



NATIONAL COMPREHENSIVE CENTER  
FOR **TEACHER QUALITY**

# Measuring Teacher Effectiveness

Laura Goe, Ph.D.  
Research Scientist, ETS

March 30, 2009

Copyright © 2009 National Comprehensive Center for Teacher Quality. All rights reserved.



# How Do We Measure Teacher Effectiveness?

- "It's a hard nut to crack. The things that are easy to measure don't matter that much, and the things that matter aren't easy to measure."

Adam Gamoran, interim dean at the University of Wisconsin School of Education, talking to *Teacher Magazine* in December 2008

# Measuring Teacher Effectiveness

- Measuring teacher effectiveness is impacted by the following:
  - What is valued
  - Our technological advances and limitations
  - The data, evidence, and information we have or can acquire
  - The resources (staff, money, time, policy levers) available to us
  - The cooperation of the teachers themselves

# Teacher Qualifications and Effectiveness

- We know from research that qualifications may be useful for screening purposes but are poor predictors of teacher effectiveness.
- Thus, it is important to measure teacher effectiveness not by initial qualifications but by performance in classrooms and schools.
- However, what aspects of performance constitute teacher effectiveness are the source of much debate.

# Lack of Agreed Upon Definition

- There is little consensus about how to define teacher effectiveness.
- In many recent research studies, teacher effectiveness is not defined at all.
- The “default definition” seems to be as follows: teachers’ contribution to students’ growth as measured by standardized achievement tests.

# Five-Point Definition of Highly Effective Teachers

1. Have high expectations for all students and help students learn, as measured by value-added or alternative measures.
2. Contribute to positive academic, attitudinal, and social outcomes for students, such as regular attendance, on-time promotion to the next grade, on-time graduation, self-efficacy, and cooperative behavior.

# Five-Point Definition of Highly Effective Teachers

3. Use diverse resources to plan and structure engaging learning opportunities; monitor student progress formatively, adapting instruction as needed; and evaluate learning using multiple sources of evidence.
4. Contribute to the development of classrooms and schools that value diversity and civic-mindedness.

# Five-Point Definition of Highly Effective Teachers

5. Collaborate with other teachers, administrators, parents, and education professionals to ensure student success, particularly the success of students with special needs and those at high risk for failure.

Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved March 27, 2009, from <http://www.tqsource.org/publications/EvaluatingTeachEffectiveness.pdf>

# Most Common Measures

- **Value-Added Models:** Of little value in helping teachers improve their practice because value-added scores tell us nothing about what goes on in teachers' classrooms.
- **Teacher Observations:** Great for formative evaluation but are more expensive to conduct (personnel time, training, calibrating) and are only as good as the instruments and the observers.

# Emphasis on Student Achievement

- With the current federal emphasis on test scores (see the No Child Left Behind Act), teacher effectiveness is often discussed in terms of teachers' contributions to students' learning as measured by test scores.
- This unfortunately means that other ways teachers contribute to student learning and well-being or to the culture and stability of the school are often not measured at all and may be given too little consideration.

# What Value-Added Models Cannot Tell You

- Value-added models are really measuring classroom effects, not teacher effects.
- Value-added models can't tell you why a particular teacher's students are scoring higher than expected.
  - Maybe the teacher is focusing instruction narrowly on test content.
  - Or maybe the teacher is offering a rich, engaging curriculum that fosters deep student learning.
- How the teacher is achieving results matters.

# Other Measures

- Many types of evidence\*—including classroom observations, portfolios, administrator recommendations, analysis of teachers' assignments, analysis of students' work, documentation of teachers' positive contributions to the school, documentation of teacher leadership and mentoring—can be used in addition to student test scores.

\* For descriptions and discussions of instruments for measuring various aspects of teacher performance, see Goe, Bell, and Little (2008).

# The Quality of Teacher Observations

- There are a number of good teacher observation instruments (Charlotte Danielson's *Framework for Teaching*, CLASS, Praxis III).
- But even if you have great instruments, the usefulness of the scores for high-stakes evaluation depends on the following:
  - How well trained the observers are
  - How many times the teacher is observed
  - Whether calibration and retraining occurs

# Validity of Observations

- **Validity:** The extent to which evidence and theory support an interpretation of scores for the use of the measure (i.e., a higher standard for evidence should be used when high-stakes decisions are associated with scores).
- Do higher observation scores mean more effective teaching? Not necessarily!

# Potential Problems With Observation Instruments

## Instrumentation Is Key to Validity.

- The instrument doesn't measure what is important and valued, given the purpose of the observation.
  - What is being measured does not match with what teachers believe to be important.
  - What is being measured does not match with what observers/raters/evaluators believe to be important.

# Potential Problems With Observation Instruments

## Training Is Important to Validity.

- Raters are poorly trained in the use of the observation instrument.
  - In order for scores to be valid for any purpose, and particularly for teacher compensation, raters must all be scoring the same way.
  - In other words, a teacher should get virtually the same score, no matter which rater does the observation.
  - Proper training ensures that each rater considers the same evidence and scores it in the same way.

# Potential Problems With Observation Instruments

## Calibration Is Important to Validity.

- Raters have “drifted,” meaning the raters have forgotten how to use the instrument properly.
  - Including extraneous evidence
  - Neglecting to include important evidence
  - Allowing personal bias to creep in
- Result: The rater may be too stringent or lax in scoring.
- Rater calibration means evaluating raters’ scores and retraining when appropriate.

# Measuring Teacher Effectiveness: System Design Suggestions

- Involve teachers and stakeholders in developing the evaluation system.
  - Increases teacher/stakeholder buy-in
  - Increases validity of the system
- Use multiple indicators, not just an observation score or a value-added rank.
  - There are many other important factors you can measure economically (see the five-point definition of teacher effectiveness).

# Measuring Teacher Effectiveness: System Design Suggestions

- Differentiate among teachers by grade level and experience.
  - Standards may be the same, but what constitutes effectiveness may be different.
- Use appropriate weights to give more significance to components of the system that are most important (i.e., on-time graduation may be weighted for secondary teachers and not weighted for elementary teachers).

# Measuring Teacher Effectiveness: System Design Suggestions

- Measure what is most important to you, your administrators, your teachers, and other education stakeholders.
  - The system will drive improvement.
  - Teachers will strive to improve in areas that they know will count.
  - Ensure that what they are striving for is truly important in your definition of successful teaching.

# Measuring Teacher Effectiveness: System Design Suggestions

- Give teachers opportunities to improve in areas in which they score poorly.
  - Provide assistance in determining problem areas and planning strategies to address them.
  - Provide resources to help them improve.
    - Mentors
    - Targeted professional development
    - College courses

# Measuring Teacher Effectiveness: System Design Suggestions

## ➤ For observations:

- Set aside funds to support training and calibrating of observers.
- Devise a system that involves multiple observations throughout the year, not just one.
- Use at least two different evaluators for each teacher and average the scores.

# Summary

- There are no shortcuts to measuring teacher effectiveness.
- Design a system that drives effective instruction, not just measures it.
- The fewer indicators you use, the greater the potential for errors.
- Include other outcomes in your teacher evaluation system, not just scores from observations.



**Laura Goe, Ph.D.**

**P: 609-734-1076**

**E-Mail: [lgoe@ets.org](mailto:lgoe@ets.org)**

**National Comprehensive Center for  
Teacher Quality**

1100 17th Street NW, Suite 500

Washington, DC 20036-4632

877-322-8700 > [www.tqsource.org](http://www.tqsource.org)