

What Passes for School Reform: “Value-Added” Teacher Evaluation and Other School Reform Absurdities (##)

September 9, 2010

“Value-Added” Teacher Evaluation and Other School Reform Absurdities

By Alfie Kohn

The less people know about teaching and learning, the more sympathetic they're likely to be to the kind of “school reform” that's all the rage these days. Look, they say, some teachers (and schools) are lousy, aren't they? And we want kids to receive a better education – including poor kids, who typically get the short end of the stick, right? So let's rock the boat a little! Clean out the dead wood, close down the places that don't work, slap public ratings on these suckers just like restaurants that have to display the results of their health inspections.

On my sunnier days, I manage to look past the ugliness of the L.A. Times's unconscionable public shaming of teachers who haven't “added value” to their students, the sheer stupidity and arrogance of Newsweek's cover story on the topic last spring, the fact that the editorials and columns

about education in every major newspaper in the U.S. seem to have been written by the same person, all reflecting an uncritical acceptance of the Bush-Obama-Gates version of school reform.

I try to put it all down to mere ignorance and tamp down darker suspicions about what's going on. If I squeeze my eyes tightly, I can almost see how a reasonable person, someone who doesn't want to widen the real gap between the haves and have-nots (which is what tends to happen when attention is focused on the gap in test scores), might look at what's going on and think that it sounds like common sense.

Unfortunately, the people who know the most about the subject tend to work in the field of education, which means their protests can be dismissed. Educational theorists and researchers are just "educationists" with axes to grind, hopelessly out of touch with real classrooms. And the people who spend their days in real classrooms, teaching our children – well, they're just afraid of being held accountable, aren't they? (Actually, proponents of corporate-style school reform find it tricky to attack teachers, *per se*, so they train their fire instead on the unions that represent them.) Once the people who do the educating have been excluded from a conversation about how to fix education, we end up hearing mostly from politicians, corporate executives, and journalists.

This type of reform consists of several interlocking parts, powered by a determination to "test kids until they beg for mercy," as the late Ted Sizer once put it. Test scores are accepted on faith as a proxy for quality, which means we can evaluate teachers on the basis of how much value they've added – "value" meaning nothing more than higher scores. That, in turn, paves the way for manipulation by rewards and punishments: Dangle more money in front of the good teachers (with some kind of pay-for-performance scheme) and shame or fire the bad ones. Kids, too, can be paid for jumping through

hoops. (It's not a coincidence that this incentive-driven model is favored by economists, who have a growing influence on educational matters and who still tend to accept a behaviorist paradigm that most of psychology left behind ages ago.)

"Reform" also means diverting scarce public funds to charter schools, many of them run by for-profit corporations. It means standardizing what's taught (and ultimately tested) from coast to coast, as if uniformity was synonymous with quality. It means reducing job security for teachers, even though tenure just provides due-process protections so people can't be sacked arbitrarily. It means attacking unions at every opportunity, thereby winning plaudits from the folks who, no matter what the question, mutter menacingly about how the damned unions are to blame.

And of course it means describing as "a courageous challenge to the failed status quo" what is really just an intensification of the same tactics that have been squeezing the life out of our classrooms for a good quarter-century now. That intensification has been a project of the Obama administration, even though, as Rep. John Kline (R-MN) remarked the other day, in its particulars it comes "straight from the traditional Republican playbook."

We can show that merit pay is counterproductive, that closing down struggling schools (or firing principals) makes no sense, that charters have a spotty record overall (and one much-cited study to the contrary is deeply flawed), that high-stakes testing has never been shown to produce any benefit other than higher scores on other standardized tests (and even that only sporadically). To make these points is not to deny that there are some lousy teachers out there. Of course there are. But there are far more good teachers who are being turned into bad teachers as a direct result of these policies.

*

How do such strategies get to be called “school reform” – as opposed to “one particular, highly debatable version of school reform”? Partly, as I say, because those in the best position to challenge them have been preemptively silenced, but also because the so-called reformers are expert at framing the issue. They know that if the focal question is “Don’t you agree that a lot of schools stink?” or “Shouldn’t we hold teachers and schools accountable?” then they have the advantage. They can present their slash-and-burn tactics as “better than nothing” (as if nothing were the only alternative) or as “tough medicine” (even though what they’re peddling is worse than the disease it’s supposed to cure).

What if we asked other questions instead? We could do so about any of the policies I’ve mentioned, but for now let’s consider the idea of judging teachers with a “value-added” method.

Question 1: Does this model provide valid and reliable information about teachers (and schools)? *Most experts in the field of educational assessment say, Good heavens, no. This year’s sterling teacher may well look like crud next year, and vice versa. Too many variables affect a cohort’s test scores; statistically speaking, we just can’t credit or blame any individual teacher.*

Unfortunately, many of the experts who point this out tend to stop there, even though the problem runs far deeper than technical psychometric flaws with the technique. For example.

. . .

Question 2: Does learning really lend itself to any kind of “value-added” approach? *It does only if it’s conceived as an assembly line process in which children are filled up with facts and skills at each station along a conveyor belt, and*

we need only insert a dipstick before and after they arrive at a given station (say, fourth grade), measure the pre/post difference, and judge the worker at that station accordingly. The very idea of “value-added measures,” not just a specific formula for calculating them, implicitly accepts this absurd model.

Question 3: Do standardized tests assess what matters most about teaching and learning? If not, then no value-added approach based on those tests makes any sense. As I’ve argued elsewhere – and of course I’m hardly alone in doing so – test results primarily tell us two things: the socioeconomic status of the students being tested and the amount of time devoted to preparing students for a particular test.

Regarding individual students, at least three studies have found a statistically significant positive relationship between high scores on standardized tests and a relatively shallow approach to learning. Regarding individual teachers, let’s just say that some of the best the field has to offer do not necessarily raise their kids’ test scores (because they’re too busy helping the kids to become enthusiastic and proficient thinkers, which is not what the tests measure), while some teachers who are very successful at raising test scores are not much good at anything else. Finally, regarding whole schools, if test scores rise enough, and for long enough, to suggest a trend rather than a fluke, the rational response from a local parent would be, “Uh-oh. What was sacrificed from our children’s education in order to make that happen?”

It won’t do to fall back on the tired slogan that test scores may not be perfect, but they’re good enough. The more you examine the construction of these exams, the more likely you are to conclude that they do not add any useful information to what can be learned from other, more authentic forms of assessment. In fact, they actively detract from our understanding about learning (and teaching) because their

results are so misleading.

Notice, by the way, that everyone who declares that we ought to reward good teachers and boot the bad ones is assuming that all of us agree on what “good” and “bad” mean. But do we? I’d argue that a dipstick, test-based model is endorsed by newspapers, by public officials, and by billionaires who have bought their seat at the policy-making table (seat, hell; they own the table itself) precisely because we often don’t agree.

Imagine a teacher who gives students plenty of worksheets to complete in class as well as a substantial amount of homework, who emphasizes the connection between studying hard and getting good grades, who is clearly in control of the class, insisting that students raise their hands and wait patiently to be recognized, who prepares detailed lesson plans well ahead of time, uses the latest textbooks, gives regular quizzes to make sure kids stay on track, and imposes consequences to enforce rules that have been laid out clearly from the beginning. Plenty of parents would move mountains to get their children into that teacher’s classroom. I’d do whatever I could to get my children out.

Of course people disagree about good education, just as they may not see eye to eye about which movies or restaurants are good. We may never change each other’s minds, but we ought to have the chance to try, to discuss our criteria and reflect on how we arrived at them. As Deborah Meier likes to point out, disagreement is both valuable and inevitable in a democratic society. Undemocratic societies attempt to conceal the disagreement, imposing a single, simple standard from above – and, worse, use that standard to make decisions that can ruin people’s lives: which teachers will be humiliated or even fired, which kids will be denied a diploma or forced to repeat a grade, which schools will be shut down. A productive discussion about who’s a good teacher (and why) is less likely to take place when the people with the power

get to enforce what becomes the definition of quality by default: high scores on bad tests.

I don't expect the founder of a computer empire like Bill Gates, or a lawyer like Joel Klein, or a newspaper editor to understand the art of helping children to understand ideas, or of constructing tasks to assess that process. I just expect them to have the humility, the simple decency, not to impose their ignorance on the rest of us with the force of law.

To fight back, an awful lot of teachers who have been celebrated for their students' high scores – those teachers who can't be accused of sour grapes – will have to stand up and say, "Thanks, but let's be honest. All of us who work in schools know that you can't tell how good a teacher is on the basis of his or her kids' test results. In fact, by being forced to think about those results, my colleagues and I are held back from being as good as we can be. By singling me out for commendation – and holding other teachers up to ridicule – you've lowered the quality of schooling for all kids."

Back to Blogs