



NATIONAL COMPREHENSIVE CENTER
FOR **TEACHER QUALITY**

Multiple Measures of Teacher Effectiveness

Laura Goe, Ph.D.

New York Department of Education Task Force

Webinar ♦ 12/3/2010

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



Laura Goe, Ph.D.

- Former teacher in rural & urban schools
 - Special education (7th & 8th grade, Tunica, MS)
 - Language arts (7th grade, Memphis, TN)
- Graduate of UC Berkeley's Policy, Organizations, Measurement & Evaluation doctoral program
- Principal Investigator for the National Comprehensive Center for Teacher Quality
- Research Scientist in the Learning & Teaching Research Center at ETS

The goal of teacher evaluation

*The **ultimate** goal of all teacher evaluation should be...*

**TO IMPROVE
TEACHING AND
LEARNING**

Research Behind the Push for New Evaluation Measures and Systems

- The Widget Effect report (Weisberg et al., 2009) "...examines our pervasive and longstanding failure to recognize and respond to variations in the effectiveness of our teachers." (from Executive Summary)
- Value-added research shows that teachers vary greatly in their contributions to student achievement (Rivkin, Hanushek, & Kain, 2005).

Measures of teacher effectiveness

- **Evidence of growth in student learning and competency**
 - Standardized tests, pre/post tests in untested subjects
 - Student performance (art, music, etc.)
 - Curriculum-based tests given in a standardized manner
 - Classroom-based tests such as DIBELS
- **Evidence of instructional quality**
 - Classroom observations
 - Lesson plans, assignments, and student work
- **Other evidence (varies, based on local values)**
 - Administrator/supervisor reports
 - Surveys of students and/or parents
 - An “evidence binder” created & presented by the teacher

Measures

- Many types of evidence*—including portfolios, administrator recommendations, analysis of teachers' assignments, analysis of students' work, documentation of teachers' positive contributions to the school, student and parent reports, and 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).

Teaching Standards

- All measures used as part of the teacher evaluation system should align with the state's teaching standards
- One way to do this is to create instruments that indicate which standard(s) are being measured
- If available instruments are used instead, go through and indicate which standard(s) are being measured in order to determine coverage and alignment

Teacher observations

➤ Strengths

- Great for teacher formative evaluation
- Helps evaluator understand teachers' needs across school or across district

➤ Weaknesses

- Only as good as the instruments and the observers
- Considered "less objective"
- Expensive to conduct (personnel time, training, calibrating)
- Validity of observation results may vary with who is doing them, depending on how well trained and calibrated they are

Threats to validity in observations

- The instrument is not sensitive enough to detect differences in teacher effectiveness
- The instrument doesn't measure what is important and valued to all stakeholders, given the purpose of the observation
- The raters are poorly trained in the use of the observation instrument
- The raters have "drifted" over time, meaning they are no longer using the instrument in the way it was intended to be used

Example: University of Virginia's CLASS observation tool

	Emotional Support	Classroom Organization	Instructional Support
Pre-K and K-3	<p>Positive Climate</p> <p>Negative Climate</p>	<p>Behavior Management</p> <p>Productivity</p>	<p>Concept Development</p> <p>Quality of Feedback</p> <p>Language Modeling</p>
Upper Elementary/ Secondary	<p>Teacher Sensitivity</p> <p>Regard for Student (Adolescent) Perspectives</p>	<p>Instructional Learning Formats</p>	<p>Content Understanding</p> <p>Analysis and Problem Solving</p> <p>Quality of Feedback</p>

Example: Charlotte Danielson's Framework for Teaching

Domain 1: Planning and Preparation

includes comprehensive understanding of the content to be taught, knowledge of the students' backgrounds, and designing instruction and assessment.

Domain 3: Instruction is concerned with the teacher's skill in engaging students in learning the content, and includes the wide range of instructional strategies that enable students to learn.

Domain 2: The Classroom

Environment addresses the teacher's skill in establishing an environment conducive to learning, including both the physical and interpersonal aspects of the environment.

Domain 4: Professional

Responsibilities addresses a teacher's additional professional responsibilities, including self-assessment and reflection, communication with parents, participating in ongoing professional development, and contributing to the school and district environment.

Example: Kim Marshall's Rubric

<i>Planning & Preparation for Learning</i>				
	Highly Effective	Effective	Improvement Necessary	Does Not Meet Standards
a. Knowledge	Is expert in the subject area and has a cutting-edge grasp of child development and how students learn.	Knows the subject matter well and has a good grasp of child development and how students learn.	Is somewhat familiar with the subject and has a few ideas of ways students develop and learn.	Has little familiarity with the subject matter and few ideas on how to teach it and how students learn.
b. Strategy	Has a well-honed game plan for the year that is tightly aligned with state standards and assessments.	Plans the year so students will meet state standards and be ready for external assessments.	Has done some thinking about how to cover high standards and test requirements this year.	Plans lesson by lesson and has little familiarity with state standards and tests.

Show me the evidence!

- Example of an instrument for evaluating teachers: Kim Marshall's rubric
- What constitutes *evidence* for the following?
 - Classroom management (j): *Gets students to buy into a highly effective system of incentives linked to intrinsic rewards.*
 - Monitoring...(i): *Works with colleagues to analyze and chart assessment data, draw action conclusions, and share them with others.*

Show me the evidence! (cont'd)

- What constitutes *evidence* for the following?
 - Family & Community Outreach (e): *Frequently involves parents in supporting and enriching the curriculum as it unfolds.*
 - Professional Responsibility (j): *Devours best practices from fellow professionals, workshops, reading, study groups, the Internet, and other sources.*

Portfolios

➤ Electronic Teaching Portfolio (generic)

- <http://electronicportfolios.com/portfolios/site2000.html>
- “The process of developing electronic teaching portfolios can document evidence of teacher competencies and guide long-term professional development. The competencies may be locally defined, or linked to national teaching standards. Two primary assumptions in this process are: 1.) a portfolio is not a haphazard collection of artifacts (i.e., a scrapbook) but rather a reflective tool which demonstrates growth over time; and 2.) as we move to more standards-based teacher performance assessment, we need new tools to record and organize evidence of successful teaching, for both practicing professionals and student teachers.”

Portfolios (cont'd)

➤ Kansas Performance Teaching Portfolio

- <http://www.ksde.org/LinkClick.aspx?fileticket=E8ha-oFtX7Q%>

➤ New teachers complete a set of guided tasks

- Task #1. Contextual Information & Learning Environment Factors
 - Who you will be teaching
- Task #2. Designing Instruction
 - What you will be teaching
- Task #3. Teaching and Learning
 - How you will be teaching
- Task #4. Reflection and Professionalism
 - Examine and reflect

Portfolios (cont'd)

➤ Washington Teacher Portfolio System (ETS)

- www.waproteach.org
- Practicing teachers in Washington must pass the assessment within five years of initial licensure in order to gain professional certification. Candidates are required to submit three "entries." The entries, which measure the Washington Professional Certificate Standards and Criteria, include Professional Growth and Contributions; Building a Learning Community; and Curriculum, Instruction, and Assessment. Scores are assigned for each of the 12 criteria that comprise the three Professional Certificate Standards. Candidates receive feedback to help them determine which entries to retake should they not meet that standard.

Portfolios (cont'd)

➤ New Mexico's Professional Development Dossier

- <http://teachnm.org/programs/professional-development-dossier.html>
- "The Professional Development Dossier (PDD) is the cornerstone of licensure advancement for teachers in the 3-Tiered Licensure System."
- "...intent of the Requirements and Guidelines is to allow you to tell a set of stories from your practice — stories that are illustrated and validated by concrete evidence from your work with your students and others in your classroom. The PDD is not a test with the right answer. Instead it allows you to construct a picture of your teaching."

Evidence binders: “Next generation” portfolios

- Not all evidence can be collected in an observation
- Focus on *specific evidence*, not random “stuff”
- Evidence binders can be physical or virtual
- Provide a place for teachers to document
 - Evidence of practice, such as using formative assessment or differentiating lessons
 - Evidence of professional growth, such as participation in learning communities
 - Evidence of contribution to the school and community, such as working on intervention teams

Using Evidence Binders

- Teachers have a choice about what activities to complete but consulting teachers or administrators may recommend particular activities for professional growth
- The number of activities a teacher completes may vary depending on other components in the 60%
- Scoring processes have to be determined, i.e., ensuring comparability across a district

Artifacts

➤ SCOOP Notebook

- <http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED495853>
- “The goal of the project was to use artifacts and related materials to represent classroom practice well enough that a person unfamiliar with the teacher or the lessons can make valid judgments about selected features of practice solely on the basis of those materials”
- “...our analyses indicate that the reliability and validity of the ratings are currently not sufficient to justify the use of the Scoop Notebook for making high-stakes decisions about individual teachers. It might be possible, however, to use the notebook in combination with other measures of teacher background, attitudes, or content knowledge, as part of a system of indicators that could support valid inferences about individual classroom practice in high-stakes situations.

Artifacts (cont'd)

➤ Analysis of Unit Plan and Lesson Plan

- http://www.state.tn.us/education/frameval/doc/comprehensive_assessment.pdf
- The evaluator forms for analyzing and comments on teachers' plans are on pgs. 49-51
- Teacher and evaluator examine the Unit Plan and Lesson Plan together and evaluator completes the document

Artifacts

➤ Teacher Work Sample

- <http://www.wou.edu/~girodm/library/ethnog.html>
- “A teacher work sample is a vehicle to help teacher candidates articulate, document, investigate, and reflect on their teacher actions and their impact on student learning.”
- “Typically, a work sample has the following components: 1) a description of the context in which the teacher candidate will be teaching; 2) an articulation of the learning goals to be targeted; 3) a rationale statement, defending the goals as important, and the pedagogical moves as sound, in light of the context; 4) lesson plans; 5) pre and post tests; 6) analysis of testing results and student learning gains, and; 7) reflections on teaching efforts, student learning, and personal professional growth.”

Measures that help teachers grow

- Measures that motivate teachers to examine their own practice against specific standards
- Measures that allow teachers to participate in or co-construct the evaluation (such as “evidence binders”)
- Measures that give teachers opportunities to discuss the results with evaluators, administrators, colleagues, teacher learning communities, mentors, coaches, etc.
- Measures that are directly and explicitly aligned with teaching standards
- Measures that are aligned with professional development offerings
- Measures which include protocols and processes that teachers can examine and comprehend

Evidence-based improvement

- Create a culture of using evidence to drive improvements in teaching and learning
- Content is necessary but not sufficient—it's no substitute for improved teaching practices
- Good teacher evaluation should provide evidence about teachers' practice that can be used to help teachers improve their instruction
- In the CTES, keep the focus on improving instructional quality, not raising test scores

Decisions

- What will be the criteria for approval of a measure or instrument?
 - Reliability/validity
 - Local capacity and resources to use the measure
 - Alignment with and coverage of standards
- How much local flexibility will there be to choose components?
- How much flexibility will there be in weighting components to result in a final score?

Decisions (cont'd)

- Who will conduct observations?
 - Ensure validity and reliability by requiring training (possibly certification) to conduct observation
- How often will teachers be observed?
 - Might be different for new vs. tenured teachers
- Will teachers have choices, i.e., whether to be observed or complete a portfolio?

References

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417 - 458.

<http://www.econ.ucsb.edu/~jon/Econ230C/HanushekRivkin.pdf>

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. Brooklyn, NY: The New Teacher Project.

<http://widgeteffect.org/downloads/TheWidgetEffect.pdf>

Resources

Little, O., Goe, L., and Bell, C. (2010). *A practical guide to evaluating teacher effectiveness*. Washington, DC: National Comprehensive Center for Teacher Quality.

<http://www.tqsource.org/publications/practicalGuide.pdf>

Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality.

<http://www.tqsource.org/publications/EvaluatingTeachEffectiveness.pdf>

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality.

[http://www.tqsource.org/publications/
LinkBetweenTQandStudentOutcomes.pdf](http://www.tqsource.org/publications/LinkBetweenTQandStudentOutcomes.pdf)

Observation instruments

Charlotte Danielson's Framework for Teaching

<http://www.danielsongroup.org/theframeteach.htm>

CLASS

<http://www.teachstone.org/>

Kim Marshall Rubric

[http://www.marshallmemo.com/articles/Kim%20Marshall%20Teacher%20Eval%20Rubrics%20Jan%](http://www.marshallmemo.com/articles/Kim%20Marshall%20Teacher%20Eval%20Rubrics%20Jan%20)

Questions?





Laura Goe, Ph.D.

P: 609-734-1076

E-Mail: 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