

JOURNAL OF
POLICY ANALYSIS
AND
MANAGEMENT

Point/counterpoint: Should 'value-added' models be used to evaluate teachers?

By Allison Armour-Garb, Guest Editor
Fall 2009

Until now, accountability in education has largely been about institutions – schools, school districts, states. But reformers are beginning to talk about making it much more *personal*, measuring (or attempting to measure) the performance of individual teachers.

Spurred in part by the American Recovery and Reinvestment Act of 2009, states are developing education data systems that can match teachers to students, and that can track students' test scores from year to year. Some stakeholders – notably, teachers' unions – are concerned that these capabilities lay the groundwork for evaluating teachers based on the academic progress of their students, as measured by standardized tests.

'Value-added' models, which estimate teachers' effectiveness based on student score gains, are intuitively appealing because they attempt to get at a central question: How much are teachers contributing to their students' progress? States and districts are finding innovative ways to use the data: in Louisiana, to assess the effectiveness of teacher preparation programs; in New York City, to help teachers improve; and in Knox County, Tennessee, to study the distribution of effective teachers in high poverty schools.

But controversy surrounds the use of value-added data for decisions that carry high stakes for individual teachers, such as teacher tenure, or merit pay plans like those in Florida, Denver, and Tennessee's Hamilton County. Union opposition has led to restrictions on using value-added data in teacher pay, evaluation, or personnel decisions in California and in teacher tenure decisions in New York.

Though the idea of tracing students' performance back to the individual teacher may intuitively seem simple, value-added models are quite complicated technically in ways that can cause impatient reformers' eyes to glaze over. States and districts may therefore undertake to build value-added models and use value-added data without appreciating the statistical and testing (or "measurement") issues that are involved, as Heather C. Hill points out in the following exchange. This problem is compounded by a national shortage of expertise in educational testing and accountability systems.

How should policymakers decide whether to use value-added data to evaluate teachers? Douglas N. Harris (pro) and Hill (con) agree that policy decisions about teacher value-added should be analyzed using the criteria – *validity* and *reliability* – that measurement experts apply to determine whether a particular test is suitable for a given purpose. To those cri-

This is a preprint of an article published in the *Journal of Policy Analysis and Management*, Volume 28, Number 4. © 2009, Association for Public Policy Analysis and Management. See: <http://www3.interscience.wiley.com/cgi-bin/home>.

teria, Harris would add *cost*, a *comparison of teacher value-added with alternative methods of ensuring instructional quality*, such as performance observations and credentials, and consideration of *exactly how teacher value-added would be used*.

Harris argues that value-added models provide an inexpensive tool that could turn out to work better for some purposes than alternative methods. Hill agrees with Harris that credentials and some forms of performance observation have drawbacks, but, she contends, value-added models are not sufficiently reliable to sort good teachers from bad ones.

To access Harris' and Hill's arguments on this issue, go to the complete article at <http://www3.interscience.wiley.com/journal/34787/home>. Nonsubscribers will be given the option to purchase 24-hour access to the article for \$29.95.

Douglas N. Harris, assistant professor at the University of Wisconsin-Madison, co-chaired the National Conference on Value-Added (2008). He is a principal investigator at Teacher Quality Research and serves on the working group that advises school districts participating in the Teacher Incentive Fund, a federal pilot program that supports performance-based compensation systems. He is the author of "The policy uses and 'policy validity' of value-added and other teacher quality measures," forthcoming in D. H. Gitomer's Measurement Issues and Assessment for Teacher Quality.

Heather C. Hill, associate professor at the Harvard Graduate School of Education, is an expert in the measurement of instruction – primarily the measurement of mathematical knowledge for teaching. Her research interests include instructional improvement and the implementation and evaluation of education policy. She is the co-author, with David K. Cohen, of Learning policy: When state education reform works.