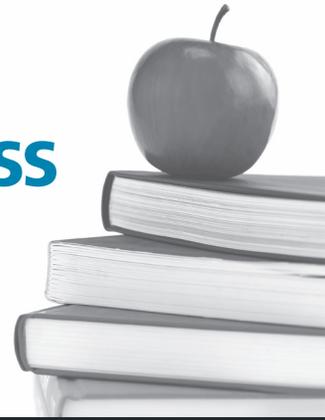


# Measuring Teaching Effectiveness

## Understanding Common, Uncommon, and Combined Methods

JONATHAN SCHWEIG



Research has validated the widespread belief that effective teaching matters. But what does effective teaching look like? And how can we measure it? Education practitioners, policymakers, and researchers have suggested a wide range of methods. Many of these have been incorporated into teacher feedback and evaluation systems.



### Three methods are used most widely in the United States.

The three most widely used measures in the United States are structured classroom observations, teacher contributions to student achievement growth, and student perceptions of teacher effectiveness and classroom instructional climate. According to the National Council on Teacher Quality, as of 2019, 44 states require classroom observations, 33 require measures of student achievement growth, and seven require student surveys as components of teacher feedback and evaluation systems. An additional 24 states permit the use of student surveys in teacher evaluation.



### There are many other possible measures that require more research attention.

Teaching effectiveness can also be inferred from tests of teachers' knowledge or skills; teachers' participation in professional development, committees, or mentoring; instructional artifacts, including lesson plans and assignments; teacher self-reporting, including instructional logs; and input provided by parents, peers, or administrators.

However, these measures have not been as thoroughly examined by researchers as have test-based and observation methods. Because of this, they are less frequently incorporated into teacher evaluation systems.



### How to combine measures may be as important as which measures to combine.

Education agencies use a wide range of models to combine multiple measures into a single teacher effectiveness appraisal. However, research on combining measures is evolving as a field, with no clear consensus about the best way to do this to achieve an accurate overall rating of teaching effectiveness. Many states and districts use weighted averages to combine information from multiple sources into a single score. Others use discrete decision rules for each measure and then specify an overall decision rule based on those separate decisions. However, combining measures does not always produce more-precise estimates of teacher quality or improve decisionmaking, and measurement error from different sources does not "cancel out" when combining measures. Researchers are exploring ways in which different approaches to combining measures affect the accuracy, consistency, and fairness of evaluation scores.

#### WANT TO LEARN MORE?

To access other fact sheets in this series, as well as reports, multimedia products, and more or visit [www.rand.org/teacher-effectiveness](http://www.rand.org/teacher-effectiveness).



## Multiple measures provide a more accurate picture of teaching effectiveness than any single measure.

Deciding between methods for measuring teacher effectiveness requires trade-offs in terms of accuracy, precision, implementation challenges, and scalability. Each provide a different perspective on teaching—for example, test scores tell us something about the outcomes attained by students in a teacher's class but do not tell us whether students enjoy being in class or whether teachers' instructional practices are aligned with district guidelines. Student surveys and classroom observations, respectively, can shed light on those topics. No single method provides a complete picture of a teacher's effectiveness; fair, accurate, and actionable appraisals of teaching quality depend on having information from multiple, complementary sources of information.

### Related Readings

Doan, Sy, Jonathan D. Schweig, and Kata Mihaly, "The Consistency of Composite Ratings of Teacher Effectiveness: Evidence From New Mexico," *American Educational Research Journal*, 2019, pp. 1–31.

Goe, Laura, Courtney Bell, and Olivia Little, "Approaches to Evaluating Teacher Effectiveness: A Research Synthesis," Washington, D.C.: National Comprehensive Center for Teacher Quality, June 2008. As of September 6, 2019:  
<https://gtlcenter.org/sites/default/files/docs/EvaluatingTeachEffectiveness.pdf>

Kane, Thomas J., Daniel F. McCaffrey, Trey Miller, and Douglas O. Staiger, *Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment*, Seattle, Wash.: Bill & Melinda Gates Foundation, Measures of Effective Teaching project, January 2013.

Kraft, Matthew A., and Gilmour, Allison F., "Revisiting the Widget Effect: Teacher Evaluation Reforms and the Distribution of Teacher Effectiveness," *Educational Researcher*, Vol. 46, No. 5, June–July 2017, pp. 234–249.

National Council on Teacher Quality, "Measures of Professional Practice: 2019 Teacher and Principal Evaluation Policy," data set, webpage, undated. As of September 6, 2019:  
<https://www.nctq.org/yearbook/national/Measures-of-Professional-Practice-95>

Steinberg, Matthew P., and Matthew A. Kraft, "The Sensitivity of Teacher Performance Ratings to the Design of Teacher Evaluation Systems," *Educational Researcher*, Vol. 46, No. 7, October 2017, pp. 378–396.

### About This Research

This document describes the three most-common methods for measuring teacher effectiveness in the United States and discusses the importance of carefully considering the ways that measures are combined to create an overall index. This research in the public interest was supported by RAND, using discretionary funds made possible by the generosity of RAND's donors and the fees earned on client-funded research.

### About RAND Education and Labor

This study was undertaken by RAND Education and Labor, a division of the RAND Corporation that conducts research on early childhood through postsecondary education programs, workforce development, and programs and policies affecting workers, entrepreneurship, and financial literacy and decisionmaking.

More information about RAND can be found at [www.rand.org](http://www.rand.org). Questions about this document should be directed to Jonathan\_Schweig@rand.org, and questions about RAND Education and Labor should be directed to [educationandlabor@rand.org](mailto:educationandlabor@rand.org).

The RAND Corporation is a research organization that develop solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is a nonprofit, nonpartisan, and committed to the public interest. RAND's publications do not necessarily reflect the opinions of its research clients or sponsors.

© 2019 RAND Corporation | Photo: TongRo Images Inc/Getty Images

**Limited Print and Electronic Distribution Rights:** This document and trademark(s) contained herein are protected by law. This representation of RAND intellectual property is provided for noncommercial use only. Unauthorized posting of this publication online is prohibited. Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please visit [www.rand.org/pubs/permissions](http://www.rand.org/pubs/permissions).



[www.rand.org](http://www.rand.org)