COMMENTARY

Taking Teacher Evaluation to Extremes

By Kenneth Mitche

New York state's recent success in winning approximately $700 million in federal **Race to the Top** money might be attributed to its concept for a new teacher-evaluation system that holds teachers accountable for the performance of their students on state and standardized measures in math and English in grades 4-8.

In that vein, New York's school district leaders were strongly encouraged to sign a "memorandum of understanding" last summer to ensure that they "understood" the general concept behind the proposal, even though the details of the plan remained vague. Superintendents and boards of education were assured that all would be fine in the end as New York would receive its money—and then worry about the details. The state requested time, patience, and a little faith from those running the school districts.

In an article on the New York State School Boards Association's website, state Commissioner of Education David Steiner **challenged skeptics** of the not-yet-fully-defined plan: "There is a strong 'counter' movement, supported by many teachers, the schools of education that support them, and a sizable group of parents. They are skeptical that multiple-choice tests can capture the rich skills and knowledge that children should encounter, and doubtful that mathematical equations based on tests can ever be an adequate way to measure teacher performance. They are convinced that the most-effective learning often occurs in project- and team-based environments and that we need to get away from test prep and instead focus on critical thinking and metacognitive skills." He encouraged resistance to the temptation to "reify such distortions." Commissioner Steiner and other New York state officials simply want the time and support to develop the details.

Why are we taking so long to get started when the solution is so simple? Teachers teach, and students are tested; hence, teachers can be evaluated on their students' success. It is obvious, right? To assist the commissioner in developing a set of responses to the impatient or the audacious who question that which has no been answered, here is a modest solution: Let's rate all the teachers on how well their students perform. Why wait for details?

New York's plan already includes the elementary classroom teachers, so why not start assessing other disciplines? Physical education teachers can be rated according to how well their students increase their speed in the 100-yard dash, run the mile, or make consecutive foul shots. We can skip lessons on concepts that can't be measured, such as perseverance, courage, or risk-taking. Keep it simple and efficient. We'll weigh students when they begin the year and rate the teachers on the weight students lose or the muscle mass they gain. (No excuse will be accepted about the meals that are served at home or exercise habits instilled by the family. With time, we will surely have a value-added formula to address non-school-related factors.)

Just as the **No Child Left Behind Act** targeted 2013-14 as the year when all students will achieve proficiency in reading and math, we must declare that obesity and sloth will be eradicated by a specific date that's not too far off in the future, but not too close, allowing us to "develop the details" while we figure this out. Since many proponents of the Race to the Top disagree that nonschool factors such as poverty, emotional obstacles, student motivation, or neurological damage from low birth weight or poor nutrition are excuses for a teacher's failure, all children will become fit and fast, or

we'll find teachers who will get them there. It is a very simple and obvious solution, once we develop the details.

We can apply this to the rating of teachers in kindergarten, where all children will learn to develop socially and emotionally and be ready to learn and with no excuses about their readiness or natural rate of cognitive development. Our music teachers will be rated on the abilities of students to play state-approved scales with precision, all of which will be determined by an objective evaluation process. If there is extra time in the school year, and students have demonstrated proficiency on their scales, then perhaps there will be time for interpretive or improvisational enrichment. But first, we need the basics.

Of course, all disciplines at all levels will have similar forms of precise measurement. We can become more efficient in the study of poetry by defining metaphors without wasteful discussion—Robert Frost's "wall" is simply a wall. Just imagine how much poetry could be taught if we skipped circular and inefficient analysis. We can eliminate vitriolic political discourse in history classes by testing student knowledge of presidential facts and the "real" causes of the Civil War. Learning through discourse and debate is a messy process and has the potential to create conflict. We have a chance to clean it up and simplify.

To ensure that we have the best teachers, we cannot stop with the public schools. With the plan on the books and eventual development of specific details for important assessments that will, of course, be scientifically based, we will not only be able to hold public school teachers accountable, pay them what they deserve, and get rid of the ones whose students cannot pass these tests, but also to measure the teachers who taught the teachers—college professors.

Perhaps there would be better teaching if the undergraduate professors had done a better job preparing public school teachers. Then there are the schools of education and their teachers. Let's rate the education professors on how well the students of those they taught have performed. We can even rate the universities. It's as logical as the plan to rate public school teachers. Stacked like babushka—Russian nesting dolls, we will have a system that rates to the core of the teaching deficiencies that have brought our nation to the brink of ruin.

Skeptics might argue that there are too many confounding variables. Professors may blame the admissions office for allowing such students into their programs or may criticize the high schools for not preparing students for the rigors of college. Well, that's not right. Public schools may not use excuses, such as home environment, so the college teachers will no longer be able to scapegoat high schools. Colleges will use their version of a value-added assessment model. It can separate the irrelevant variables and is a simple solution.

All the external factors that seem to affect performance—at any level and in any discipline—can be disaggregated. Psychometricians in New York or Washington will surely develop a set of algorithms that can circumvent any argument that advances the notion that non-school-based factors affect learning. (If the quants can use their brilliant mathematical minds to run Wall Street, why not schools?) The District of Columbia's former education czar, Michelle Rhee, argued that the concern that poverty affects learning was "crap." And lots of smart and influential people agree with her and seem to have found a data-driven solution, but, of course, with time, patience, and a long-term target date to allow them to work out the details.

Kenneth Mitchell is the superintendent of schools in the South Orangetown Central School District in New York.

RELATED STORIES

"Commentary: Value-Added: It's Not Perfect, But It Makes Sense," December 15, 2010.

"Commentary: Public Displays of Teacher Effectiveness," December 15, 2010.

"A Look at the Fuss Over Value-Added Teacher Data," October 22, 2010.

"The Problems With Value-Added Assessment," (Bridging Differences Blog) October 5, 2010.

"Commentary: Value Added?," September 14, 2010.