

Schooling Beyond Measure

By **Alfie Kohn**

[This is a slightly expanded version of the published article.]

As we tend to value the results of education for their measurableness, so we tend to undervalue and at last ignore those results which are too intrinsically valuable to be measured.

— Edmond G. A. Holmes,
chief inspector of elementary schools
for Great Britain, 1911

The reason that [standardized test](#) results tend to be so uninformative and misleading is closely related to the reason that these tests are so popular in the first place. That, in turn, is connected to our attraction to — and the trouble with — [grades](#), [rubrics](#), and various practices commended to us as [“data-based.”](#)

The common denominator? Our culture’s worshipful regard for numbers. Roger Jones, a physicist, called it “the heart of our modern idolatry . . . the belief that the quantitative description of things is paramount and even complete in itself.”

Quantification can be entertaining, of course: Readers love top-ten lists, and our favorite parts of the news are those with numerical components — sports, business, and weather. There’s something comforting about the simplicity of specificity. As the educator Selma Wassermann observed, “Numbers help to relieve the frustrations of the unknown, for nothing feels more certain or gives greater security than a number.” If the numbers are getting larger over time, we figure we must be making progress. Anything that resists being reduced to numerical terms, by contrast, seems vaguely suspicious, or at least suspiciously vague.

In his book *Trust in Numbers*, historian Theodore Porter points out that quantification has long exerted a particular attraction for Americans. “The systematic use of IQ tests to classify students, opinion polls to quantify the public mood...[and] even cost-benefit analyses to assess public works — all in the name of impersonal objectivity — are distinctive products of... American culture.”

In calling this sensibility into question, I’m not denying that there’s a place for quantification. Rather, I’m pointing out that it doesn’t always seem to know its place. If the question is “How tall is he?”, “six-foot-two” is a more useful answer than “pretty damn tall.” But what if the question were “Is that a good city to live in?” or “How does she feel about her sister?” or “Would you rather have your child in this teacher’s classroom or that one’s?”

The habit of looking for numerical answers to just about any question can probably be traced back to

overlapping academic traditions like behaviorism and scientism (the belief that all true knowledge is scientific), as well as the arrogance of economists or statisticians who think their methods can be applied to everything in life. The resulting overreliance on numbers is, ironically, based more on faith than on reason. And the results can be disturbing.

In education, the question “How do we assess (kids, teachers, schools)?” has morphed over the years into “How do we measure...?” We’ve forgotten that assessment doesn’t require measurement — and, moreover, that the most valuable forms of assessment are often qualitative (say, a narrative account of a child’s progress by an observant teacher who knows the child well) rather than quantitative (a standardized test score). Yet the former may well be brushed aside in favor of the latter — by people who don’t even bother to ask what was on the test. It’s a number, so we sit up and pay attention. Over time, the more data we accumulate, the less we really know.

You’ve heard it said that tests and other measures are, like technology, merely neutral tools, and all that matters is what we do with the information? Baloney. The measure affects that which is measured. Indeed, the fact that we chose to measure in the first place carries causal weight. His speechwriters had President George W. Bush proclaim, “Measurement is the cornerstone of learning.” What they should have written was, “Measurement is the cornerstone of the kind of learning that lends itself to being measured.”

One example: It’s easier to score a student writer’s proficiency with sentence structure than her proficiency at evoking excitement in a reader. Thus, the introduction of a scoring device like a rubric will likely lead to more emphasis on teaching mechanics. Either that, or the notion of “evocative” writing will be flattened into something that can be expressed as a numerical rating. Objectivity has a way of objectifying. Pretty soon the question of what our whole education system ought to be doing gives way to the question of which educational goals are easiest to measure. That means, in the words of University of Colorado professor Kenneth Howe, putting “the quest for accurate measurement - and control - above the quest for educationally and morally defensible policies.”

A few years ago, a writer in Education Week recalled a conversation with the director of testing for a state’s education system who “agreed that being able to make a public presentation was likely to be a more important skill for adults than knowing how to factor a polynomial. ‘But,’ he added, ‘I know how to test the ability to factor a polynomial.’” Only the latter, therefore, was going to be assessed — and therefore taught.

I’ll say it again: Quantification does have a role to play. We need to be able to count how many kids are in each class if we want to know the effects of class size. But the effects of class size on what? Will we look only at test scores, ignoring outcomes such as students’ enthusiasm about learning or their experience of the classroom as a caring community?

Too much is lost to us — or warped — as a result of our love affair with numbers. And there are other casualties as well:

1. We miss the forest while counting the trees: Rigorous ratings of how well something is being done tend to distract us from asking whether that activity is sensible or ethical. Dubious cultural values and belief systems are often camouflaged by numerical precision, sometimes out to several decimal places. Stephen Jay Gould, in his book *The Mismeasure of Man*, provided ample evidence that meretricious findings are often produced by impressively meticulous quantifiers.

2. We become obsessed with winning: An infatuation with numbers not only emerges from but also exacerbates our cultural addiction to [competition](#). It’s easier to know how many others we’ve

beaten, and by how much, if achievements have been quantified. But once they're quantified, it's tempting for us to spend our time comparing and ranking — trying to triumph over one another rather than cooperating.

3. We deny our subjectivity. Sometimes the exclusion of what's hard to quantify is rationalized on the grounds that it's "merely subjective." But subjectivity isn't purged by relying on numbers; it's just driven underground, yielding the appearance of objectivity. An "86" at the top of a paper is steeped in the teacher's subjective criteria just as much as his comments about that paper. Even a score on a math quiz isn't "objective": It reflects the teacher's choices about how many and what type of questions to include, how difficult they should be, how much each answer will count, and so on. Ditto for standardized tests — except the people making those choices are distant and invisible.

Subjectivity isn't a bad thing; it's about judgment, which is a marvelous human capacity that, in the plural, supplies the lifeblood of a democratic society. What's bad is the use of numbers to pretend that we've eliminated it.

Skepticism about — and denial of — judgment in general is compounded these days by an institutionalized distrust of teachers' judgments. Hence the tidal wave of standardized testing in the name of "accountability." Part of the point is to bypass the teachers, and indeed to evaluate them, too. The exalted status of numerical data also helps to explain why teachers are increasingly being trained rather than educated.

*Interestingly, some thinkers in the business world understand all of this. The late W. Edwards Deming, guru of Quality management, once declared, "The most important things we need to manage can't be measured." If that's true of what we need to manage, it should be even more obvious that it's true of what we need to teach.

It should be, but it isn't. As a result, we're left vulnerable to the misuse of numbers, a timely example being the pseudoscience of "value-added modeling" of test data — debunked by experts but continuing to sucker the credulous. The trouble, however, isn't limited to lying with statistics. Quantification can be a problem even when it's done honestly and competently. Better tests — or tests that are formative rather than summative — won't solve the problem. Neither will rating based on more ambitious or humanistic criteria.

At the surface, yes, we're obliged to do something about bad tests and poorly designed rubrics and meaningless data. But what lies underneath is an irrational attachment to tests, rubrics, and data, per se — or, more precisely, our penchant for reducing to numbers what is distorted by that very act.

Copyright © 2012 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. We can be reached through the [Contact Us](#) page. www.alfiekohn.org — © Alfie Kohn