Wanted: Ways to Assess the Majority of Teachers

By Stephen Sawchuk

The debate about “value added” measures of teaching may be the most divisive topic in teacher-quality policy today. It has generated sharp-tongued exchanges in public forums, in news stories, and on editorial pages. And it has produced enough policy briefs to fell whole forests.

But for most of the nation’s teachers, who do not teach subjects or grades in which value-added data are available, that debate is also largely irrelevant.

Now, teachers’ unions, content-area experts, and administrators in many states and communities are hard at work examining measures that could be used to weigh teachers’ contributions to learning in subjects ranging from career and technical education to art, music, and history—the subjects, in other words, that are far less frequently tested.

The work, which has taken place quietly, in contrast to the larger value-added conversation, is renewing interest in alternative sources of achievement information.

“A standardized test is very useful to take a snapshot of the inventory of knowledge and skills that a student has, but it’s not as useful to change teacher practice and improve strategies and determine what needs to happen next,” said Laura Goe, a principal investigator for the federally financed National Comprehensive Center for Teacher Quality, who has worked with several states on teacher evaluations.

Such work is also giving way to questions about how to make sure that alternative measures—such as projects, portfolios scored according to guidelines, and classroom-based assessments—are both rich sources of data and provide comparable information about teachers and their impact on student learning.

“I would say this has been a very challenging area because it’s so high-risk in a sense,” said Lawrence T. Waite, the program manager of a teacher-evaluation initiative housed at the New York State United Teachers, or NYSUT. “Using growth is new. Certainly, teachers look at student learning every day, but we’re now doing this to look at teachers’ impact on student learning, which raises the stakes on the use of that assessment instrument.”

Gauging Evaluations
Charged with using multiple measures for evaluating teachers, states and district are mulling over methods for estimating the impact of teachers on student learning.

**STUDENT-LEARNING OBJECTIVES**

Individual teachers, or sometimes teams, set goals for student growth, subject to certain parameters, with their principals. The goals are approved by principals and audited using agreed-upon assessments.

**Used for bonus-pay systems in:**

- Charlotte-Mecklenburg, N.C.
- Austin, Texas
- Denver

**Used for professional development by:**

- Teach For America (goals set with program directors)

**Used for teacher evaluation in:**

- Rhode Island

**COMMON GROWTH MEASURES**

Committees of teachers review a variety of measures of student growth and make recommendations to either the district or state, which approves a list for each subject and grade. Measures have comparability across classrooms and/or districts.

- Albany City, Hempstead, Marlboro, North Syracuse, and Plattsburgh school districts, N.Y.
- Delaware

**STANDARDS-BASED TESTS**

Pre- and post-course tests aligned to standards are developed in most subjects and grades. Results are comparable across the district.

- Hillsborough County, Fla.
- District of Columbia (planned)

**SCHOOLWIDE STUDENT GROWTH**

Under this system, teachers in nontested grades and subjects have their student-growth targets based on schoolwide growth on state standardized tests or on another schoolwide index.
If the debate seems wonkish and academic, the consequences are anything but. In a wave of legislative action last year, nearly a dozen states took steps to require teacher-evaluation systems to consider evidence of students’ academic growth.

Value-added measures rely on state standardized tests to generate the individual teacher estimates and are typically available only in reading and mathematics in grades 4-8.

One widely cited statistic puts the proportion of those who teach in nontested grades and subjects at about 70 percent, but technical issues can push the figure much higher. Under the District of Columbia’s IMPACT teacher-evaluation system, just 15 percent of teachers have individual value-added data, according to school district officials.

And because legislation in several states, including New York, Rhode Island, and Tennessee, requires evaluations to weigh multiple measures of a teacher’s contribution to student learning, all teachers in those states will need more than just value-added statistics to show that they are effective.

**Bringing Teams Together**

In New York and Rhode Island, the American Federation of Teachers has played a role in helping shape all aspects of 10 districts’ teacher-evaluation frameworks.

Through a $5 million federal Investing in Innovation grant, as well as a cut of the AFT’s own Innovation Fund program, the union has brought together teams of education leaders from participating districts. Each team consisted of the superintendent, teachers’ union president, and other representatives, including classroom teachers.

“Our group was excited and exhilarated and scared to death, all at once,” said Cathy Corbo, the president of the local teachers’ union in Albany, one of the districts participating in the work. “The process has been pretty intense.”

NYSUT, the state affiliate overseeing the grant, brought in 20 additional content-area teachers to help guide the work of developing measures for non-tested grades and subjects. The teachers discussed textbook units, teacher-made assessments, district exams, and New York state’s Regents exams, among other things.

In some subjects, tools now used for other purposes could potentially be adapted. History teachers, for instance, could gauge students’ progress at making connections between primary- and secondary-source documents, using Advancement Placement materials.
Officials have found that such examples often come with a catch. “What we found is that many of those assessments were measuring the endpoint and not the beginning point [of student learning],” said Mr. Waite of NYSUT. “Then the teachers had to grapple with, if the Regents exam is going to be the final test, what would be the pretest?”

Devising Frameworks

Certain subjects, like physical education, writing, art, and career-tech fields are more performance-based and would likely require performance assessments to become part of the mix of measures.

In the Charlotte-Mecklenburg district in North Carolina, a performance-pay program in 20 schools helps teachers, in consultation with their principals, set a rigorous achievement goal for their students and select or craft a variety of assessments to determine if they’ve met that goal, including performance measures.

At West Mecklenburg High School, for instance, Christine E. Kapakos, a teacher of culinary arts and family and consumer sciences, has developed a yardstick to gauge students’ knife skills. It measures aspects like knife safety, technique, and whether students can perform a variety of cuts from julienne to rondel with uniformity.

“They are such an important part of what we do in culinary,” Ms. Kapakos said. “I feel I’m trying to train students to get a job, so what do employers look for? They look at how well they handle knives, at their speed.”

Teachers in the district can use evidence of student progress on such measures to fulfill aspects of the state’s teacher evaluation, though there is as yet no formal link between the two.

Student learning objectives are among the more heavily studied options for measuring teacher impact on learning. Studies have shown a correlation between teachers’ crafting of rigorous learning objectives—pioneered as part of Denver’s ProComp system—and higher performance by their students.

“They’re a very precise process, intended to evoke critical and evidence-based thought about a teacher’s students and lead to objectives for student growth,” said William J. Slotnik, the executive director of the Community Training and Assistance Center, which helped in the development of ProComp and the Charlotte-Mecklenburg pay plan. “It’s trying to bring a higher level of science to a practice—goal-setting—that’s existed in districts for a long time.”

Rhode Island, which is adopting student-learning objectives as part of its teacher-evaluation system, will be among the first states to use them formally for teacher accountability.

Comparability at Issue
One of the thornier issues the states and districts are confronting has to do with vetting the measures for wide-scale use.

Rules in states like New York and guidelines for the federal Race to the Top competition, which is also helping to pay for much of the teacher-evaluation work around the nation, stress the comparability of measures across classrooms. That could be a harder goal to achieve with individually set student-learning objectives than measures that are adopted districtwide or statewide.

For New York districts involved in the AFT grant, the recommended measures will go through several more reviews by NYSUT teachers before they are made available to the districts to use in their teacher-evaluation systems, said Mr. Waite. Presumably, the agreed-upon measures will be identical for all teachers who share a course or grade.

State officials in Delaware have taken a similar approach to vetting measures. A work group of teachers there is evaluating and will make recommendations on measures to the state, which will approve a selection. Then, groups of teachers in the relevant subjects will take the lead in ensuring that the measures are implemented consistently across classrooms.

The importance of comparability from classroom to classroom is one of the reasons that leaders in the Hillsborough County, Fla., district are relying on end-of-course tests, first developed in the 1980s, to supply most of the teacher-growth information for use in the district’s teacher-evaluation system. The district recently added pre- and post-test measures to the courses to accord with a state performance-pay plan.

“It’s important to have that consistency,” said MaryEllen Elia, the superintendent of the 191,000-student district, about the end-of-course tests. “If we determine this is the level of excellence in our district, it’s important to know exactly where you are. It provides a context to see how you’re doing relative to others in similar classrooms.”

**Challenges Await**

Some areas remain a challenge to address.

Ms. Corbo of Albany said that among her union’s members, finding appropriate techniques for assessing teachers of students with disabilities remains a key concern, as does ensuring teachers aren’t penalized for the demographic makeup of their classes.

Finally, the officials say that collecting information on the validity and reliability of the measures will be especially important, given that many of them haven’t been used for teacher evaluation.

“You have to spend time looking at the properties of the measure,” said Ms. Goe of the National Comprehensive Center for Teacher Quality. “Is it something that is going to give us useful information, or is it going to be another ‘widget effect,’ where everyone comes out looking good?”
Coverage of policy efforts to improve the teaching profession is supported by a grant from the Joyce Foundation, at www.joycefdn.org/Programs/Education.

Vol. 30, Issue 19, Pages 1,15