
Comparing estimates of teacher value-added based on criterion- and norm-referenced tests

David Stuit

Basis Policy Research

Mark Berends

University of Notre Dame

Megan J. Austin

University of Notre Dame

R. Dean Gerdeman

American Institutes for Research

Key findings

Three analytic strategies were used to compare estimates of teacher value-added based on a criterion-referenced state assessment and a widely used norm-referenced test. They found that:

- Single-year estimates from the state assessment and norm-referenced test were moderately correlated (correlation coefficients of 0.44 to 0.65).
- On average, 33.3 percent of estimates ranked in the same quintile on both tests in the same school year.
- No teachers had estimates above the sample average with 95 percent confidence on one test and below the sample average with 95 percent confidence on the other test.

REL 2014–004

The National Center for Education Evaluation and Regional Assistance (NCEE) conducts unbiased large-scale evaluations of education programs and practices supported by federal funds; provides research-based technical assistance to educators and policymakers; and supports the synthesis and the widespread dissemination of the results of research and evaluation throughout the United States.

January 2014

This publication was prepared for the Institute of Education Sciences (IES) under contract ED-IES-12-C-0004 by Regional Educational Laboratory Midwest, administered by American Institutes for Research. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government. The publication is in the public domain. Authorization to reproduce in whole or in part for educational purposes is granted.

This REL report is in the public domain. While permission to reprint this publication is not necessary, it should be cited as:

Stuit, D., Berends, M., Austin, M. J., & Gerdeman, R. D. (2014). *Comparing estimates of teacher value-added based on criterion- and norm-referenced tests* (REL 2014–004). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved from <http://ies.ed.gov/ncee/edlabs>.

This report is available on the Regional Educational Laboratory website at <http://ies.ed.gov/ncee/edlabs>.

Summary

Recent changes to state laws on accountability have prompted school districts to design teacher performance evaluation systems that incorporate student achievement (student growth) as a major component. As a consequence, some states and districts are considering teacher value-added models as part of teacher performance evaluations. Value-added models use statistical techniques to estimate teachers' (or schools') contributions to growth in student achievement over time.

Designers of new performance evaluation systems need to understand the factors that can affect the validity and reliability of value-added results or other measures based on student assessment data used to evaluate teacher performance. This study provides new information on the degree to which value-added estimates of teachers differ by the assessment used to measure their students' achievement growth.

To compare estimates of teacher value-added based on two different assessments, the study selected districts whose students took the criterion-referenced Indiana Statewide Testing for Educational Progress Plus (ISTEP+) and the norm-referenced Measures of Academic Progress (MAP) in the same school year. The analysis examines reading and math achievement data for grades 4 and 5 in 46 schools in 10 Indiana districts for 2005/06–2010/11.

The study used three analytic strategies to quantify the similarities and differences in estimates of teacher value-added from the ISTEP+ and MAP: correlations of value-added estimates based on the two assessments, comparisons of the quintile rankings of value-added estimates on the two assessments, and comparisons of the classifications of value-added estimates on the two assessments according to whether their 95 percent confidence intervals were above, below, or overlapping the sample average.

Consistent with prior research, the study found a moderate relationship between value-added estimates for a single year based on the ISTEP+ and MAP, with average yearly correlation coefficients of 0.44 to 0.65. The comparison of quintile rankings found that an average of 33.3 percent of estimates of teacher value-added ranked in the same quintile on both tests in the same school year. Results were more consistent for estimates in the top and bottom quintiles than in the three middle quintiles. Across all comparisons 28.1 percent of estimates ranked two or more quintiles higher on one test than on the other.

Teacher value-added estimates were more consistent between the ISTEP+ and MAP when considering the precision of the estimates, as measured by confidence intervals. None of the estimates had a 95 percent confidence interval falling above the sample average on one test and a 95 percent confidence interval falling below the sample average on the other.

Overall, the findings indicate variability between the estimates of teacher value-added from two different tests administered to the same students in the same years. Specific sources of the variability across assessments could not be isolated because of limitations in the data and research design. However, the research literature points to measurement error as an important contributor. The findings indicate that incorporating confidence intervals for value-added estimates reduces the likelihood that teachers' performance will be misclassified based on measurement error.