



# **Illinois State Board of Education**

**June 10, 2015**

**Guidance Document 15-4**

**PEAC Guidance to Schools and Districts**

**Regarding Evaluating Teacher Practice and  
Understanding Summative Ratings**

*This document is intended to provide non-regulatory guidance on the subject matter listed above.  
For specific questions, please contact the person(s) identified in the document.*

---

Tony Smith, Ph.D., State Superintendent



## PEAC Guidance to Schools and Districts

---

### Guidance Regarding Evaluating Teacher Practice and Understanding Summative Ratings

There are concerns that some teachers are being placed in the “Unsatisfactory” and “Needs Improvement” categories based on an “Unsatisfactory” or “Needs Improvement” rating in as little as one component of one domain used to assess a teacher’s practice.

As this is not a practice recommended by the Performance Evaluation Advisory Council (PEAC), the PEAC is providing the following guidance for school districts to consider as they develop and implement their evaluation systems.

This guidance focuses on the evaluation of teacher practice as it relates to a teacher’s summative rating.

### Summative Ratings

The Performance Evaluation Reform Act (PERA) requires that by no later than the 2012–13 school year, school districts must use a four-category rating system for both principal/assistant principal and teacher evaluations (“Excellent,” “Proficient,” “Needs Improvement,” and “Unsatisfactory”). PERA and the PERA Administrative Rules require that the plan consider the professional practice of the teachers and, by a district’s statutory implementation date, data and indicators of student growth.

### Instructional Framework for Use in the Evaluation of Teacher Practice

School districts must identify an instructional framework for use in the evaluation of teacher practice “that is based on research regarding effective instruction; addresses at least planning, instructional delivery, and classroom management; and aligns to the Illinois Professional Teaching Standards” (PERA Administrative Rules, Section 50.120(a)). “The instructional framework shall align to the roles and responsibilities of each teacher who is being evaluated” (PERA Administrative Rules, Section 50.120(a) (1)). In addition, “the evaluation plan shall contain a rubric to be used in rating professional practice that aligns to the instructional framework” being used (PERA Administrative Rules, Section 50.120(a) (2)).

In fact, most school districts (including many through their PERA joint committees or teachers union) have elected to use one of several instructional frameworks that are being employed in districts across the country (i.e., Danielson’s Framework for Teaching, Marzano’s Teacher Evaluation Model, or other research-based frameworks). Those frameworks define a set of teaching practice domains and provide a rubric that describes multiple levels of performance for the components being evaluated.

PERA required that evaluators be prequalified by September 1, 2012 (Illinois School Code 105 ILCS 5/24A-3). During the Illinois Growth Through Learning prequalification training, evaluators were trained to collect appropriate evidence, align evidence to a framework, and make professional judgments regarding teaching performance levels based on the evidence collected. Evaluators were expected to demonstrate interrater reliability as a part of this evaluation process, a key part of district considerations regarding the rating of teaching practice.

## Weighting the Components of Teacher Practice

Neither PERA nor its Administrative Rules define the relative weights of the components of teacher practice or how ratings of components of teacher practice are to be combined into a final rating of teacher practice. School districts are charged with making those decisions. Again, many districts are doing this through their PERA joint committees or teachers union.

PEAC recommends that the following questions be considered as district evaluation systems are developed:

1. *Should we assign a weight to the domains and/or components of the instructional framework?*

A discussion of the relative weight given to each domain of the instructional framework will require considering the relative importance of each part of the teaching framework. For example, a school district using the Danielson Framework could decide to weight Domain 1 (Planning and Preparation) at 20 percent, Domain 2 (Classroom Environment) at 20 percent, Domain 3 (Instruction) at 40 percent, and Domain 4 (Professional Responsibilities) at 20 percent. This discussion forces addressing the following key questions: How important is planning? Is a teacher's participation in the professional community as important as the classroom environment? Do the domains of the district's framework work together to inform and support each other? It should be determined if weighting domains or components reflect the values and beliefs of the district and an understanding of the district's framework as well as the goals and purposes of the evaluation system.

It may be decided to weight each component equally (i.e., not assign any weight to any of the domains or components). Again, it is important to ensure that this decision is a reflection of the values and beliefs of the district as well as the goals and purposes of the evaluation system. Many districts work closely with their joint committees or teachers union to make these decisions.

2. *What is the minimum level of performance required for each level of practice?*

A teacher's summative practice rating should combine the evaluator's assessment of evidence of teaching performance in each domain or component of the instructional framework. This determination can be made based on a set of decision rules that define the level of performance required for each summative rating using the component ratings. For example, it could be decided that a teacher who is proficient in all of the components of the instructional domain cannot receive an "Unsatisfactory" rating. As another example of a decision rule, a teacher who has "Unsatisfactory" or "Needs Improvement" ratings on three or more instructional components may not receive a "Proficient" practice

rating. In addition to using a set of decision rules, there are other ways to determine a summative practice rating, including the use of numerical scoring, examining the preponderance of evidence across components/domains, or holistic scoring. Regardless of the method used to determine summative practice ratings, it is important to consider how different potential rating outcomes reflect the district’s values and beliefs as well as the teaching strengths and weaknesses across components/domains.

3. *Should a rating on a single component of the framework determine the overall rating of practice?*

School district evaluation plans should specifically address whether an “Unsatisfactory” or “Needs Improvement” rating on one component of one domain should automatically result in an overall practice rating of “Unsatisfactory” or “Needs Improvement.” Given the structure of most of the practice frameworks in use in districts, basing the overall practice rating on one component is bad practice and inconsistent with the goal of improving teaching practice; therefore, it should not be utilized. Such a choice is not likely to result in constructive feedback and thoroughly eliminates the usefulness of the evaluation system for identifying areas in need of improvement. The components of most practice frameworks are interconnected, and evidence collected across components can provide a fuller picture of teaching practice, which can foster productive conversations between teacher and evaluator. Thus, summative practice rating methods should give consideration to the prevalence of strengths and weaknesses found across domains and components through the evidence collected during multiple observations.

## **Varying Approaches for Determining Overall Practice Ratings for Educators**

*Not endorsed or recommended—these are simply examples.*

### **Example 1: Simple Average**

**Urbana School District, Illinois**

Source: *Supportive Supervision Evaluation Plan Handbook* (May 9, 2013)

[http://www.usd116.org/files/2013BoardDocs/2013-5-21-items9\\_0.pdf](http://www.usd116.org/files/2013BoardDocs/2013-5-21-items9_0.pdf)

- Urbana’s Supportive Supervision Evaluation Plan uses the Framework for Teaching to assess teacher practice.
- There are four domains within the Framework for Teaching:
  1. Planning and Preparation
  2. Classroom Environment
  3. Instruction
  4. Professional Responsibilities
- The overall performance evaluation rating is calculated from the average of four domain scores, with the final rating assigned by using a set of score-level guidelines.

## How It Works:

- Educators receive a performance rating from 1 to 4 in each of the subcomponents within each of the four domains: 1=Unsatisfactory, 2=Needs Improvement, 3=Proficient, and 4=Excellent.
- The score for each of these four domains is calculated by averaging the scores (1–4 point scale) for each of the *subcomponents* under the domain (see the table below).
- The overall score is calculated by averaging the four domain scores.
- The final overall performance evaluation rating is determined by assigning the level associated with the *score-level guidelines* below<sup>1</sup>:
  - Unsatisfactory: score of 1.0–1.99
  - Needs Improvement: score of 2.0–2.74
  - Proficient: score of 2.75–3.49
  - Excellent: score of 3.5–4.0

### Example for Calculating the Performance Evaluation Rating

Sample calculation for Domain 1 (D1)

Domain 1	Unsatisfactory	Needs Improvement	Proficient	Excellent
Ia	1	2	3	4
Ib	1	2	3	4
Ic	1	2	3	4
Id	1	2	3	4
Ie	1	2	3	4
If	1	2	3	4

$(1a+1b+1c+1d+1e+1f)/6$ = Performance Evaluation Rating D1

$(4+4+4+3+3+3)/6$ = Performance Evaluation Rating D1

3.5= Performance Evaluation Rating D1 The resulting rating for D1 would be Excellent.

The teacher would receive a 4.0 on Domain 1 for purposes of calculating the cumulative rating.

### Example—Teacher 1: Calculating the Simple Average Overall Performance Rating

Teacher 1 received the following four domain scores:

	Unsatisfactory	Needs Improvement	Proficient	Excellent
Domain 1		2		
Domain 2		2		
Domain 3		2		
Domain 4	1			

$(D1+D2+D3+D4)/4$ =Performance evaluation rating

$(2+2+2+1)/4= 1.75$

Score falls in 1.0–1.99 range=Unsatisfactory

Final overall performance evaluation rating is Unsatisfactory.

<sup>1</sup> **Important Note:** In Urbana, if a tenured teacher is rated “Unsatisfactory” in any of the four rating domains, the teacher will be placed on a professional development plan at a minimum.

**Example—Teacher 2: Calculating the Simple Average Overall Performance Rating**

Teacher 2 received the following four domain scores:

	Unsatisfactory	Needs Improvement	Proficient	Excellent
Domain 1			3	
Domain 2			3	
Domain 3			3	
Domain 4	1			

$(D1+D2+D3+D4)/4$ =Performance evaluation rating  
 $(3+3+3+1)/4= 2.5$   
 Score falls in 2.0–2.74 range=Needs Improvement  
 Final overall performance evaluation rating is Needs Improvement.

**Note:** This example only explains the professional practice score. Additional information is available on how an overall summative rating is calculated with a combination of professional practice scores and measures of student growth.

**Example 2: Weighted Average**

**Chicago Public Schools (CPS)**

Source: *CPS REACH Students, Educator Evaluation Handbook, 2014–15*  
<http://www.ctunet.com/rights-at-work/teacher-evaluation/text/CPS-REACH-Educator-Evaluation-Handbook-FINAL.pdf>

- CPS Framework for Teaching is adapted from the Danielson Framework for Teaching.
- There are four domains within the Framework for Teaching:
  1. Planning and Preparation
  2. Classroom Environment
  3. Instruction
  4. Professional Responsibilities
- Levels of performance on the CPS Framework are Distinguished, Proficient, Basic, and Unsatisfactory (1–4).
- The overall performance evaluation rating is calculated using a weighted average of the four domain scores.

**How It Works:**

- There are four domains within the CPS Framework for Teaching. Each domain is assigned a separate weight when calculating a teacher’s professional practice score:
  1. Planning and Preparation (25 percent)

2. Classroom Environment (25 percent)
  3. Instruction (40 percent)
  4. Professional Responsibilities (10 percent)
- Each domain contains four or five subcomponents. Educators receive ratings of 1 to 4 at the component level following formal and informal observations.
  - Subcomponent ratings from four observations are averaged over the course of the evaluation cycle.
  - The averaged subcomponent scores within each domain are then averaged to yield a domain average.
  - The domain average is multiplied by the weight for that domain.
  - The weighted domain averages are added together to produce the *total professional practice score*.

**Note:** This example only explains the professional practice score. Additional information is available on how an overall summative rating is calculated with a combination of professional practice scores and measures of student growth.

**Example:**

Domain	Sample Domain Score (average of all scores given for the subcomponents in all observations)	Domain Weight	Overall Score
Planning and Preparation	3.25	25%	$3.25 \times .25 = 0.81$
Classroom Environment	3.25	25%	$3.25 \times .25 = 0.81$
Instruction	3.3	40%	$3.3 \times .4 = 1.32$
Professional Responsibilities	3.9	10%	$3.9 \times .10 = 0.30$
<b>REACH students: Professional practice score</b>			<b>3.33</b> (sum of all weighted domains)