

EDUCATION WEEK

Published Online: June 4, 2010

Published in Print: June 9, 2010, as **Rethinking Teacher Accountability—Before It's Too Late**

COMMENTARY

Needed: Fresh Thinking on Teacher Accountability

By James W. Stigler

[← Back to Story](#)

U.S. Secretary of Education Arne Duncan and Microsoft co-founder and philanthropist Bill Gates have both thrown their support behind a new accountability system for teachers. Based on research showing significant variability in teachers' effectiveness (as measured by their students' learning), Duncan and Gates propose developing measures of effectiveness to get rid of bad teachers and increase the pay of good ones. It sounds like common sense. Or does it?

This approach was called the "inspection" method by W. Edwards Deming, known as the father of the science of quality improvement. Inspection, he wrote, is not an effective way to improve quality because it has no effect on the process that caused suboptimal results in the first place. Real and continuous improvement, Deming argued, occurs only when the workers themselves study outcome variability and the processes that produce it.

Deming's alternative to useless "inspection" is simple: Start with a well-defined goal and agreed-on measures for charting progress toward the goal. Then involve workers in studying and improving the process that leads to the desired outcome, using "PDSA" cycles: First, *plan* an innovation, something worth trying; next, *do* it; then *study* the result of the change; and finally, *act*, whether by trying something else if the innovation didn't work, modifying the innovation and going through the cycle again, or implementing the innovation as a permanent change in the production process. Over time, PDSA cycles yield permanent improvements in the production process, something that never results from merely inspecting the product and tossing out the defective ones.



—Chris Whetzel

Deming's science of quality improvement has not been widely applied to education in the United States. But it has in Japan, as exemplified by the "lesson study" movement. In a process similar to Deming's PDSA cycles, Japanese teacher groups develop, implement, and test improvements in teaching methods that can be shared with the profession as a whole. Some methods are better than others. But because teaching is complex, often certain methods are better only in certain situations.

Thus, it is not enough just to improve the methods teachers use. We also need to improve teachers' capabilities for selecting the best method to use in each particular situation, and of implementing that method in an expert way. The beauty of lesson study is that it accomplishes both goals: It is a process for making incremental improvements to teaching methods over time, but it also gives teachers the opportunity to learn and grow their expertise.

Most people who have tried lesson study in the United States have seen it as simply an appealing structural alternative to conventional forms of professional development. But in Japan, lesson study is part of a wider, albeit informal, accountability system. As described to me years ago by research colleagues there, teachers teaching the same content (for instance, all the 4th grade teachers in a school) meet together to develop monthly exams based on their own curricular goals. (Japan has a national curriculum, so all teachers teach the same topics at roughly the same time.) All teachers administer the common exam, and then meet as a group to score the exams and examine the performance of students in each classroom.

These are not standardized tests, and they are not national. Because the teachers themselves construct common exams, they see them as fair and valid assessments of learning outcomes. The comparisons across classrooms are also perceived as fair. Classes in Japan are larger than in the United States, and students are assigned randomly, so there is no reason to expect differences in learning across classrooms unless they are the result of something the teachers are doing differently from one another. As teachers jointly examine their classroom averages, they look at the variability across classes and discuss what might explain greater-than-expected differences. They might wonder, for example, why so many students in one teacher's class failed to find common denominators before adding fractions, whereas most students in the other classes did not make that error.

Questions like these lead to open discussions that may at times be difficult. Sometimes a particular teacher's classroom will consistently produce the lowest average score month after month, and this can be the focus of a lesson-study-type investigation. A low-performing teacher may ask others to observe his teaching and give him feedback. And other teachers in the group are willing to help. Through this process, a lower-performing teacher has an opportunity to learn and improve. But what if he doesn't want to improve, or isn't willing to put the requisite work into improving his performance? As one teacher told me, such teachers usually leave the profession. Think how it must feel to sit with your colleagues and produce subpar results month after month. This would indicate, perhaps, that you should try a different line of work.

What a simple accountability system this is. It is not based on standardized tests, and does not require the statistical adjustments used by value-added assessments most commonly proposed in this country. The results are private; they aren't published in the newspaper, or even necessarily relayed to the principal. Like Deming's PDSA cycles, the Japanese system provides interpretable data directly to those in a position to act on it: the teachers

"Even if we achieve a perfect test of teacher effectiveness, will it get us where we need to go?"

themselves. The test results do not affect teachers' pay. But then, does anyone really think American teachers would "teach better" for higher pay? Most teachers teach because they want to help students learn. If they aren't succeeding, it's because they don't know how to do any better, not because they are holding out for higher pay.

In this elegant Japanese accountability system, teachers are given realistic feedback on their performance. But, equally important, they are given a regular context in which, assisted by their colleagues, they can use outcome data to drive improvements in practice.

Implementing an accountability system like this in the United States would require a change in the culture of teaching. Teachers would need to willingly make their work public, at least to their colleagues, and be willing to engage in frank discussion and analysis with colleagues. They would need to change their view of teaching from being an idiosyncratic art to a practice that can be studied and improved over time. They would need to take pride in the quality of education at the school level, and not just at the individual teacher level.

The accountability system favored by Duncan and Gates would also change the culture of teaching: It might become less collaborative (if that is possible), as teachers sought to outperform their colleagues for higher pay. Lost in their approach would be the individual satisfactions that can come from working as a team to achieve tangible student benefits. Under the Duncan-Gates approach, many good teachers likely would leave the profession, not wishing to be held accountable for variations in outcomes not under their control.

Before beginning this difficult journey toward an untested accountability system, perhaps we should rethink the direction we are going. Even if we achieve a perfect test of teacher effectiveness (and if we don't, it won't be for lack of money), will it get us where we need to go? Shouldn't we at least look seriously at some alternative accountability systems that provide mechanisms for improvement, not just a means of inspection and sorting?

James W. Stigler is a professor of psychology at the University of California, Los Angeles, and a senior partner at the Carnegie Foundation for the Advancement of Teaching.

Vol. 29, Issue 33, Pages 30,36

RELATED STORIES

"Commentary: Beyond the Rhetoric of National Standards," May 19, 2010.

"Commentary: Attention, Gates: Here's What Makes a Great Teacher," February 3, 2010.

"In 'Lesson Study' Sessions, Teachers Polish Their Craft," February 11, 2004.

