

Getting Teacher Evaluation Right: Adding Value to Teaching and Learning



Getting Value from Value-Added

Value-added methods offer benefits for:

- Measuring progress rather than status
- Researching effects of practices and programs in large-scale data sets

At the individual level, VAM has difficulty:

- Disentangling other influences on student learning
- Measuring gains (especially on state tests)
- Attributing gains to individual teachers

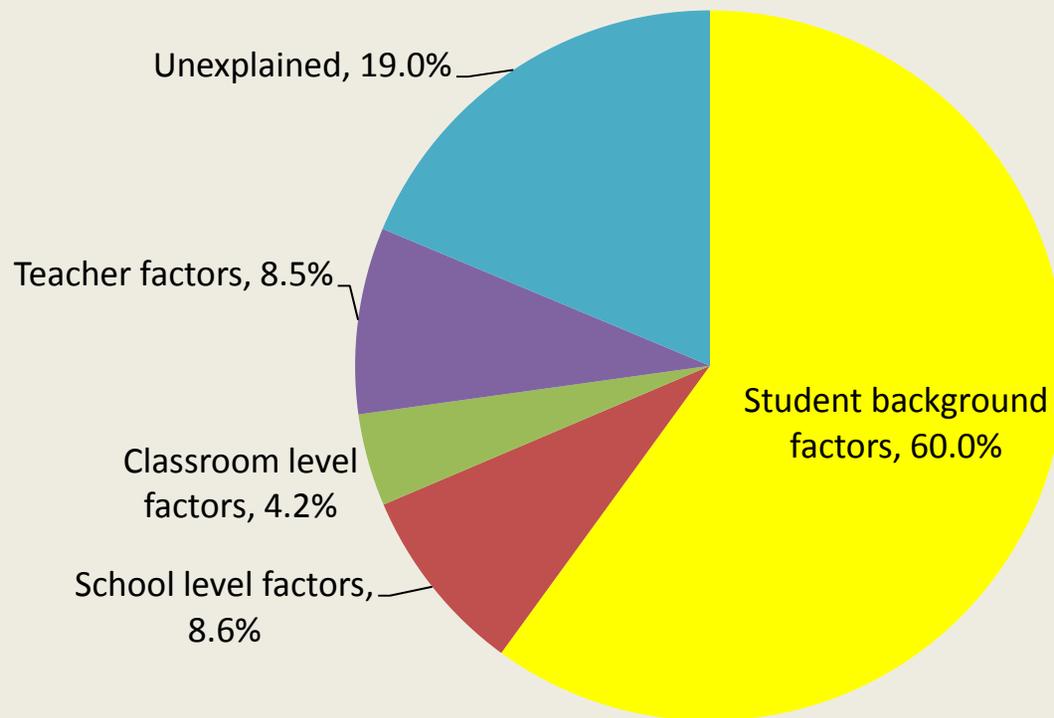
Teacher Effectiveness in Context

Studies find that teachers' value-added "effectiveness" is influenced by:

- Collaboration with peers
- Class size, curriculum, and time with students
- Tutoring; out-of-school learning; summer loss
- School organization & administration
- Student characteristics and attendance
- The measure of achievement used and the statistical model used

Many Things Influence Achievement Gains

Variance in Achievement*
(10th grade Mathematics)

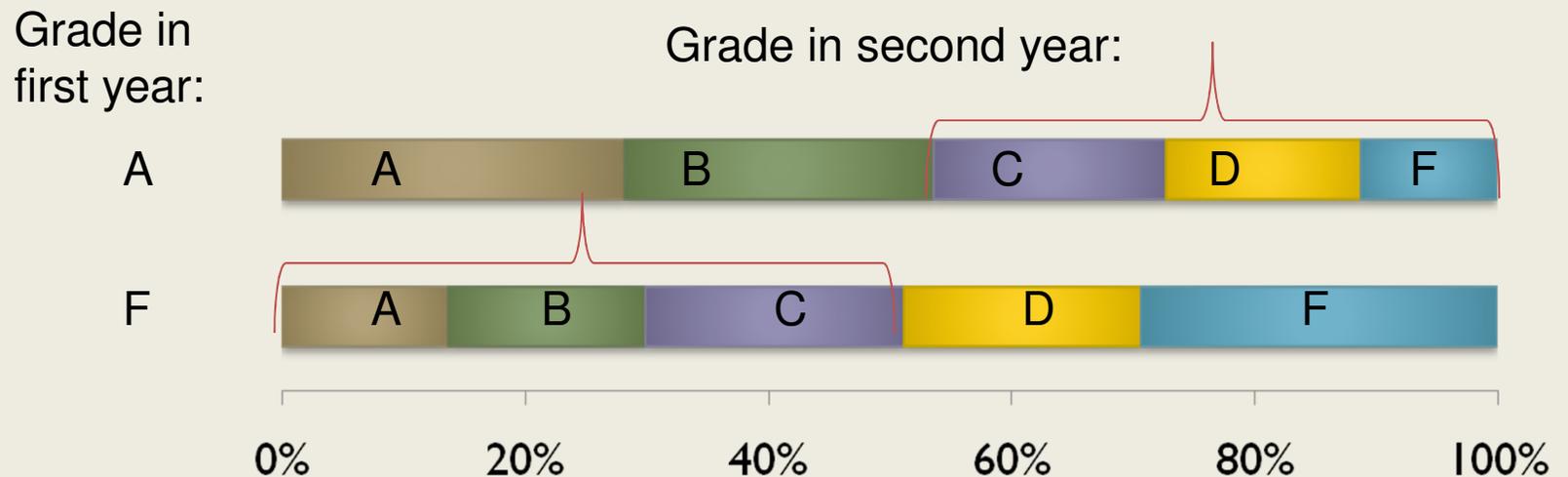


- Model Controls for Prior Achievement
- 95% of school, classroom, and teacher effects were unobservable

Source: Goldhaber, Brewer, & Anderson, *Education Economics*, 7 (3), 1999.

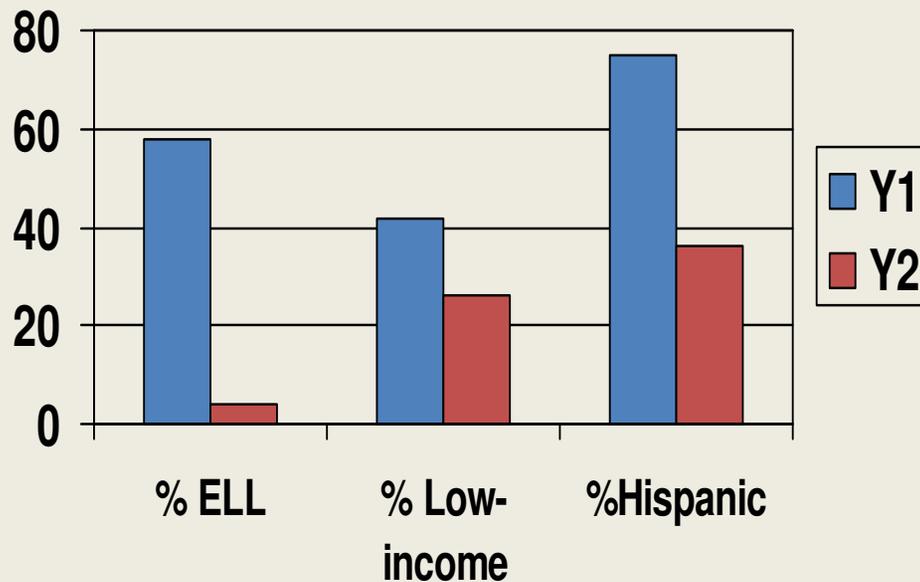
Value-Added Estimates Are Highly Unstable

- Many teachers indicated as effective or ineffective in one year are not for others
- Consider classification of teachers into 5 categories (A-F) in two consecutive years.



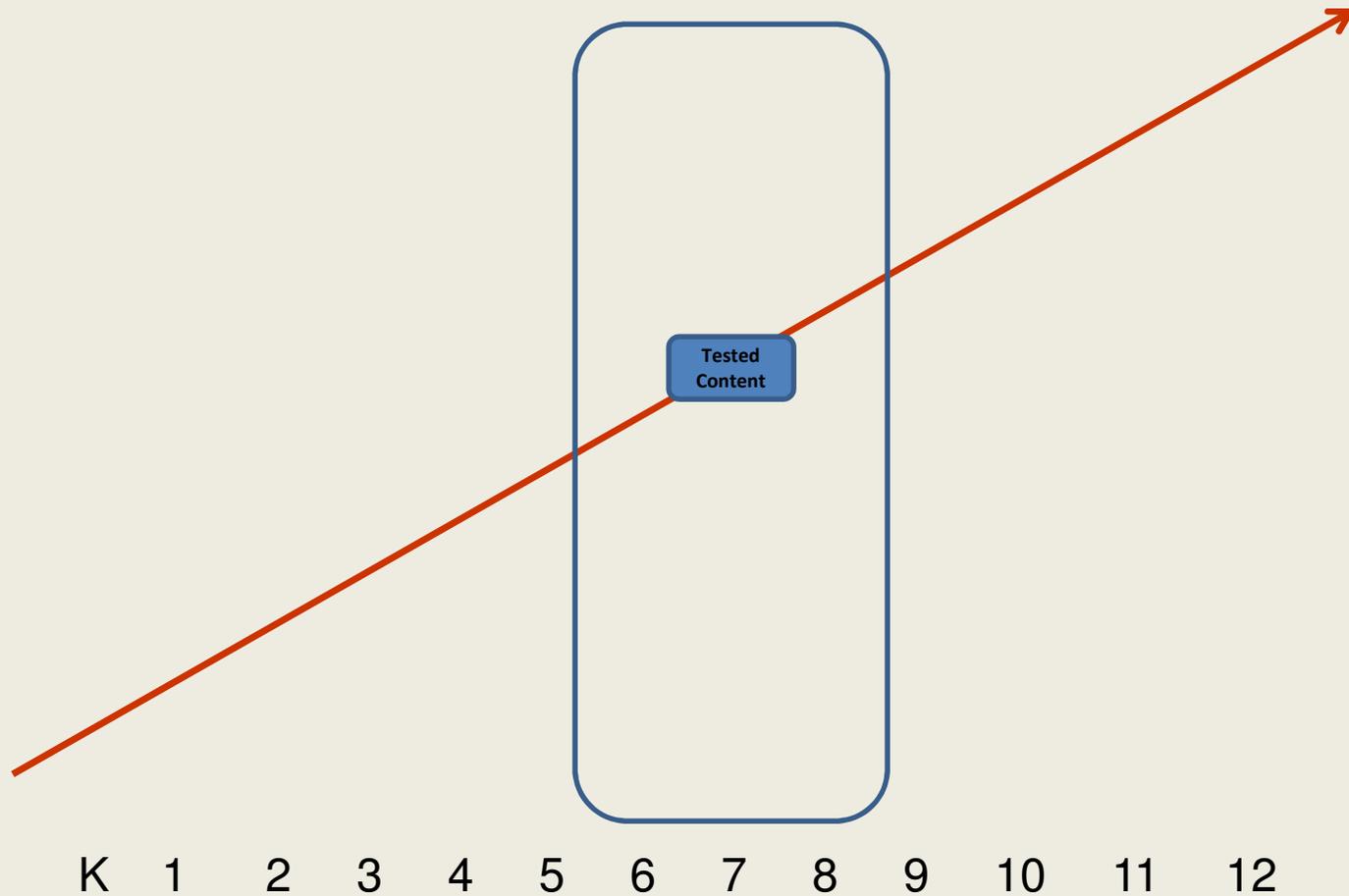
Grades A-F correspond to quintiles 1-5. Source: Tim Sass (2008).

A Teacher's Measured "Effectiveness" Can Vary Widely: 'Maria's' Case



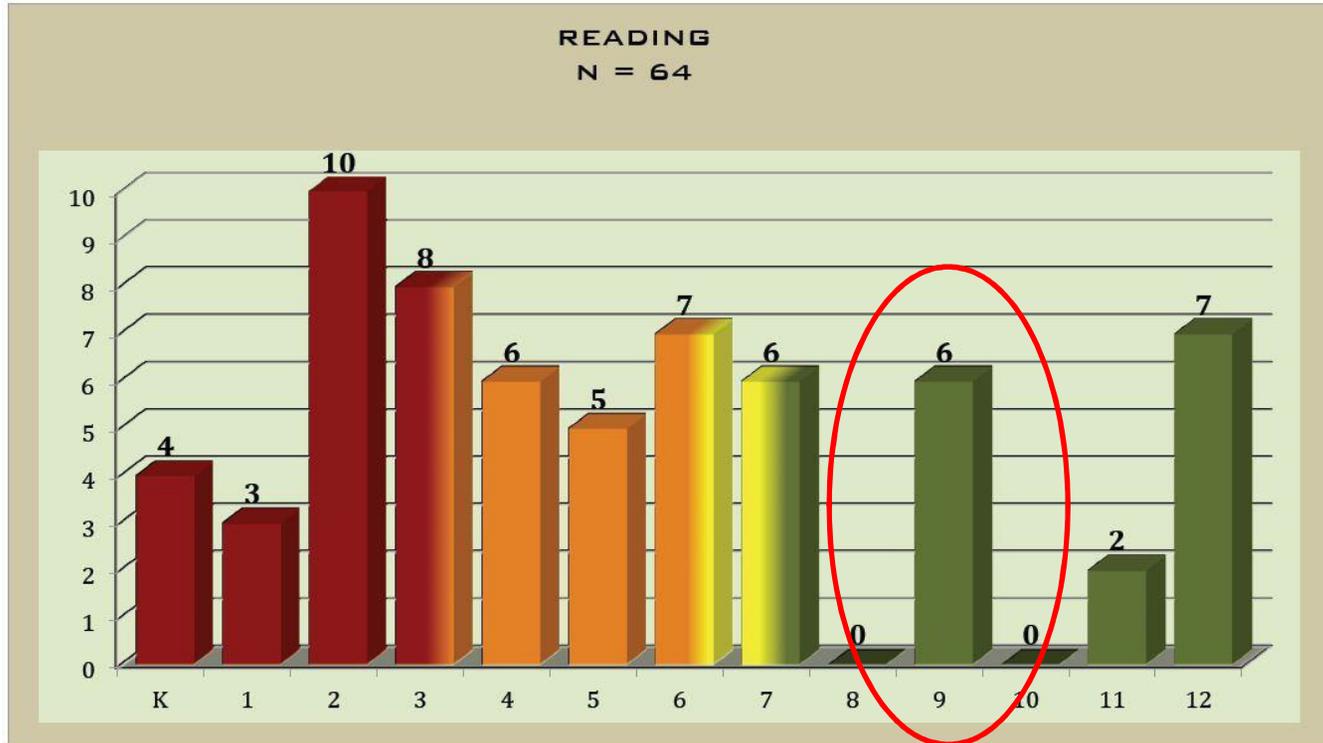
- Same high school
- Same course (English I)
- Not a beginning teacher
- Model controls for:
 - Prior achievement
 - Demographics
 - School fixed effects

Grade-Level Tests Cannot Measure Growth for Many Students



Entering 9th Graders: Reading, 2011

MEASURES OF ACADEMIC PROGRESS (MAP) TEST
GRADE 9, FALL 2011
GRADE LEVEL PERFORMANCE FREQUENCIES IN MATHEMATICS AND READING



Use of Invalid Metrics Mismeasures Teachers: Carolyn's Case

"Ms. Abbott is a great teacher."

"Ms. Abbott is my son's teacher, and she has been fabulous. He has never learned more math than he has this year.

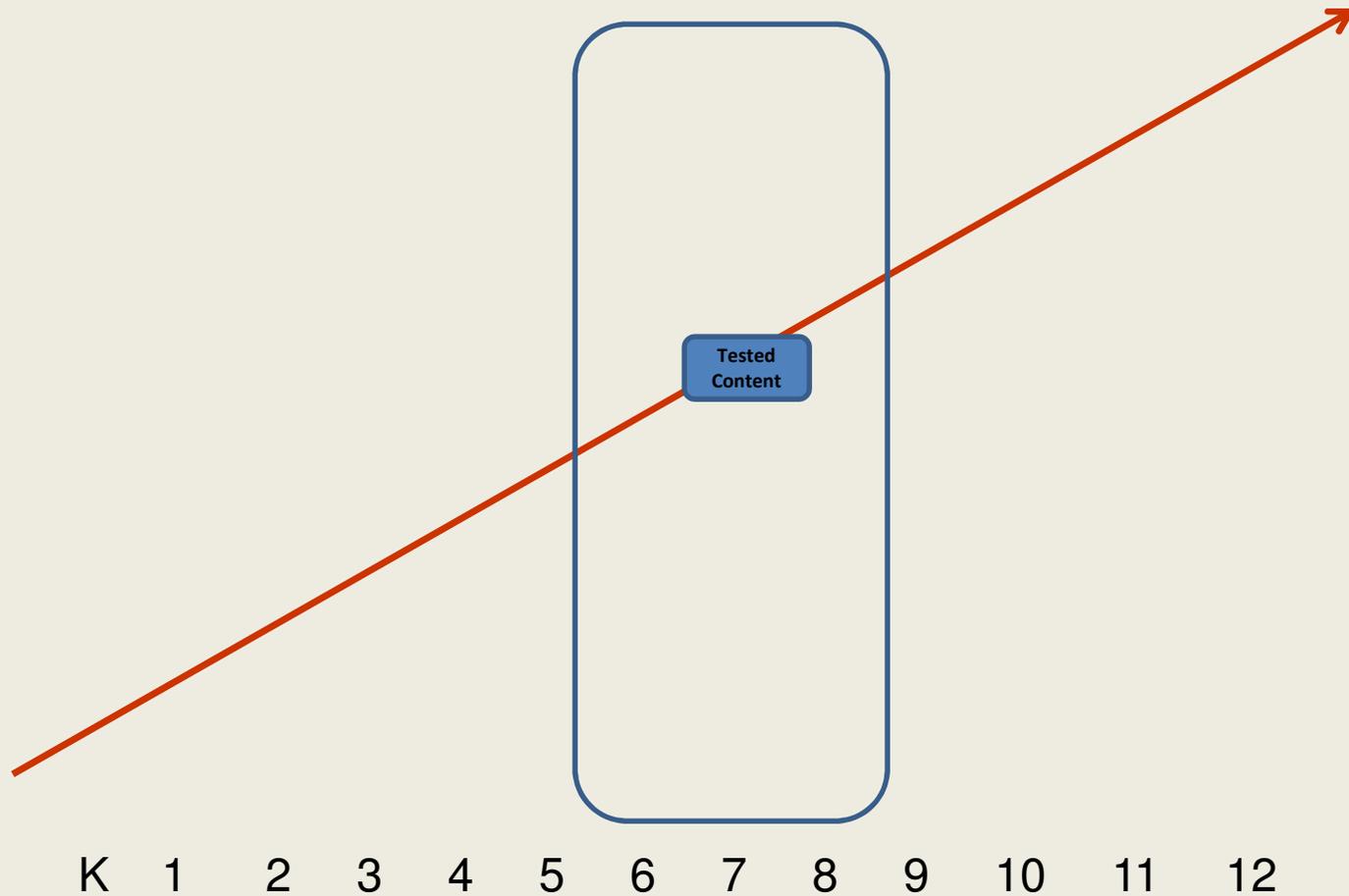
"Ms. Abbott is one of my daughter's favorite teachers. She helped my daughter develop a love of math, and the kids really enjoy her class."

"I always loved Ms. Abbott. She is and was an amazing teacher, and no teacher evaluation will ever change that."

"I am a sixth grader at Anderson, and I have been looking forward to having math class with Ms. Abbott since the fourth grade.... She is one of the best math teachers ever, in my opinion, and I would do anything to keep her here at Anderson."

The "Worst" 8th Grade Math Teacher
in New York Based on VAM

Grade-Level Tests Cannot Measure Growth for Many Students



Recent findings from Houston

- Teachers teaching in grades in which English Language Learners (ELLs) are transitioned into mainstreamed classrooms are the least likely to show “added value.”
- Teachers teaching larger numbers of special education students in mainstreamed classrooms are also found to have lower “value-added” scores in those years.
- Teachers teaching gifted students have small gains because their students have already reached the top of the scale.

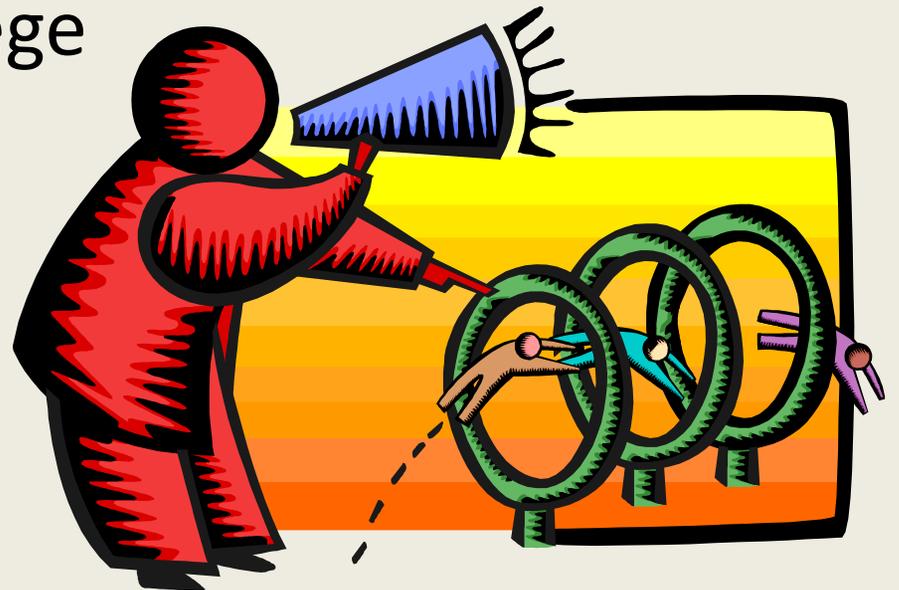
(Amrein-Beardsley & Collins, Educational Policy Analysis Archives, 2012)

Counterproductive Incentives

- “I’m scared to teach in the 4th grade. I’m scared I might lose my job if I teach in an [ELL] transition grade level, because my scores are going to drop, and I’m going to get fired.”
- “When they say nobody wants to do 4th grade – nobody wants to do 4th grade! Nobody.”

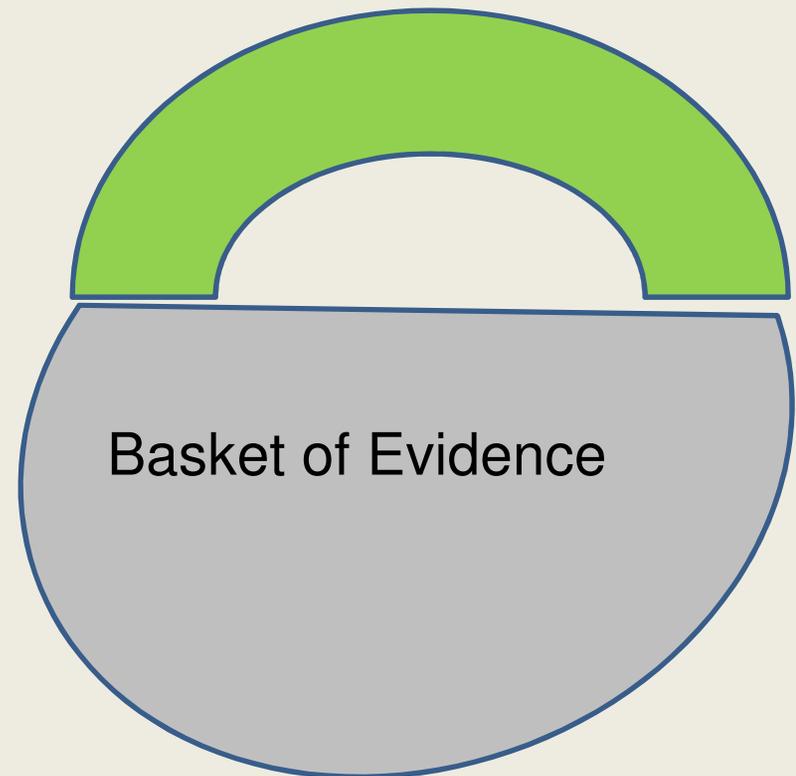
Problems with Current Tests

- Current tests, in the context of high-stakes accountability ...
 - fail to measure higher order skills
 - narrow curriculum to topics / formats tested
 - cause neglect of college and career ready skills



Solution: Use Multiple Measures to Assess Student Learning

- Authentic assessments of learning: essays, research projects, investigations; portfolios
 - Local, state, national assessments (pre - post measures; AP; IB; DRA)
 - English proficiency gains
 - Accomplishment of learner goals (IEP, other)
 - Documentation of learning
- Goal-Setting Matched to Curriculum and Students Taught



An Integrated System

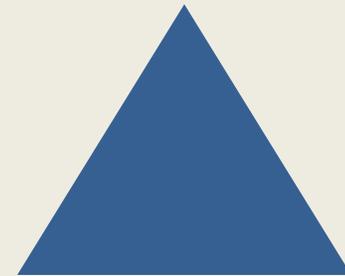
- Standards-based measures of practice
- Evidence of collegial contributions
- Measures of learning & growth

Linked by goals,
considered in context



Integrated evidence of:

practice



professional
contributions

student
learning

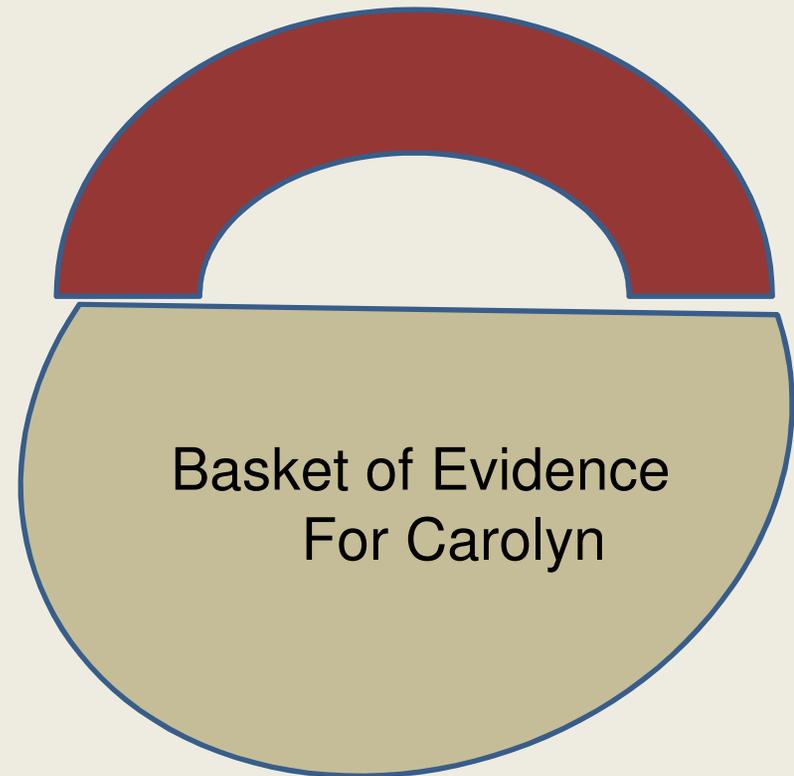
Example: Maria's Basket

- English language proficiency gains
- Qualitative Reading Inventory gains
- 9th grade Autobiography paper / exhibition (video) illustrating progress across multiple drafts (linked to curriculum artifacts and teacher feedback)
- School fall and spring scored writing samples



Example: Carolyn's Basket

- Regents Exam outcomes with fall pre-test on Regents test material
- Pre- and post-test for unit on quadratic equations (Carolyn's focus this year) linked to curriculum and teaching evidence from that unit
- Scores & work samples for the 8th grade math/science inquiry project, with focus on students' applications of mathematics, as compared to the previous year
- Student survey of attitudes about math (pre- and post)



Student Learning Objectives (SLOs)

- A goal-setting process in which teachers assemble learning evidence to evaluate how students are making progress.
- Teachers set targets on assessment measures—for entire classrooms or specific subsets of students, for an entire course or a specific set of skills within the course—taking into account students' starting points.
- The goals and evidence of student learning are examined, along with evidence of teachers' practices, in the evaluation process.



What might SLOs (and other Collections of Data) include?

Primarily, work that emerges directly from ongoing classroom practice:

- Performance-based assessments, such as papers, presentations, projects, and investigations;
- Classroom tests or rubric-scored tasks (before and after a unit; semester; or year)
- Portfolios of student work, with samples illustrating knowledge and skills before, sometimes during, and after a learning experience;
- School, district, state, or national assessments

-- American Institutes of Research

How Evaluation Can Support Collaboration

[Student learning evidence] can be used to promote collaboration and reflection on practice among educators. By analyzing work in teams and by setting schoolwide, subject-level, grade-level, or team-based goals, educators can work together to better understand and meet the needs of all students. The goal-setting and inquiry cycle encourages educators to share strategies and seek guidance and assistance from specialists who support student learning....

Long Beach Approach

- Individuals, grade-level / department teams, and schools set goals
- Multiple sources of evidence are examined collectively to evaluate progress for students and to make plans for next year. These can include:
 - Teacher observations of learning
 - Anecdotal and cumulative records
 - Success and progress on a continuum of learning or a course of study
 - Teacher, department, or school-made pre- and post-tests
 - Curriculum-related tests
 - Use of audio-visual documentation
 - Student self-evaluations
 - Evaluative discussion with students and parents
 - Records of students' past learning performances
 - Files of students' work collected to show growth
 - Action research

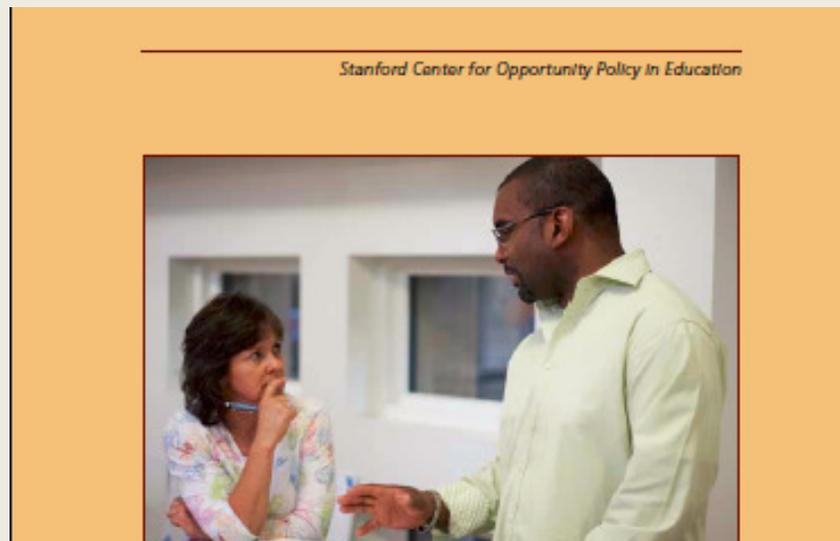


Hillsdale High School Approach (San Mateo, CA)

- Teachers work with principal to set specific goals around 2 CSTP standards (e.g. student engagement; instructional practice)
- Teachers write a self-evaluation which becomes an iterative conversation with the principal via googledocs as observations occur and new strategies are tried
- Observation of practice and collection of evidence about learning drive conversations throughout the year
- Teachers assemble a portfolio of evidence about how they have met the standards, which integrates evidence of practice, student learning, and professional contributions

New Mexico Approach: South Valley Academy

- 3-tiered state portfolio-based licensing system
- Within the school, PDP (Professional Development Plan) used for goal-setting; focused collaborative work; and assessment of results -- A practitioner action research evaluation model



Case in Point

- Andres' focus on improving vocabulary learning in chemistry
 - measurable goals [58% to 80% mastery]
 - a variety of strategies, developed with peers
 - tracking of outcomes on multiple measures [14 artifacts collected in his portfolio with analysis of student learning]

Andres' Reflection

“These collaborations are perhaps the most powerful and useful part of the PDP process. Staff became aware of the many learning gaps that we are noticing in our students, and we began working collaboratively to share and find ways to address those learning gaps. Many times other staff members noticed things that were happening that I hadn't noticed before, but as soon as they said it, I was like, “Yeah, that's happening with my students, too.” In many of the meetings, teachers together can start noticing trends, and then the project is even more meaningful....

Having a forum to share effective strategies for student learning opens the door for every teacher at the school to help every other teacher get better. When other people start doing new things because of other people's PDPs, this makes the process even more powerful.”

If We Take Teaching Effectiveness Seriously, We Should Ensure...

- High-quality preparation and mentoring for all beginners
- Teaching assignments and conditions that support success for students and teachers
- Sustained, collegial learning opportunities for all educators
- Well-designed schools with thoughtful, coherent curriculum
- Evaluation systems that are
 - Standards-based, integrating practice and outcome data
 - Linked to coaching and professional development
 - Rigorous at tenure and other key junctures
 - Designed to provide assistance, due process, & timely decisions