


EDUCATION WEEK

Gates Study Offers Teacher-Effectiveness Clues

By [Stephen Sawchuk](#)

“Value added” gauges based on growth in student test scores and students’ perceptions of their teachers both hold promise as components of a system for identifying and promoting teacher effectiveness, according to preliminary findings from the first year of a major study.

The [analysis](#) , released today by the Bill & Melinda Gates Foundation, shows that teachers’ value-added histories strongly predicted how they would perform in other classrooms or school years—as did students’ perceptions of their teachers’ ability to maintain order in the classroom and provide challenging lessons.

The findings are part of the Seattle-based foundation’s \$45 million [Measure of Effective Teaching study](#). The project seeks to identify the most accurate measures of superior teaching. (“[Multi-City Study Eyes Best Gauges of Good Teaching](#),” Sept. 2, 2009.)

While underscoring the preliminary nature of the findings, Gates officials said they were heartened to see that some of the measures being studied do appear predictive of good teaching.

“I was hugely excited and encouraged” by the findings, said Vicki Phillips, the foundation’s director of education programs. “It has implications for what people can be doing right now. It begins to answer questions teachers have had. And I think it shows that valid teacher feedback doesn’t need to be limited to test scores alone.”

Among its education philanthropy, the Gates Foundation provides grant support to Editorial Projects in Education, the publisher of *Education Week*.

Complex Studies

The preliminary findings are based on data from five of the six districts participating in the study. They are New York City; Charlotte-Mecklenburg, N.C.; Hillsborough County, Fla; Dallas; and Denver.

A team of researchers directed by Thomas J. Kane, the foundation’s deputy director of research and data for its education program, analyzed student scores on state tests given in grades 4-8 in the 2009-10 school year, using value-added modeling.

Such modeling purports to control for a student's past performance and other factors so that learning gains can be attributed to specific teachers.

The researchers also analyzed student-perception data gathered from 2,519 classrooms, grades 4-8. Those data were gathered by using the Tripod survey instrument, developed by Harvard University researcher Ron Ferguson, in which students score teachers on a 1-to-5 scale on such aspects as whether teachers make the point of their lessons clear, are caring and considerate of students, and explain material in several different ways.

The analysts found that, in every grade and subject studied, teachers' value-added histories were strongly predictive of their performance in other classrooms. While they found a degree of volatility in the estimates from year to year, that volatility "is not so large as to undercut the usefulness of value-added as an indicator of future performance," the study says.

Similarly, the researchers found that student perceptions of a given teacher were generally consistent across his or her classes, and that students gave high ratings to teachers whose classes consistently made learning gains.

Certain strands of the Tripod instrument were particularly well correlated with the value-added results: Student perceptions of teachers' ability to manage a classroom and provide challenging academic content were strongly linked to those teachers' ability to raise scores.

One of the study's findings appears to challenge the conventional wisdom that teachers can boost scores by "teaching to the test."

The analysis found that the value-added estimates of teacher effectiveness held up even when students were given supplemental tests with harder tasks than those on the state tests, including conceptual questions and open-ended writing tasks. Meanwhile, student reports of classes spent heavily on test preparation were generally weak predictors of teachers' ability to raise scores.

Combining Results

The value-added findings, in particular, come in the midst of a divisive debate in the K-12 field about whether such methods should count in a teacher's evaluation.

In a series of recent reports, statisticians and test experts have lined up on both sides of that issue. Critics say the valued-added methods are opaque and too error-prone to be used in teacher evaluation,

while proponents say they can lend objectivity and clarity to such evaluations when combined with other measures.

The Gates Foundation's findings on student perceptions, in the meantime, raise new questions for states and districts about factors that could be considered in teacher evaluations. Spurred by federal grant programs, some states have moved toward including teacher observations and even value-added methods in evaluations. Far fewer, however, are considering student-perception data.

One of the benefits of the student ratings, the report says, is that they're not limited to certain grades or subjects, as testing data generally are.

So far, the study also appears to support the notion, advocated by teachers' unions and others, that evaluations should be based on multiple measures. The analysis concludes that combining both sources of information—value-added and student results—yielded a more finely grained estimate of teacher effectiveness than using the student-perception information alone.

"Teachers have been saying that they're not opposed to the performance-based aspects of evaluation if they had measures that were fair, respectful, and multiple," said the Gates Foundation's Ms. Phillips. "This shows that you can put multiple measures together in a way that honors great teaching."

Findings to Come

One key area not yet studied is the accuracy of teacher-observation ratings on a variety of different teaching frameworks. The foundation's research partners are still collecting and scoring videotaped observations of some 13,000 lessons in 2009-10 as part of that effort.

Other measures under study include teachers' pedagogical content knowledge and their perceptions of their working conditions. Those are important factors for study, the report says, because they provide crucial information to teachers about how to get better at their craft.

The most immediate beneficiaries of the information are the four districts and one charter school network participating in the Gates Foundation's \$290 million Intensive Partnerships for Effective Teaching grants. Hillsborough County; Memphis, Tenn.; Pittsburgh; and five charter networks in Los Angeles are overhauling their teacher-development and -compensation systems with the funding, and will be expected to tailor their plans based on the study's findings.

Gates officials plan to release a second report next spring as the research project moves into its second year. It will also begin to examine the results of an experiment, already under way, to gauge student

performance when students are randomly assigned to teachers identified as being more or less effective.

Final results will be released in the winter of 2011-12.

“Things we’ve intuitively known, or thought about, or wished for about teacher effectiveness—there’s now some empirical evidence that they are valid,” Ms. Phillips said.