



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

Brad White, Illinois Education Research Council

Jen Cowhy, Consortium on Chicago School Research

Sue Sporte, Consortium on Chicago School Research

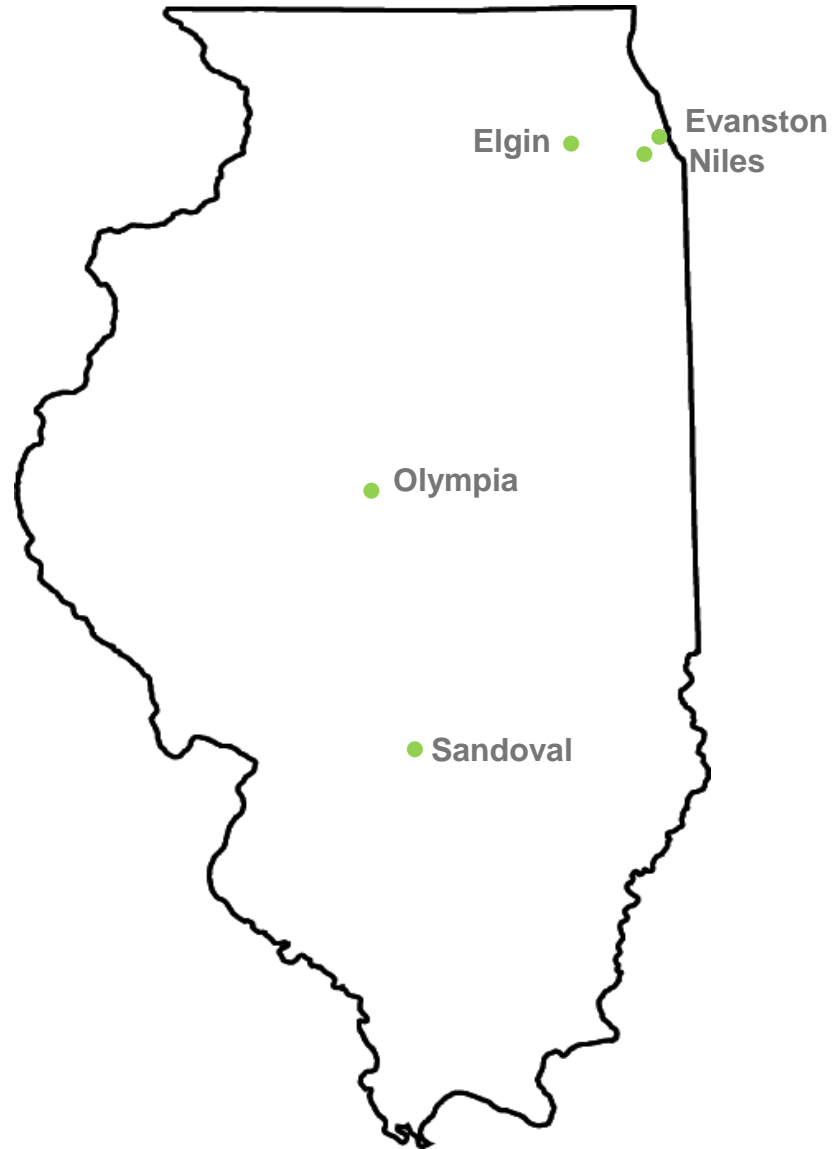
# Purpose of Study

- 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 “environmental 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) and locales (1 urban, 2 suburban, 2 rural)
- Range of student populations (30%-70% FRL and 30%-95% white)

# Evaluation System Descriptors

	Elgin	Evanston	Niles	Olympia	Sandoval
<b>Locale</b>	Urban	Suburban	Suburban	Rural	Rural
<b>Implementation Status-Observations</b>	Fully implemented	Early implementation	Piloting	Fully implemented	Designing
<b>Teacher Performance</b>	YES	YES	YES	YES	YES
<b>Student Growth</b>	NO	YES	YES (piloting 2012)	NO	Designing 2012
<b>Evaluation Tool</b>	Modified Danielson	Modified Danielson	Modified Danielson	Modified Danielson	Modified Danielson
<b># of Formal Evaluations for New Teachers</b>	3 per year	2	8 to 12	1	3 observations and 9 meetings
<b># of Formal Observations for Tenured Teachers</b>	1-2 every other year	1	2	1	1 obs 1 mtg every other year
<b>Who Observes?</b>	Principals and Admin	Principal and/or Outside Evaluator	Peers and Admin	Principals/ APs	Principal
<b>HR link</b>	Remediation Plans	Salary, PD	PD, renewal	Hiring, renewal	PD, tenure

©CCSR

# *Expected Opportunities? Challenges?*

# Overview

- **These systems were seen as a huge upgrade over the status quo**
- **But there were still some challenges...**
  1. Securing buy-in and understanding
  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 Buy-In & Understanding

- Buy-in problematic, especially in the first year
- Training typically did not include teachers
- Distrust between teachers and administration; anxiety surrounding PERA
- Success dependent upon principal understanding and buy-in

# Cultivating Buy-In & Understanding:

- **Strategy: Gather all perspectives**
  - Balanced representation and diverse perspectives on design committee important
  - Facilitators, consultants, and other districts useful
  - Ongoing monitoring necessary
    - Integrate input to revise and modify policy as needed

# Cultivating Buy-In & Understanding:

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

# Cultivating Buy-In & Understanding:

- **Strategy: Train teachers early, continuously, and consistently**
  - Transparency in design process
  - Ongoing training and communication
    - Consider consistent training with teacher-administrator teams
  - Pitfalls of voluntary/optional training sessions

# Cultivating Buy-In & Understanding:

- **Strategy: Align teacher evaluation with other district priorities**
  - Link to existing human resources systems (hiring, mentoring, etc.) or other initiatives
  - Can serve as the “glue that holds the rest together” and reinforces other initiatives

*“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 Buy-In & Understanding:

- **Strategy: Start soon and implement gradually**
  - Consider phasing in with low-stakes pilot
  - Don't start or change mid-year
  - Start planning/designing ASAP

# Questions?



# 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
- 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 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 and toward accountability

# Using Observations to Improve Instruction

- **Strategy: Build Evaluator Capacity**
  - A *lot* of training for evaluators(12-35 hours)
    - Provided by CEC, the Danielson Group, independent consultants, hand-picked trainers from within the district
    - 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

# 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 (Niles, Sandoval, Elgin)
    - Formal/informal mentoring and discussions (Olympia)
    - Growth through Learning/Teachscape for PERA

# Using Observations to Improve Instruction

- **Strategy: Link observations to professional development**
  - Reinforce the view of evaluation and professional development as a “cycle”
  - Move beyond merely tracking compliance and toward using data to help improve instruction
  - Use evaluation ratings (along with other data) to direct district-level professional development
  - Don't just merely identify weaknesses and poor performers, 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: Conduct more classroom observations**
  - Reduce concerns about the accuracy of ratings and build teacher trust
  - Offer more productive feedback
  - Catch problems early, provide more opportunities to improve

# Using Observations to Improve Instruction

- **Strategy: Conduct more classroom observations**
  - Many 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



# Questions?

# 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
  - 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*
  - Add support for new/struggling teachers
  - Some districts used assistant principals or other evaluators to help lighten the load

# 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)*

# Questions?

# 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
  - Communicating your 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)*

# Questions?

# Topics still unaddressed?

# Chicago Pilot 2008-2010

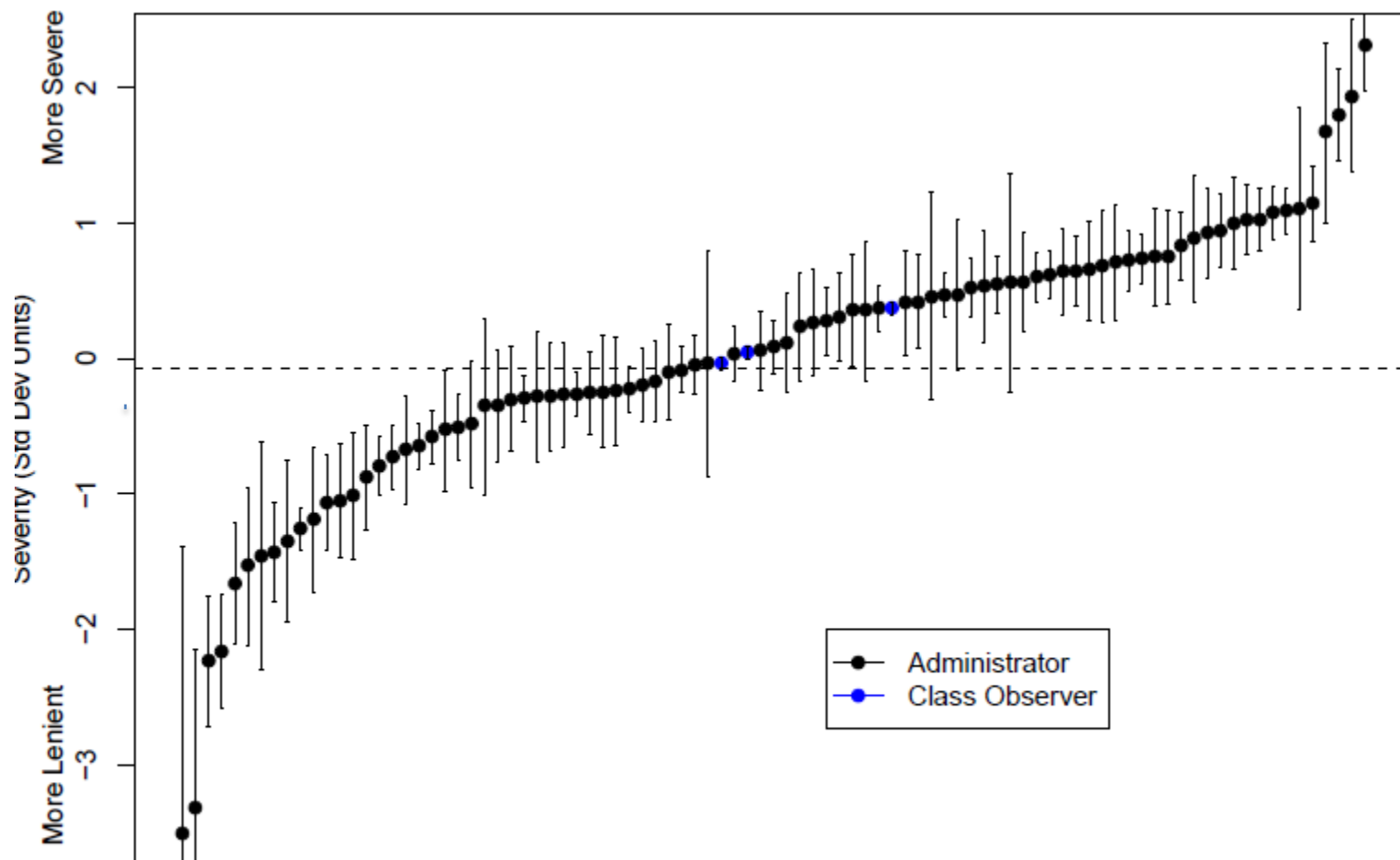
- Chicago piloted Danielson rubric for evaluating teachers in 44 K-8 schools in 2008, expanded to 100 schools in 2009.
- Interested in:
  - Establishing reliability
  - Establishing validity
  - Studying implementation



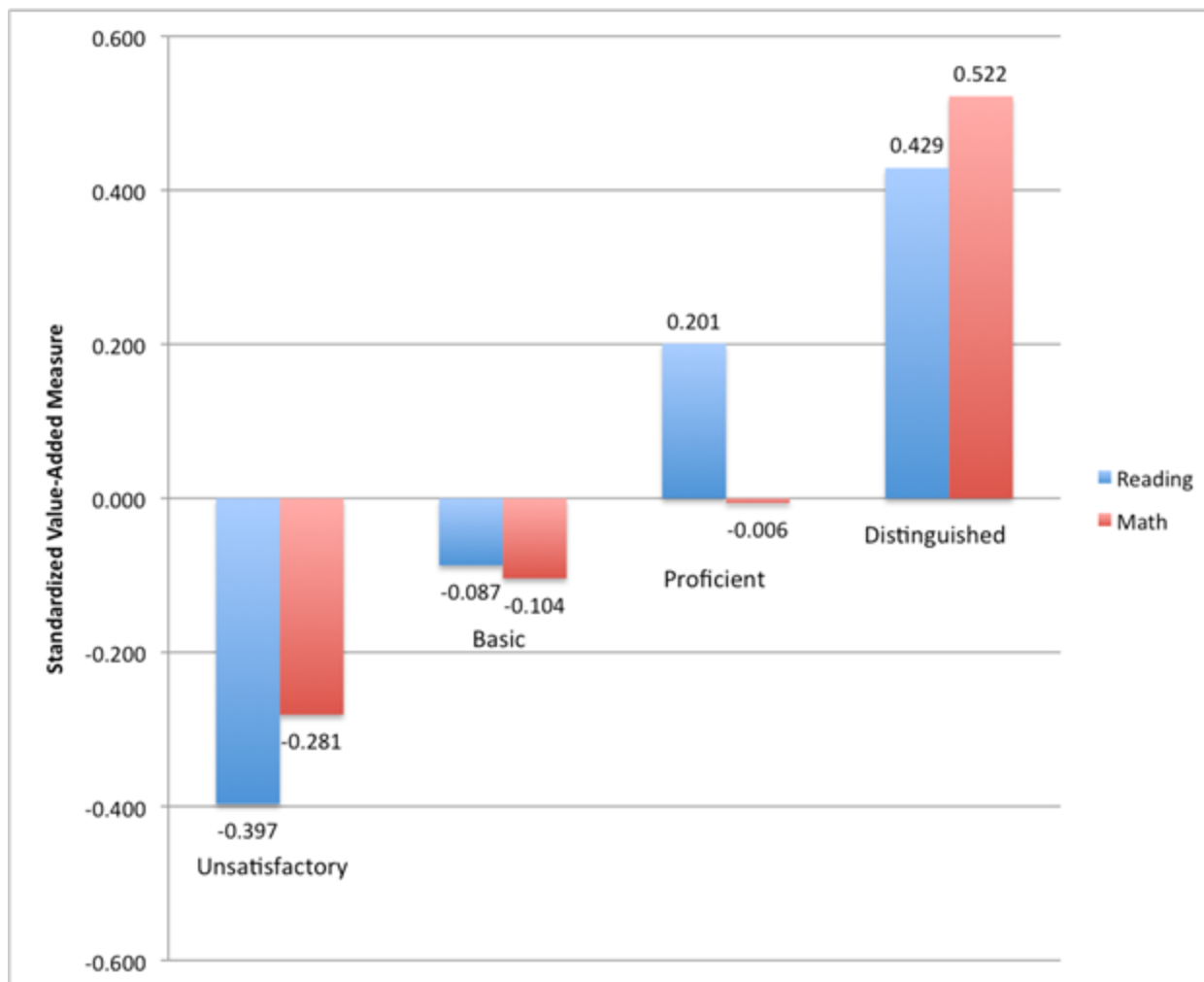
# Reliability/Validity/Implementation

- Establishing reliability
  - Compared ratings of principals with ratings of 3 full time external observers
- Establishing validity
  - Compared ratings on Danielson rubric with teacher level value-add on ISAT, component by component
- Studying implementation
  - Logistics
  - Case studies in 8 schools

# Reliability: Principals vary in the precision and severity of their ratings



# Validity: Higher observation ratings related to higher value-added



# Implementation: Early Design Decisions: Observation-logistics

- 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

# District Implementation Challenges- Similar to Other Districts in 2012

- **Case studies from 2012 study NOT Chicago, but many issues same as CPS in 2008**
  - Lack of teacher knowledge a barrier to trust in the system—need centralized communication
  - Teacher positive 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?

# Implications

**Reforming teacher evaluation in all 5 districts is a work in progress; complicated problems remain**

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



## Implications -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 and experience levels instead of and focus only on weakness?

# Implications -3-

- Incorporating student growth
  - What role can the state play in helping districts develop high quality assessments?
  - How can district and school leaders learn enough to be able to answer teacher and parent concerns about fairness and accuracy?
- Some resources:
  - <http://www.carnegieknowledgenetwork.org/>

# 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 Questions? Comments?

# 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>*

*Jen Cowhy, Research Assistant  
The University of Chicago Consortium on Chicago School Research  
jcowhy@uchicago.edu*