



Designing and Implementing “Next Generation” Teacher Evaluation Systems: Lessons Learned from Case Studies of Illinois Districts

Brad White, Illinois Education Research Council

Sue Sporte, Jen Cowhy, and David Stevens, University of Chicago
Consortium on Chicago School Research

Overview of PERA

- Sets parameters for teacher (and principal) evaluation in Illinois
 - Must have a 4-level rating system with minimum specified evaluation cycle
 - Must include both teacher performance (observation-based) and student growth components
 - Student growth must account for at least 25% in first two years of implementation, at least 30% thereafter
 - Multiple types of growth measures specified
- All districts must implement teacher component in 2016-17 school year
 - If a district cannot agree on a plan, it will default to the state model (which includes 50% student growth)

Purpose of Study

- We focused solely on teacher (NOT principal) evaluation
- Analyze key lessons learned in Illinois districts identified as leaders in teacher evaluation
- Inform other districts as they design and implement "next generation" teacher evaluation systems under PERA.

Methodology: Case Studies

- Spring-summer 2012
- Started with policy scan of 13 districts nominated by those with statewide knowledge
- Narrowed to five districts (see next slide) based on: geographic and demographic diversity, program components, stage of implementation
- In-depth interviews on policy design, implementation, and perceptions – four to six subjects in each district
 - District administrators (n=9)
 - Teachers/union representatives (n=11)
 - Principals/evaluators (n=7)

Case Study Districts



Characteristics of the Case Study Districts

■ District characteristics

- Range of sizes (> 30,000 to <1,000 students)
- Range of locales (1 urban, 2 suburban, 2 rural)
- Geographic range (3 northeast, 1 central, 1 southern)
- Range of student populations (30%-70% FRL and 30%-95% white)

Characteristics of the Evaluation Systems

- Various stages of implementation
- All used some form of the Danielson Framework for teacher performance measure
- Only one included student growth in teacher evaluations
- Observation cycles tended to hew to state minimum guidelines

Expected Opportunities? Challenges?

Overview

- **These systems were seen as a huge upgrade over the status quo**
- **But there were still some challenges...**
 1. Securing understanding and buy-in
 2. Using observations to improve instruction
 3. Building the capacity of evaluators
 4. Incorporating student growth
- *The rest of this presentation will focus on the specific challenges in these four areas – and strategies the case study districts used to overcome these obstacles*

Challenge #1: Cultivating Understanding & Buy-In

- Initial training typically did not include teachers
- Success highly dependent upon principal understanding and buy-in
- Some distrust between teachers and administration, anxiety surrounding PERA
- Buy-in problematic, especially in the first year

Cultivating Understanding & Buy-In

- **Strategy: Gather all perspectives**
 - Ensure balanced representation and diverse perspectives on design committee
 - Gather ideas from external facilitators, consultants, and other districts
 - Ongoing monitoring necessary

Cultivating Understanding & Buy-In

- **Strategy: Develop a shared vision of quality instruction**
 - Clear, evidence-based observation standards, rubrics, and performance levels (such as the Danielson Framework) were viewed as especially helpful
 - Common vision and language for quality instruction can provide next steps for improvement
 - Teachers held accountable for high expectations & evaluators accountable for performance management

Cultivating Understanding & Buy-In

- **Strategy: Teachers need to be informed early, continuously, and consistently**
 - Transparency in design process
 - Consider training with teacher-administrator teams
 - Pitfalls of voluntary/optional training sessions
 - In order for system to reach its full potential to improve instruction, teachers need to understand standards and rubrics, not just the evaluation processes

Cultivating Understanding & Buy-In

- **Strategy: Align teacher evaluation with other district priorities**
 - Link to existing human resources systems (hiring, mentoring, professional development, etc.) or other district initiatives (Common Core, school improvement plan, etc.)
 - Can reinforce other initiatives and serve as the “glue that holds the rest together” around quality instruction

“When we...constructed [wording of] the new evaluation tool, what we did was, we tried to take the different things that are going on in the district...the things that we value...and we really tried to build it into the evaluation tool, and by doing that, really sort of cementing it for us as a district.”

Cultivating Understanding & Buy-In

- **Strategy: Start soon and implement gradually**
 - Consider phasing in with low-stakes pilot
 - Try not to start or make major changes mid-year
 - Start planning/designing ASAP to give yourself enough time to refine the plan and hit the ground running

Comments/Questions about Understanding & Buy-In?

Challenge #2: Using Observations to Improve Instruction

- Systems good at pinpointing weaknesses, less successful transforming these into strengths
 - Teachers craved constructive criticism, but principals found it difficult to have “tough conversations” with teachers to coach teachers on how to progress from one performance level to the next
- Still some concerns with the accuracy and usefulness of evaluation ratings
 - Potential subjectivity or lack of inter-rater reliability, rating inflation, favoritism, lack of fidelity to the system, inadequate training and preparation

Challenge #2: Using Observations to Improve Instruction

- Meeting the needs of all teachers – both new and veteran, high- and low-performing
 - Teachers at all levels of performance want feedback about how to improve
- Though everyone *wants* to focus on improvement, worry that PERA and SB 7 might shift focus away from formative aspect and toward accountability

Using Observations to Improve Instruction

- **Strategy: Build Evaluator Capacity**
 - A *lot* of training for evaluators (Growth through Learning for PERA)
 - Multiple modules
 - understanding the observation process and teaching standards and tools
 - distinguishing between performance levels
 - collecting appropriate evidence
 - providing formative feedback
 - Can also be helpful to inform the work of teachers, board members, PD providers, and prep programs as well

Using Observations to Improve Instruction

- **Strategy: Build Evaluator Capacity**
 - Interactions with other evaluators to calibrate ratings and feel more confident in decisions
 - Jointly observing and rating teachers
 - Formal/informal mentoring and discussions
 - Teachscape component of GTL for PERA
 - A learning process for evaluators as well as teachers

Using Observations to Improve Instruction

- **Strategy: Use evaluation ratings to guide professional development**
 - Reinforce cycle of evaluation and professional development
 - Move beyond merely complying with policy and toward using data to help improve instruction
 - Remember that this is not solely about identifying weaknesses or poor performers – can (and should) be used to learn from best practices

“There’s so much talk about evaluation and finding those teachers who shouldn’t be in the classroom, and...I think it’s best used in the reverse. What this does, it identifies the teachers who are most competent, who have the best practice. ...I mean, you have someone in your building you knew was a really good teacher, but what was it about them? What was it about their practice that...possibly others could benefit from? So, now we have that information, and hopefully the district leverages it....[T]he biggest benefit is learning from those who are highly skilled at teaching.” (Teacher)

Using Observations to Improve Instruction

- **Strategy: Collect more data (more observations, multiple measures)**
 - Produces more accurate evaluation ratings
 - Allows evaluators to offer more productive feedback
 - Reduces concerns about the accuracy of ratings and build teacher trust
 - Catch problems early, provide more opportunities to improve and to measure growth

Using Observations to Improve Instruction

- **Strategy: Collect more data**
 - Many districts saw the need, but few districts were able to provide
 - Some suggested possibilities...
 - Frequent, unannounced observations (if evaluators are trained)
 - Re-prioritization/reconceptualization of principal's role as instructional leader
 - Peer evaluation

Comments/Questions about Using Observations to Improve Instruction?

Challenge #3: Reducing the Burden on Principals

- Ultimate impact highly dependent upon principals and their implementation

“[It is] how the administrator proceeds through that is equally as important as the evaluation tool, because if this evaluation tool still becomes just that checklist, which it easily can...then, you know, it’s no different really than anything we’ve done in the past. But if the administrators truly embrace it as an opportunity to provide that growth – and that’s a lot of responsibility on their part because that’s going to take more time on their part.” (Teacher)

Challenge #3:

Reducing the Burden on Principals

- Principals have many additional responsibilities in these new evaluation systems
 - More (and more thorough) classroom observations and conferences than in the past
 - Complex logistics of scheduling
 - Some had training and found it useful, others had trouble adequately pacing
 - Principals held more accountable for performance management and for prioritizing instructional leadership
- Competing priorities and “daily realities” make it difficult to prioritize teacher evaluation

Reducing the Burden on Principals

- **Strategy: Streamline the system wherever possible**
 - Allows evaluators to conduct more observations, be more efficient, and focus on improvement
 - Makes process easier for teachers as well
 - Technology: internet, iPads, evaluation software and apps
 - Elgin: completely paperless – all tools and forms online

Reducing the Burden on Principals

- **Strategy: Streamline the system wherever possible**
 - User-friendly references to simplify and summarize complex documents and procedures
 - Sandoval: Implementation toolkits and guidebooks
 - Olympia: Customized calendars to outline schedules for the year and make sure everyone sticks to deadlines

“[The online forms were] one of the pieces that made the whole program successful. It isn’t so much about the values and beliefs of the [teacher evaluation] program—which I think are essential in this document—but these are enhancing pieces that allow user ease. And when you make something easy to use, it becomes less threatening and you don’t have that undertow, and then the document can take over and you can begin doing the good work.” (Elgin)

Reducing the Burden on Principals

- **Strategy: Use Multiple Observers**
 - To reduce burden on principals; *and*
 - Increase the number of classroom observations; *and*
 - Potentially add support for new teachers and struggling teachers
 - Role for other administrators and department chairs or other teacher leaders

Reducing the Burden on Principals

- **Strategy: Use Multiple Observers**
 - Peer Assistance and Review (PAR) in Niles:
 - Two full-time peer “consulting teachers” (CTs) working with new teachers in 2011-12 for evaluation and support (along with mentors)
 - Expanded to four CTs this year, working with some veterans who need improvement
 - Selected well-respected teachers as CTs
 - Eight to twelve observations each year
 - Experience is that this system is able to provide more support *and* hold teachers more accountable

“People [who were released through PAR] had every single possible support that we could’ve given them in place, and at the end, they still either were not making the progress at the rate that we expected them to be...and so I feel more comfortable that that’s happened. Like, I feel more comfortable releasing the people after a year of, like, having all this stuff in place”
(Niles)

Comments/Questions about Reducing the Burden on Principals?

Challenge #4: Incorporating Student Growth in to Teacher Evaluation

- Only Evanston had done this; reluctance among others to be the first and “be the ones inventing the whole wheel”
 - Especially while trying to get Common Core curricula and PARCC assessments established first
- Numerous concerns and unanswered questions:
 - Buy-in to using student growth and concerns about attribution
 - Fairness and rigor across content areas and student populations
 - Finding valid and reliable assessments for “non-tested” subjects
 - Small sample sizes
 - Relationship between student growth and observation measures
 - Understanding assessments, growth models, and PERA requirements

Incorporating Student Growth

- Accept that it's coming, focus on strategies and supports to help get it right:
 - Multiple measures (PERA requirements)
 - Training on how to understand assessment data and use for improvement
 - Making sure that everyone understands the plan
- Niles: well-positioned to incorporate growth
 - Have been using student growth measures for program improvement purposes for a while
 - Teachers and administrators are familiar and comfortable with these measures
 - Have had time to establish local norms

“[It] gives us our own local data and help[s] us make decisions on how students are achieving, under which teachers. ...We’re dealing with our teachers, in our schools, in our situations, and what would it be and how would it be, for example, if the child were in a different school, with a different teacher.” (Niles)

Comments/Questions about Student Growth or Other Topics Not Yet Addressed?

Chicago Pilot 2008-2010

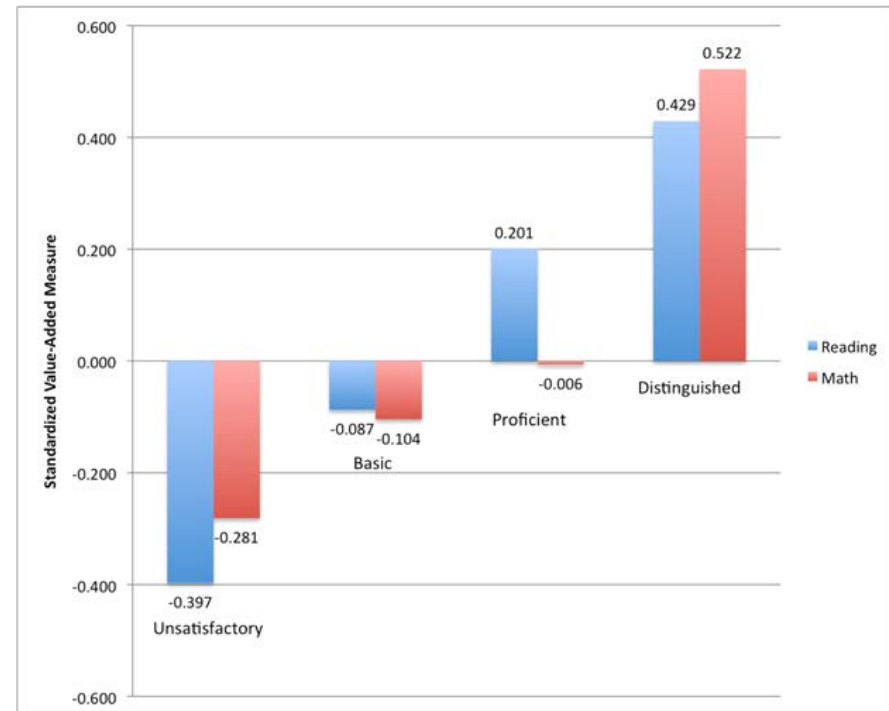
- Chicago piloted Danielson rubric for evaluating teachers in 44 K-8 schools in 2008, expanded to 100 schools in 2009.
- Consortium on Chicago School Research studied:
 - Reliability
 - Validity
 - Implementation

Reliability

- Compared ratings of principals with ratings of 3 full time external observers
- Found that principals vary in the precision and severity of their ratings

Validity

- Compared ratings on Danielson rubric (component by component) with teacher-level value-added on ISAT
- Found that teachers with higher observation ratings tended to have higher value-added scores



Implementation Logistics

- Case studies in 8 schools
- Early design decisions: Before actual implementation could start, some key decisions necessary
 - All observations were scheduled
 - Multiple observations and external observers required
 - Web-based ratings database created to ease principals' reporting burden and assist in district data collection

Implementation Challenges

- Findings from 2008 study of CPS similar to case study districts discussed today:
 - Limited teacher understanding was a barrier to trust in the system—need centralized communication strategy
 - Teacher attitudes often dependent on principal skill and buy-in
 - Principals often struggled with logistics, time management
 - Principals wanted more support in their ‘new’ role as instructional coach

Regardless, Most Participants Positive

- **89% of principals in pilot agreed that:**
 - Quality of conversations with teachers had improved
- **Teachers agreed:**
 - Conferences led to conversations about instruction that were more reflective, based on shared language, evidence-based
- **November 2011 report on website**
 - <http://ccsr/uchicago.edu/publications>

Questions about CPS studies?

Discussion

Reforming teacher evaluation is a work in progress and complicated problems remain, including:

- Communication and teacher understanding
 - How can boards, district administrators, and teachers unions augment principals' communications with teachers?
- Principals' role with limited time: gatekeeper, coach
 - How can the observation process be streamlined?
 - What can be done about resistant principals?
 - How can boards/districts/principal prep programs help principals develop new skills?

Discussion -2-

- Possible lack of ratings consistency
 - What processes can districts put in place to build trust by ensuring consistency across buildings and over time?
 - How can time be stretched to allow maximum number of observations to increase precision, accuracy and trust?
- Differentiated supports for teachers at all levels
 - How can districts encourage the use of the system for all performance levels instead of appearing to focus solely on weakness?

Discussion -3-

- Incorporating student growth
 - What role can the state play in helping districts develop high quality assessments?
 - How can board members and educators learn enough to be able to answer teacher and parent concerns about fairness and accuracy?
- Some good resources:
 - Carnegie Foundation webinars (www.carnegieknowledge.org)
 - CALDER Conversations (www.caldercenter.org)
 - Doug Harris' *Value-Added Measures in Education: What every educator needs to know*

Going Forward

- These early districts informed by external organizations and out-of-state districts – next generation should be able to rely more on each other's experiences
- **Make time to talk with each other**

Final Comments/ Questions?

Contact Info:

Brad White, Senior Researcher
Illinois Education Research Council
brawhit@siue.edu
<http://ierc.siue.edu>

Sue Spote, Director of Research Operations
The University of Chicago Consortium on Chicago School Research
ssporte@ccsr.uchicago.edu
<http://ccsr.uchicago.edu>

Download the full report online:

http://www.siue.edu/ierc/pdf/2012_Teacher_Evaluation_Policy_Brief1.pdf

-or-

http://ccsr.uchicago.edu/sites/default/files/publications/Teacher%20Evaluation%20Policy%20Brief1_0.pdf