

Rater Agreement in Washington State’s Evaluation System

Rater Agreement Background

The new Washington State Evaluation System for teachers and principals will require new learning and ongoing support to increase the effectiveness of implementation.

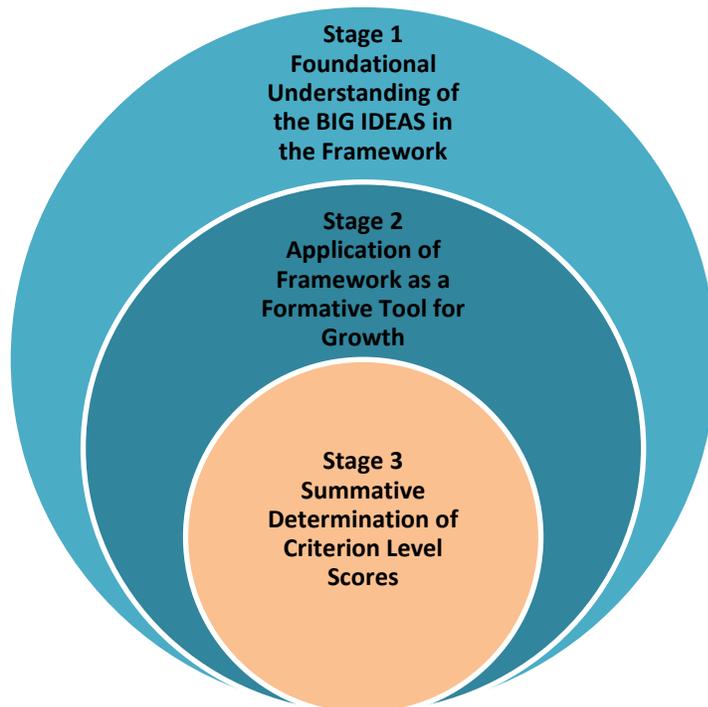
The Teacher and Principal Evaluation Project (TPEP) has relied heavily on the growing body of research, the framework authors, and the practical input from practitioners in the pilot sites to create a “working definition” of rater agreement for the 2012 – 13 school year.

TPEP is guided by ESSB 5895, which passed this last spring and became law on June 7, 2012. The new law requires that evaluators of both teachers and principals **“must engage in professional development designed to implement the revised systems and maximize rater agreement”** (RCW 28A.405.120 and RCW 28A.405.130).

The research and practical application of implementing this law is of primary concern for the Office of Superintendent of Public Instruction (OSPI), TPEP Steering Committee organizations, and the instructional and leadership framework authors.

In partnership with the three instructional framework authors and the Association of Washington School Principals (AWSP), OSPI will use the following working definition and steps to guide our work *this year* with the RIG districts:

Rater Agreement Definition and Stages



Rater Agreement

The extent to which the scores between the raters have consistency and accuracy against predetermined standards. The predetermined standards are the instructional and leadership frameworks and rubrics that define the basis for summative criterion level scores.

Stage 1	
Evaluators should know and be able to:	<i>Understand</i> the Big Ideas of the Instructional or Leadership Frameworks and the inter-dependency of the frameworks, rubrics, and state criteria. Accuracy in scoring depends on this foundational level. This foundational level of understanding is the key to the future calibration and application. <i>It is recommended that this training occur prior to evaluating teachers or principals.</i>
Success Indicators:	Understanding of: <ul style="list-style-type: none"> • The appropriate use of the instructional language and framework vocabulary • The alignment of framework to state criteria • The interdependence of Dimensions/Indicators (CEL), Domains/Components (Danielson & Marzano) • The key ideas in each framework and what evidence would look like/sound like • The definition of essential aspects of the framework • The possible evidence for aspects of the framework
Districts should be creating plans to:	Provide the Stage I training to all evaluators before evaluating staff.
Stage I Training	
<p>OSPI, through the services of the Instructional Criteria Framework Feedback Specialists, will provide two-day Stage 1 training for all who evaluate classroom teachers. This two-day overview provides an understanding of the “Big Ideas” of the Instructional or Leadership Frameworks and the inter-dependency of the frameworks, rubrics and state criteria (dates pending for principal evaluator training).</p> <p>RIG 1 Districts – August 2012 (Completed)</p> <p>RIG 2 Districts (fall) – February 2013 (Specific frameworks will be decided for each region by November and registration will be available on the OSPI events manager).</p>	
Spokane	Feb 5-6
ESD 105	Feb 4-5
ESD 112	Feb 5-6
Lacey (Olympia)	Feb 6-7
ESD 114	Feb 5-6
ESD 121	Feb 6-7
ESD 123	Feb12-13
ESD 171	Feb 6-7
ESD 189	Feb 12-13

RIG 2 Districts (spring) – June 2013**All others - August 2013 (pending legislative funding)****Stage 2**
(subject to funds available)

Evaluators should know and be able to: Apply the framework and rubric in a *formative* process. This includes the capacity to give feedback, provide immediate support for implementing the feedback, and provide long-term professional development. Real life observation must be included in the application of this stage. ***Optimal training for this stage happens during the year-long evaluation cycle.***

Success Indicators: Formative Application of:

- The appropriate use of language of instruction and framework vocabulary
- The interdependence of dimensions/domains and indicators/components
- The key ideas in each criteria/indicators/components and what evidence would look like/sound like
- The essential aspects of the framework
- The evidence for all aspects of the framework
- Providing formative feedback based on evidence
- Development and implementation of short-term professional development plans

Districts should be creating plans to: Move evaluators to accuracy and consistency through ongoing calibration conversations involving real-life or video-based observation. Consistency among raters is possible at this level.

Stage II Training

OSPI, through the services of the Instructional Criteria Framework Feedback Specialists, will provide up to 30 hours* of ongoing training on the frameworks for all who evaluate classroom teachers, principals, and assistant principals.

***The RIG 1 districts are currently piloting the Stage 2 training with principals and district administrators. A final determination of 2013 – 14 Stage 2 training will be made during the 2012 – 13 school year and be based, in part, on legislative budget approval. Stage 2 training is also dependent on the amount of growth needed by the evaluators. Since this is a performance-based system, a principal's evaluation should be based, in part, on their progress toward rater agreement.**

RIG 2 Districts (fall)

OSPI will offer up to 2 days (15 hours) of Stage 2 training March 2013 – August 2013 for all RIG 2 (Fall) districts who complete the Stage 1 training in February 2013.

RIG 2 Districts (spring)

OSPI will offer up to 2 days (15 hours) of Stage 2 training August 2013 – November 2013 for all RIG 2 (Spring) districts who complete the Stage 1 training in June 2013.

Stage 3	
Evaluators should know and be able to:	<i>Utilize the rubric for summative purposes. This includes assessing artifacts, assessing observation data, and scoring of the Washington State Criteria. At this level evaluators should provide immediate support for implementing the feedback and developing long-term professional development plans for teachers. OSPI does not provide training for this stage although portions of Stage 2 training may have applicability toward Stage 3. Stage 3 should be integrated into the evaluation of principals (Criterion 5) and district administrators.</i>
Success Indicators:	Summative Application of: <ul style="list-style-type: none"> • Appropriately use language of instruction and framework vocabulary • The interdependence of dimensions/domains and indicators/components • The key ideas in each criteria/indicators/components and what evidence would look like/sound like • The essential aspects of the framework • The evidence for all aspects of the framework and ability to determine summative criterion level scores using that evidence.
Districts should be creating plans to:	Develop and sustain rater agreement. This includes summative scoring against a pre-determined standard.

ICFFs/LCFFs (Instructional or Leadership Criteria and Framework Feedback Specialist): A team of Washington State practitioners who are trained by framework author designees to provide facilitation around Stage 1 and 2 on the specific instructional or leadership framework chosen by OSPI as the approved instructional or leadership frameworks. These practitioners applied and were selected by a state-level committee through a rigorous process. **The ICFFs and LCFFs are not experts in the new evaluation law. They are gaining expertise around the frameworks by the framework authors as the foundational tool for the new evaluation systems for teachers and principals.** If your district would like more information on the details of the new evaluation system please contact the TPEP office at OSPI.

Teacher and Principal Evaluation, OSPI
 Michaela Miller, TPEP, National Board, and BEST Director
Michaela.miller@k12.wa.us

Jim Koval, TPEP Director
Jim.koval@k12.wa.us

Rater Agreement Glossary

Accuracy: A measure of observer data quality indicating the extent to which an observer is assigning scores that agree with scores assigned to the same observation by an expert rater; the extent to which rater’s scores agree with the true or “correct” score for the performance.

Consistency: A measure of observer data quality indicating the extent to which an observer is assigning scores that agree with scores assigned to the same observation of practice by another typical observer. Consistency among the untrained is not what we are looking for. A goal of rater agreement is to ensure both accuracy and consistency.

Calibration: A process by which the regular practice of an observer’s scoring is monitored and verified that the observer is still scoring accurately and consistently according to the standards and definitions of the framework/rubrics.

Feedback: Information aligned with a rubric provided to reduce discrepancies between current performance and desired performance. Effective feedback answers three questions:

- Where am I? (What are the performance goals based on a self-assessment of the rubrics?)
- Where am I going? (How is my performance related to the rubrics?)
- Where to next? (What actions do I need to take next to increase my performance?)

Criterion: The standards for teaching and school leadership as defined by RCW 28A.405.100.

Framework: A language and structure of instruction or leadership that establishes the definitions of expertise and procedures for teachers and principals. The approved instructional and leadership framework chosen by OSPI.

Instructional

- [5 Dimensions of Teaching and Learning: Center for Educational Leadership \(UW\)](#)
- [Framework for Teaching: Charlotte Danielson](#)
- [Teacher Evaluation Model: Robert Marzano](#)

Leadership

- [Evaluating Principal Leadership in a Performance-Based School: AWSP](#)
- [Marzano School Leadership Evaluation Model: Robert Marzano](#)

Rubrics: The progression/description of practice used during an observation to capture evidence and classify teaching or leadership practice into differentiated aspects and performance levels. Typically consists of:

- Several Scales (components, domains, dimensions, indicators – there are numerous terms). See Architecture and Alignment documents on TPEP website.
- A set of score levels applied within each scale to classify performance. The score levels are described in Washington State as: Distinguished (4), Proficient (3), Basic (2), Unsatisfactory (1).

Summative Criterion Scoring: Rating given to performance based on Washington State Criterion (see definition above). These scores will be based on an ongoing and varied process using a preponderance of evidence to determine final summative scores that promotes and recognizes growth. *This is a district-determined process – Guidance coming from the Framework Author.*

Final Summative Scoring: Aggregation of the summative criterion scores. Final Summative Scoring Methodology approved by OSPI to determine final summative ratings. *Raw Score Model is a state-determined process – Guidance coming from the state WACs.*