it’s reasonable to ask just what kind of standards are at issue here. The most relevant and widely accepted distinction is between outcome and content. Outcome standards specify how well students must do. At the highest level of generality (“We support high standards”), the notion is unobjectionable but not terribly useful. When translated into specifics, it comes to mean cut scores on standardized tests and becomes downright dangerous. Outcome standards to a remarkable extent are based on [confusing harder with better](#), an error I have already discussed in these pages and need not belabor.

Content standards, by contrast, specify what students will be taught. Rather than declaring that all such standards are bad—or, as is far more common, accepting all such standards uncritically—I propose that we judge a given set of standards or frameworks according to four criteria:

1. How specific? There are many reasons policymakers seek to impose detailed curriculum mandates. They may fundamentally distrust educators: Much of the current standards movement is

just the latest episode in a long, sorry history of trying to create a teacher-proof curriculum. Alternatively, they may simply assume that more specificity is always preferable. In reality, just because it makes sense to explain to a waiter exactly how I'd like my burger cooked doesn't mean it's better to declare that students will study the perimeter of polygons (along with scores of other particular topics) than it is to offer broad guidelines for helping students learn to think like mathematicians.

The latter sort of standards, supported by practical guidance, can help students reason carefully, communicate clearly, and get a kick out of doing so. But long lists of facts and skills that teachers must cover may have the opposite effect. Thus, when Harold Howe II, the U.S. commissioner of education under President Johnson, was asked what a set of national standards should be like (if we had to adopt them), he summarized a lifetime of wisdom in four words: they should be "as vague as possible."

His caution applies to state standards as well. On the one hand, thinking is messy, and deep thinking is very messy. On the other hand, standards documents are nothing if not orderly. Keep that contrast in mind and you will not be surprised to see how much damage those documents can do in real classrooms.

Considerable research has demonstrated the importance of making sure students are actively involved in designing their own learning, invited to play a role in formulating questions, creating projects, and so on. But the more comprehensive and detailed a list of standards, the more students (and even teachers) are excluded from this process, the more alienated they tend to become, and the more teaching becomes a race to cover a huge amount of material. Thus, meeting these kinds of standards may actually have the effect of dumbing down classrooms. As Howard Gardner and his colleagues wisely observed, "The greatest enemy of understanding is 'coverage.'"

Some insist that these lists of facts and skills don't prescribe how students will be taught; the standards are said to be neutral with respect to pedagogy. But this is nonsense. If the goal is to cover material (rather than, say, to discover ideas), that unavoidably informs the methods that will be used. Techniques such as repetitive drill-and-practice are privileged by curriculum frameworks based on a "bunch o' facts" approach to education. Of course, that kind of teaching is also driven by an imperative to prepare students for tests, but no less by an imperative to conform to specific standards.

Some people sincerely believe that to teach well is to work one's way through a list of what someone decided every nth grader ought to know. But the question is not just what you or I think about that model; the question is whether that model should enjoy virtual monopoly status in American public schools. In effect, one particular, very debatable philosophy of education has been enshrined in state standards documents and has become the law of the land.

2. How quantifiable? The current accountability fad insists on mandates that are not only overly detailed but chosen according to whether they lend themselves to easy measurement. It's not just that the tests are supposed to be tied to the standards; it's that the standards have been selected on the basis of their testability. The phrase "specific, measurable standards" suggests a commitment not to excellence but to behaviorism. It is telling that this phrase is heard most often from corporate officials and politicians, not from leading educational theorists or cognitive scientists.

We are talking about a worldview in which any aspect of learning, or life, that resists being reduced to numbers is regarded as vaguely suspicious. By contrast, anything that appears in numerical form seems reassuringly scientific; if the numbers are getting larger over time, we must be making progress. Concepts like intrinsic motivation and intellectual exploration are difficult for some minds

to grasp, whereas test scores, like sales figures or votes, can be calculated and charted and used to define success and failure.

Unfortunately, meaningful learning does not always proceed along a single dimension, such that we can nail down the extent of improvement. As Linda McNeil of Rice University has observed, "Measurable outcomes may be the least significant results of learning." (That sentence ought to be printed out in 36-point Helvetica, framed, and tacked to the wall of every school administrator's office in the country.) To talk about what happens in schools as moving forward or backward in specifiable degrees is not only simplistic, in the sense that it fails to capture what is actually going on; it is destructive, because it can change what is going on for the worse.

Consider a comment made by Sandra Stotsky, the deputy commissioner of education in Massachusetts: "Explore isn't a word that can be put into a standard because it can't be assessed." This assertion is obviously false because there are plenty of ways to assess the quality of students' exploration — unless, of course, "assessment" is equated with standardized testing. But suppose for the moment that Ms. Stotsky was correct. What if we really were faced with a trade-off between an emphasis on exploration in the classroom and the demands of measurement? Most thoughtful educators would unhesitatingly choose the former, whereas those who write and enforce state standards often opt for the latter. Clearly, it is much easier to quantify the number of times a semicolon has been used correctly in an essay than it is to quantify how well the student has explored ideas in that essay. Thus, the more emphasis that is placed on picking standards that are measurable, the less ambitious the teaching will become.

3. How uniform? We have heard the phrase "standardized testing" so often that we may have become inured to the significance of that first word. To what extent do we really want our students to receive a standardized education? At a national conference last fall, a consultant announced with apparent satisfaction that now, thanks to standards-based reform, "for the first time in my experience, people on a grade level, in a subject area, or teaching a course at a high school are [feeling] a responsibility to all have the same destination." That she did not even feel it necessary to defend this goal says something about the current acceptance of a one-size-fits-all model of education.

Once again, the problem is not just with the construction of the tests, but with the uniformity of the standards. Wanting to make sure that students in low-income communities don't receive a second-rate education is a laudable objective. Wanting to make sure that all students in your state receive the same education, such that they are treated as interchangeable recipients of knowledge, is a very different matter. Even more troubling are grade-by-grade standards. Here, the prescribers are not just saying, "We expect students to know the following stuff by the time they're in 8th grade," but "We expect them to learn all the items on this list in 5th grade, all the items on that list in 6th grade," and so on. Apart from the negative effects on learning, this rigidity about both the timing of the instruction and its content creates failures unnecessarily by trying to force all children to learn at the same pace.

4. Guidelines or mandates? There are standards offered as guidelines ("See if this way of thinking about teaching can help you improve your craft"), and then there are standards presented as mandates ("Teach this or else"). Virtually all the states have chosen the latter course. The effect has been not only to control teachers, but to usurp the long-established power of local school districts to chart their own course. If there has ever been a more profoundly undemocratic school reform movement in U.S. educational history than what is currently taking place in the name of standards, I haven't heard of it.

Bullying reaches its apotheosis with high-stakes testing, the use of crude rewards and punishments

to make people ratchet up the scores. The underlying logic is captured by an ironic sign spotted on a classroom wall: “The beatings will continue until morale improves.” But the standards themselves, if handed down as requirements, embody that same determination on the part of policymakers to do things to educators and students rather than to work with them. My nominee for the most chillingly Orwellian word now in widespread use is “alignment”—as in, How can we make teachers “align” their teaching to the state standards? A remarkable number of people, including some critics of high-stakes testing, have casually accepted this sort of talk despite the fact that it is an appeal to naked power. “Alignment” isn’t about improvement; it’s about conformity.

Standards-as-mandates also imply a rather insulting view of educators—namely, that they need to be told what (and, by extension, how) to teach by someone in authority because otherwise they wouldn’t know. While plenty of teachers need help, virtually everyone is likely to resist having the state try to micromanage his or her classroom. Some will do their best to ignore the standards, while others will comply resentfully. Either way, the use of control leads to poor implementation of the standards (which, come to think of it, may not be such a bad thing). Others, including some of our best educators, will throw up their hands in disgust and find another career.

*Based on these four criteria, the standards promulgated by disciplinary groups (the councils of teachers of mathematics and English, for example) come out considerably better than the standards issued by states. This doesn’t preclude our objecting to certain aspects of the former documents, of course, nor does it mean that all state standards are equally bad, as a side-by-side comparison of, say, Minnesota’s Profile of Learning and Virginia’s Standards of Learning will confirm. [Addendum: Two years later, the choice of these particular states to illustrate a contrast proved painfully ironic when the newly elected Republican governor of Minnesota appointed a new commissioner of education: the Christian conservative who had held that same position in . . . Virginia.]

Currently, however, there is considerable pressure to implement the kind of standards that I am suggesting are the worst. Chester E. Finn Jr. and his colleagues want states to spell out “which books children should read in English class, which individuals and events to study in history, and so on”; any other standards are simply “fluff.” Pro-standards groups such as Achieve Inc. (a group of corporate officials and politicians) tend to give poor ratings to states whose standards aren’t sufficiently specific, measurable, uniform, or compulsory.

The difference between these evaluations and a report by, say, the National Rifle Association that assigns low grades to legislators who are not sufficiently pro-gun is that in the latter case everyone realizes the ratings reflect a specific and very debatable point of view. By contrast, those who mark down a state for granting too much autonomy to local school boards, or for having standards that wouldn’t satisfy behaviorists, would like us to accept this as an objective evaluation. (One could make a case that states given an A in annual evaluations really deserve an F, and vice versa—or that a state should commission a review of its standards and testing policy by one of these groups and then do precisely the opposite of what is recommended.)

An important side note here: There has been some grumbling lately about the use of off-the-shelf tests that are unrelated to the state’s standards documents—for example, in California. From a psychometric perspective, this practice doesn’t make much sense. From a pedagogical perspective, though, the only thing worse than tests that aren’t aligned to the standards are tests that are aligned to the standards. The former is silly because it is inefficient, while the latter is dangerous precisely because it is efficient . . . at accomplishing a dubious goal. Not only politicians, but also some assessment experts sometimes forget that doing something well is not the same as doing something that’s worthwhile. When the standards and tests fit together perfectly to create an airtight system of top-down, uniform, “bunch o’ facts” schooling — well, that’s when we’re really in trouble.

The tests arguably constitute the most serious and immediate threat to good teaching, such that freeing educators and students from their yoke should be our top priority. But we should not limit our critique to the testing, which is, after all, one manifestation of a larger, and seriously wrong-headed, approach to pedagogy and school reform.

I am not troubled by those who disagree with my criteria or who like a given set of standards more than I do. In fact, I welcome such challenges. What troubles me is the rarity of such discussion, the absence of questioning, the tendency to offer instruction about how to teach to the standards before we have even asked whether doing so is a sound idea.

Copyright © 2001 by Alfie Kohn. This article may be downloaded, reproduced, and distributed without permission as long as each copy includes this notice along with citation information (i.e., name of the periodical in which it originally appeared, date of publication, and author's name). Permission must be obtained in order to reprint this article in a published work or in order to offer it for sale in any form. Please write to the address indicated on the [Contact Us](#) page.
www.alfiekohn.org — © Alfie Kohn