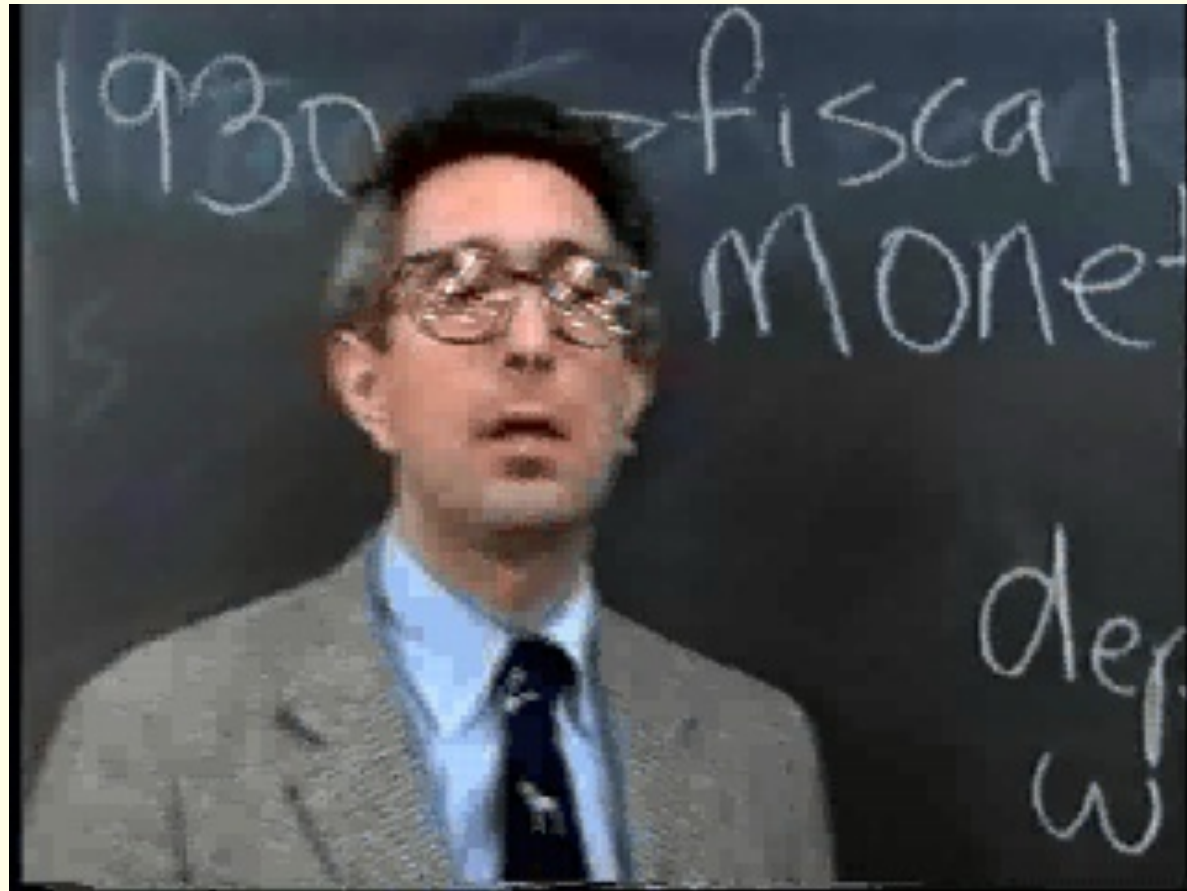

Getting Teacher Evaluation Right



The Need for More Powerful Teaching





What's Wrong with Teacher Evaluation Now?

I have had administrators who never came into my classroom for formal observations or asked me for anything more than the initial planning/goal sheet. I have had administrators observe a formal lesson and put the feedback sheet in my box without ever having spoken to me about the lesson, and I have had years where I am just asked to sign the end-of-the year evaluation sheet [without being observed].

*—Jane Fung, National Board Certified teacher
and Milken Award Winner, CA*

“The evaluations themselves are typically of little value—a single, fleeting classroom visit by a principal or other building administrator untrained in evaluation, wielding a checklist of classroom conditions and teacher behaviors that often don’t even focus directly on the quality of teacher instruction. It’s typically a couple of dozen items on a list: “Is presentably dressed,” “Starts on time,” “Room is safe,” “The lesson occupies students.... But, in most instances, it’s nothing more than marking ‘satisfactory’ or ‘unsatisfactory.’”

-- Toch & Rothman, 2008.

Why has Teacher Evaluation been So Awful?

Lack of

- Clear standards for practice
- Time
- Expertise
- Links to professional development

Little attention to student learning

Unwieldy processes for making decisions

How We Might Make Evaluation Even Worse

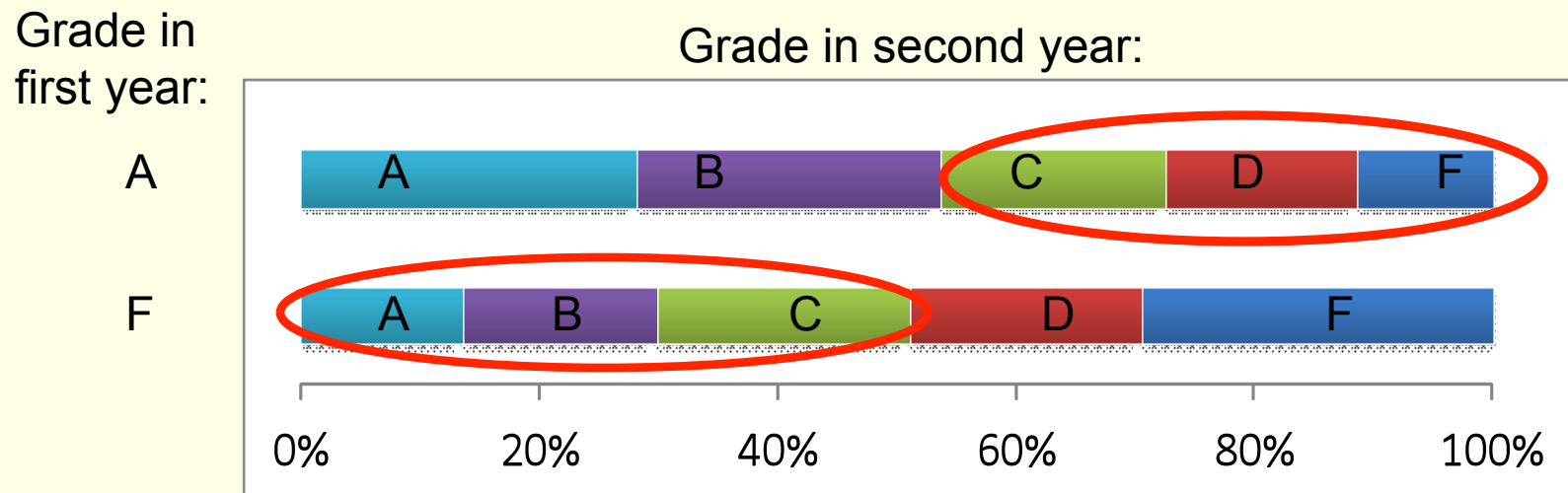
- Focusing evaluation entirely after entry
- Putting all of the weight on principals
- Designing systems that cannot be implemented
- Creating systems that
- Making decisions based on value-added test scores

Researcher Concerns

- VAM estimates of teacher effectiveness ... should not be used to make operational decisions because such estimates are far too unstable to be considered fair or reliable.
- *National Research Council
Board on Testing and Assessment,
2009*

Many teachers indicated as effective or ineffective in one year are not for others

- Value-added estimates are highly unstable.
- For example, teachers classified in 5 categories (A-F) in two consecutive years:



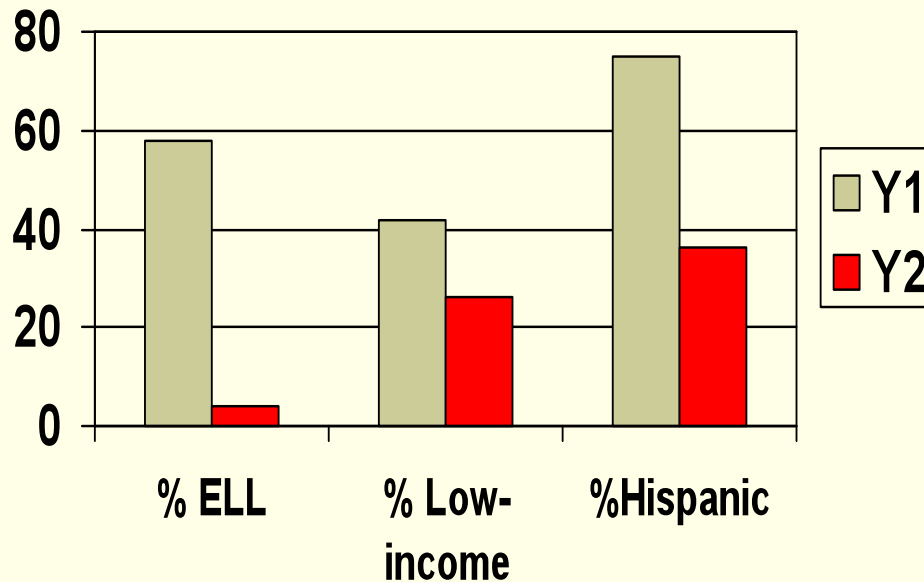
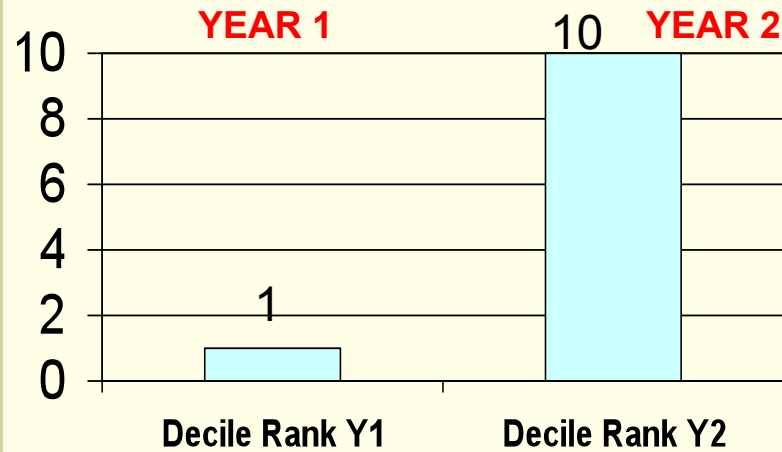
Grades A-F correspond to quintiles 1-5. Source: Tim Sass (2008).

Three True Stories

- California
- New York
- Texas



A Teacher's Measured "Effectiveness" Can Vary Widely



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

The “Worst” 8th Grade Math Teacher in New York

“Ms. Abbott is a great teacher.”

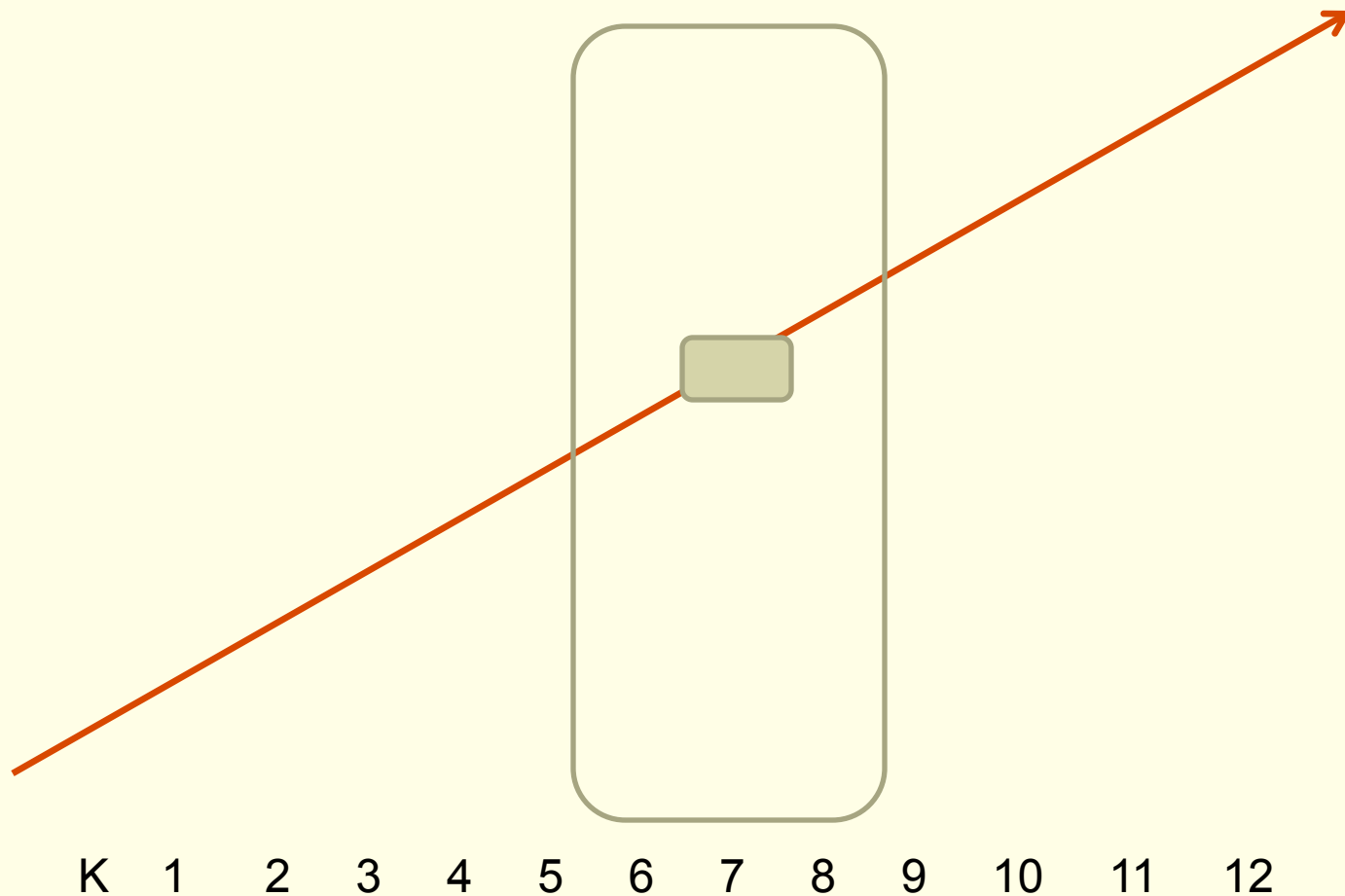
“Ms. Abbott is my son’s teacher, and she has been fabulous. He has never learned more math than he has this year.

“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.”

What Grade-Level Tests Measure



Can Random Rewards and Punishments Guide Practice?

I do what I do every year. I teach the way I teach every year. [My] first year got me pats on the back. [My] second year got me kicked in the backside. And for year three, my scores were off the charts. I got a huge bonus, and now I am in the top quartile of all the English teachers. What did I do differently? I have no clue.

-- A Houston teacher commenting on the annual changes in her value-added test scores

Worse Still: What are the Incentives Created by Bias in the Measures?

- 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)

How to Do Teacher Evaluation Right

- 1. Think Systemically
- 2. Start with Standards
- 3. Create Performance Assessments
- 4. Build Systems that focus on practice, collegial contributions, and student learning
- 5. Use Evidence of Learning Appropriately
- 6. Support Meaningful Professional Learning
- 7. Create Structures for Success
 - Training and Mentoring
 - Peer Assistance and Review

If We Took Teaching Effectiveness Seriously, We Would Ensure...

- High-quality preparation and mentoring for all beginners
- Teaching assignments and conditions that support success for students and teachers
- Sustained, practice-based 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
 - Able to provide assistance, due process, & timely decisions

Comprehensive evaluations—with standards and scoring rubrics and multiple classroom observations ... and a role for student work and teacher reflections—are valuable regardless of the degree to which they predict student achievement, and regardless of whether they're used to weed out a few bad teachers or a lot of them. They contribute to a much more professional atmosphere in schools. As a result, they make public school teaching more attractive to the sort of talent that the occupation has struggled to recruit and retain. Capable people want to work in environments where they sense they matter.... Comprehensive evaluation systems send a message that teachers are professionals doing important work.

-- Thomas Toch & Robert Rothman (2008)