



Proposed Rater Agreement in Montana State's Evaluation System, MT-EPAS Contents

- Rater Agreement Background
- Overarching Roles and Responsibilities: District Office, Principals, and Teachers
- Rater Agreement Definition and Stages
- Rater Agreement Glossary

Rater Agreement Background

Jigsaw Consulting and WM-PLC offers teachers and principals new learning and ongoing support to increase the effectiveness of implementation of the current district evaluation process by underscoring the importance of the shared experience between evaluator and evaluatee as a critical component of a professional growth model as approved by the Montana Board of Public Education State Statute, ARM 10.55.701.

To that end, the new rules support this intent:

ARM 10.55.701 4(a) The evaluation system for teachers used by a school district for licensed staff shall, at **minimum**:

- 4 (a) (ii) be aligned with applicable district goals, standards of the Board of Public Education; (4) (a) identify what skill sets are to be evaluated; (4) (a)(iv) include both formative and summative elements. JIGSAW Consulting and WM-PLC desire to:
- (1) acknowledge the critical importance of teacher and leadership quality and support professional learning as the underpinning of the new evaluation system.
 - (2) identify, in consultation with classroom teachers, principals, and assistant principals, particular areas in which the professional performance is distinguished, proficient, basic or unsatisfactory, and particular areas in which the classroom teacher needs to improve his or her performance.
 - (3) assist classroom teachers, who have identified areas needing improvement, in making those improvements.

Rater Agreement Proposal for Montana School Districts

Overarching Roles and Responsibilities: District Office, Principals, and Teachers

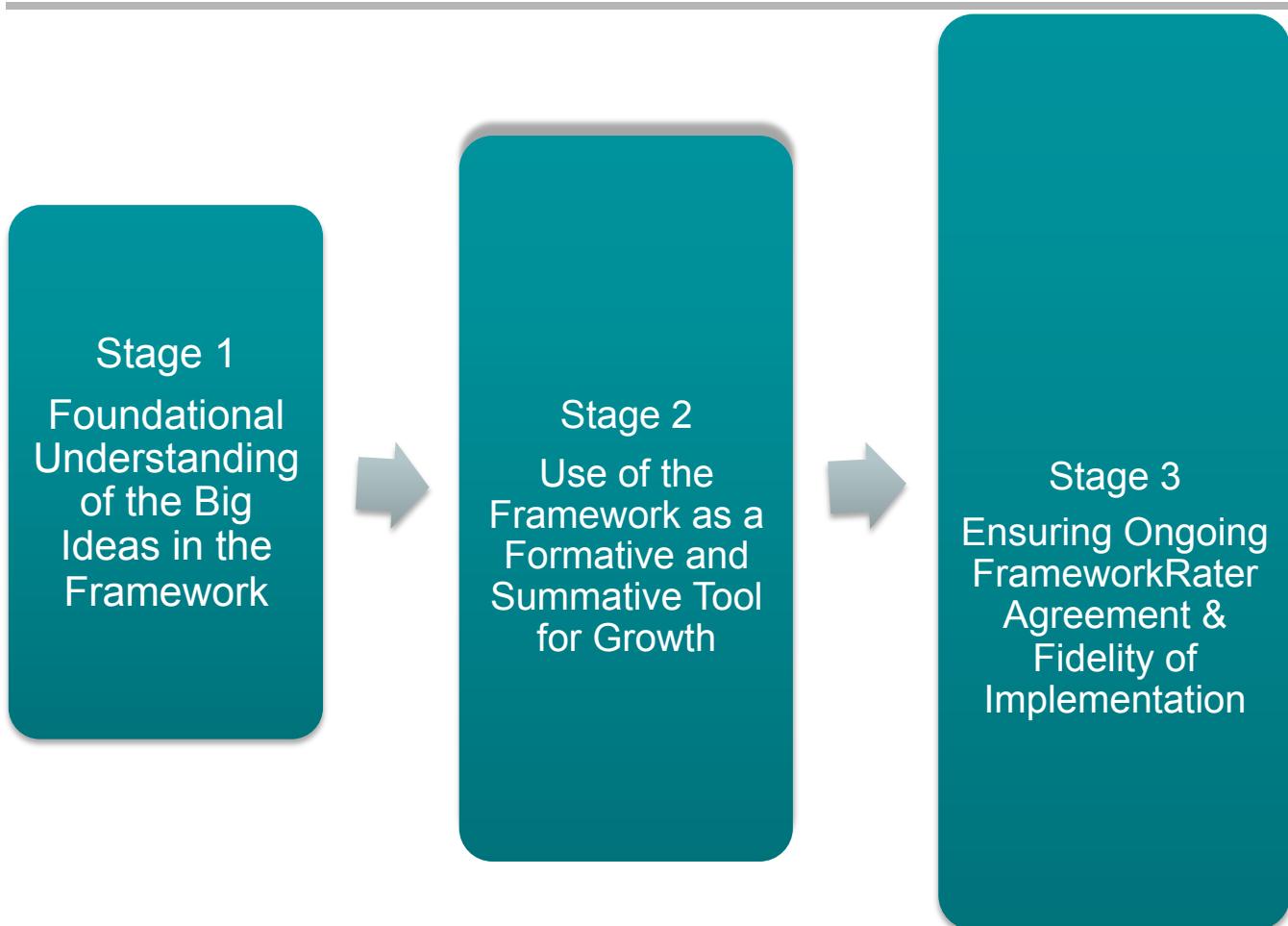
Jigsaw Consulting Offers Leader and Teacher Training on the Topics Below

	District Office	Principal	Teacher
Rater Agreement Implementation Responsibilities	<ul style="list-style-type: none"> Ensures fidelity of implementation of the chosen framework Leads the oversight of ongoing evaluator training and capacity-building across a district Moves evaluators toward maximizing rater agreement Provides the structures (e.g., principal PLC time) and resources for teachers and principals to learn, implement, and sustain growth-oriented evaluation. 	<ul style="list-style-type: none"> Leads the implementation of the growth-oriented teacher evaluation processes. Takes a collaborative learning stance with teachers in the evaluation process. Takes responsibility for learning and sustaining their own rater agreement. Provides the structures and resources for building level staff to learn, implement, and sustain the growth-oriented evaluation system. 	<ul style="list-style-type: none"> Engages in the implementation of the growth-oriented evaluation processes. Takes a collaborative learning stance with the principal and colleagues in the evaluation process. Takes responsibility for learning and applying the frameworks and rubrics to continually improve practice.
Communication	<ul style="list-style-type: none"> Provides communication and sponsorship for immediate and long-term implementation. 	<ul style="list-style-type: none"> Provides communication and feedback to support immediate and long-term implementation 	<ul style="list-style-type: none"> Provides communication and feedback to support immediate and long-term implementation.



Rater Agreement Definition and Stages

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



Stage 1

JIGSAW Consulting and WM-PLC, will provide a two-day approved Stage 1 training for all who evaluate classroom teachers. This two-day overview provides an understanding of the big ideas of the Instructional Framework and the inter-dependency of the frameworks, rubrics, and state criteria.

Evaluators should know and be able to:

Understand the big ideas of the Instructional Framework and the inter-dependency of the frameworks, rubrics, and state criteria. Accuracy in scoring depends on this foundational level. This foundational level understanding is the key to future calibration and application. ***It is recommended that this training occur prior to commencing the evaluation cycle.***

Success Indicators:

Understanding of:

- The appropriate use of the instructional language and framework vocabulary
- The interdependence of Domains/Components (Danielson)
- The definition of essential aspects of the framework
- The possible evidence for aspects of the framework
- The alignment of framework to Montana state criteria.

Districts should be creating plans to:

Provide the Stage I training to all evaluators before evaluating staff.

Stage 2

JIGSAW Consulting and WM-PLC will provide ? hours of approved, ongoing training on the frameworks for all who evaluate classroom teachers.

**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. Observation must be included in the application of this stage. ***It is recommended that this training occur 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
- Develop expertise to provide formative feedback through professional conversations/coaching on the framework and rubrics to foster teacher growth development and implementation of short-term professional development plans

Summative Application of:

- Analysis of evidence over time for the purpose of summative scoring
- Utilizing the criterion scoring document, guiding principles, and scoring matrix
- Utilize summative scoring methodology from Danielson Levels of Performance to sustain rater agreement. This includes summative scoring against the pre-determined standard identified in the framework and rubric.

**Districts should
be creating plans
to:**

Provide stage 2 training to all evaluators prior to summative scoring in the evaluation cycle. Provide for teachers to be supported in the formative evaluation process through framework training, self-assessment / goal setting, observation cycles, professional conversations, instructional coaching, building and district professional development and collaborative groups like professional learning communities or grade level / content teams.

Rater Agreement Proposal for Montana School Districts

Stage 3

Districts will continuously move evaluators to accuracy and consistency through ongoing calibration conversations involving real-life or video-based observation. JIGSAW Consulting, WMPLC and Teacher Forward Align provide training for this stage toward rater agreement.

Evaluators should know and be able to: *Accurately align and assess observation data to the standards in the rubric within the selected framework.*

Success Indicators:

- Appropriate use language of instruction and framework vocabulary
- The evidence for all aspects of the framework and ability to determine summative indicator or component scores using that evidence.
- The evidence for all aspects of the rubric and ability to determine summative criterion level scores for each state criteria.

Districts should be creating plans to: Provide for the ongoing professional development for all evaluators to demonstrate rater assurance.

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.

Artifact: Observed practice, products, or results of a certificated classroom teacher or certificated principal's work.

Evidence: Observed practice, products, or results of a certificated classroom teacher or certificated principal's work that demonstrates knowledge and skills of the educator with respect to the four-level rating system. (Danielson Levels of Performance)

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 as defined by ARM. 10.55.701 (4) (a).

Instructional framework: One of the preferred instructional frameworks adopted by the Office of Superintendent of Public Instruction to support the new evaluation system pursuant to ARM. 10.55.701 (4) (a).

The preferred instructional framework chosen by Montana Office of Public Instruction:

- [Charlotte Danielson's Framework for Teaching](#)

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)
- A set of score levels applied within each scale to classify performance. The score levels are described in MT.EPAS as: Distinguished (4), Proficient (3), Basic (2), Unsatisfactory (1)

Observe or observation: The gathering of evidence made through classroom or worksite visits for the purpose of examining evidence over time against the instructional or leadership framework rubrics.

Summative Criterion Scoring: Rating given to performance based on Danielson Levels of Performance (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 process is determined at the district level; guidance is available.

Final Summative Scoring: Aggregation of the summative criterion scores depends on multiple data measures delineated by the school district.

