

Teacher Evaluation in Chicago: Findings from the Excellence in Teaching Project

Lauren Sartain

PEAC Meeting – 15 April 2011

THE CONSORTIUM ON CHICAGO SCHOOL RESEARCH
AT THE UNIVERSITY OF CHICAGO / URBAN EDUCATION INSTITUTE

The Chicago Context: 2006 to 2010

- 2006-08: CPS-CTU Joint Committee on Teacher Evaluation
- 2008-09 SY: Year 1, the Excellence in Teaching Pilot
 - 44 CPS elementary schools randomly selected for the pilot
- 2009-10 SY: Year 2, expansion of the pilot
 - 100 CPS elementary schools implemented



Excellence in Teaching Project

- Implemented in 100 elementary schools
- Extensive training for principals
- Evidence-based conferences
- Observations with the Danielson Framework



CCSR Study: Research Questions

- What are principal and teacher perceptions of the Framework?
- What is the relationship between school context and fidelity of implementation?
- Is the Framework a reliable tool? Is it valid?



Main Findings

- Most principals were engaged – this matters for implementation
- Principals generally rated teachers consistently
- Teachers with low ratings had low value-added measures (and vice-versa)
- Ratings differentiate among teachers



Engaging Principals in Teacher Evaluation



Most Principals Are Engaged

Level of Engagement	Administrators Year 1 (n=39)	Administrators Year 2 (n=99)
High Enthusiasm	57%	52%
Mixed Emotions	28%	31%
Resistant	15%	17%



Engagement Related to Perceptions of Influence on Teaching Practice

Level of Engagement	Perceived Impact on Practice
High Enthusiasm	Impact on instructional grouping, assessment of student learning, planning
Mixed Emotions	Impact on teacher accountability, compliance
Resistant	No impact



Principals Need to Develop Instructional Leadership Skills

Level of questioning	Percent of questions	Sample questions
High	10%	What is the relationship between student engagement and classroom management in your teaching? What are some concrete steps you can take to improve each?
Middle	25%	How did you fulfill the goals you set for this lesson? Which goals did you not meet?
Low	65%	This component is for this lesson is a "Basic." Do you agree with that?



Conversations are one-way: 75% of talking and questioning done by principals

Walton School: Low Engagement, Low Institutionalization

"Nothing has changed for me. It is just more work."
–Principal

The principal "did not know enough about instruction to push for improvement" and is "overwhelmed by the Danielson Framework" –Teacher

The tool has "powerful potential" but limited use "made it just as subjective as the old tool" –Teacher



Stoller School: High Engagement, Deep Institutionalization

“If you’re saying to me that you’re a life-long learner, you’re reflective, you want to grow in this profession...evaluation is not what we do. It’s who we are.” –Principal

“It’s part of our daily conversation” –Teacher

“Regardless of whether or not CPS adopts it, she’s made it hers; she’ll stick with it” –Teacher



Many Factors Affect School Implementation

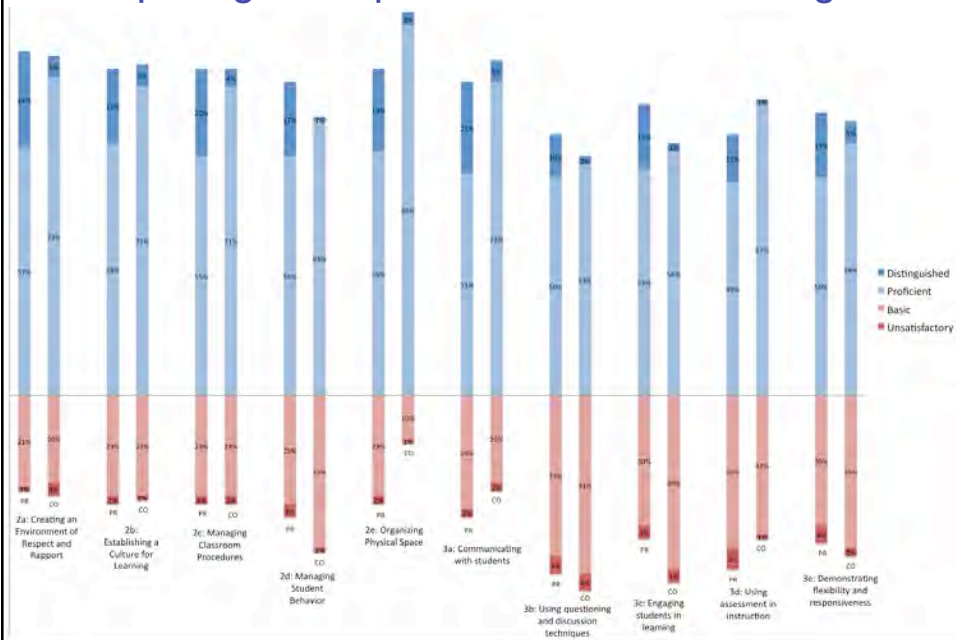
- Context matters and implementation varies
- Principal leadership defines tone and depth
- Shift from previous conception of teacher evaluation takes work
- Translating a standards-based tool into improved instructional practice is not a given



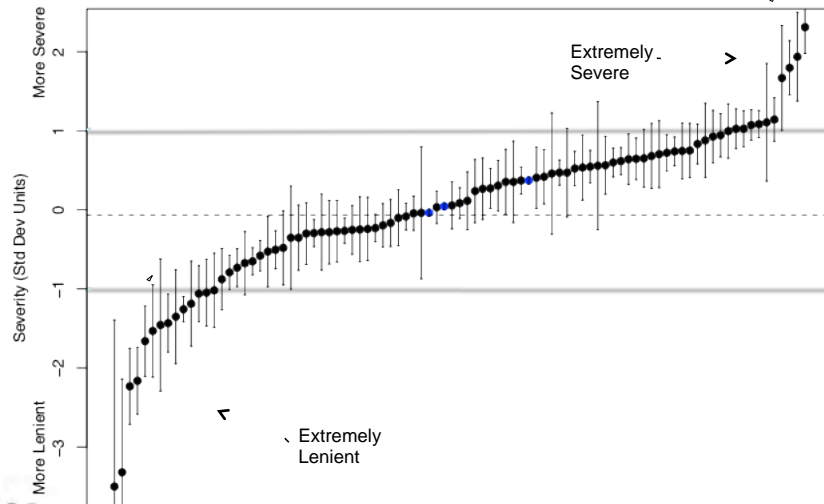
Using the Observation Tool in a Reliable and Valid Way



Comparing Principal and Observer Ratings



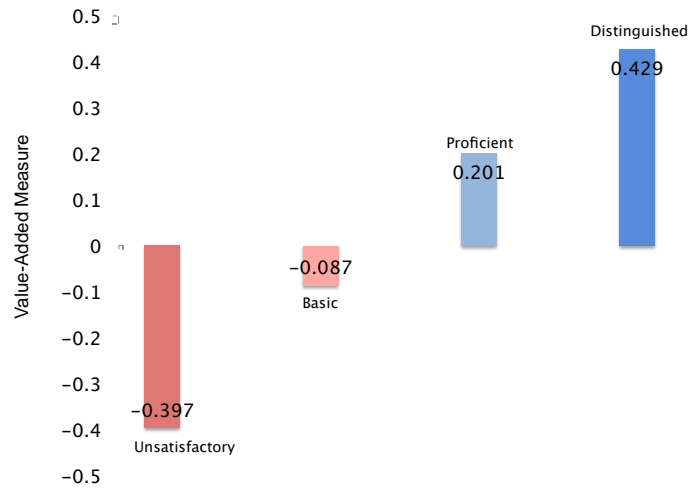
Most Principals Rate Consistently



How We Measure Validity

- What's the relationship between teacher practice and student learning?
 - Teacher practice measure: classroom observation ratings
 - Student learning measure: CPS teacher-level value add
- What we expect from a valid observation tool
 - Low observation ratings → low value add
 - High observation ratings → high value add

Flexibility & Responsiveness (3E) – Reading VA



CCSR

Framework Implemented Reliably and Validly

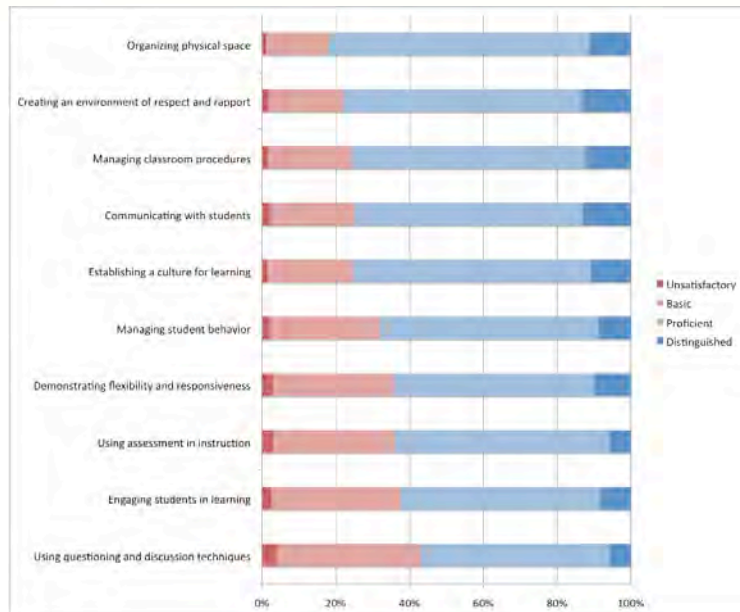
- Overall principals use the low and middle parts of the scale reliably
- Strong relationship between Framework ratings and value-added measures
- Validity and reliability is a function of how the tool is implemented
- Districts need to build in accountability to ensure high school-level implementation

CCSR

What the Ratings Tell Us About Teaching Practice



Teachers Weaker in Instruction Than Environment



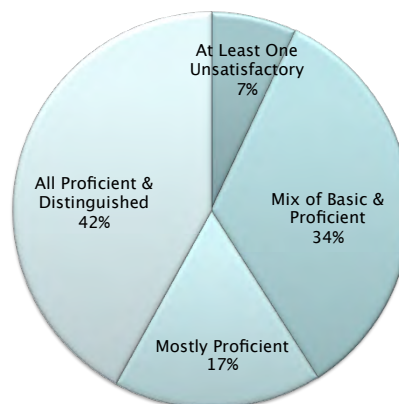
Two Sets of Criteria for Summative Ratings

- | <u>Danielson recommended</u> | <u>One urban district</u> |
|---|---|
| <ul style="list-style-type: none">• Four ratings categories• <i>Lowest category</i>: at least 1 unsatisfactory• <i>Highest category</i>: all proficient and distinguished | <ul style="list-style-type: none">• Three ratings categories• <i>Lowest category</i>: at least 1 unsatisfactory OR > 3 basic• <i>Highest category</i>: all proficient with 1 distinguished per domain |



We applied both sets to the same CPS teachers

Applying Danielson-Recommended Criteria



Applying One Urban District's Criteria



CCSR

Designing a Rating System Is Difficult

- Where you set the bar can affect a teacher's final rating
- Getting the full picture of teaching is important
 - Observations are only one, albeit important, aspect
 - Multiple measures
- Keeping sight on instructional improvement
 - A single rating does not provide teachers with information on their practice

CCSR