Note: This micro-credential requires that an educator work with at least two other educators, ideally within an existing Professional Learning Group or Critical Friends Group. In addition, the educator must have already administered a performance assessment and have samples of student work to use in the calibration protocol.

Performance Assessment Reliability
Educator ensures technical quality through reliability by analyzing student work and scoring calibration.

Key Method
The educator scores student work reliably by using protocols to engage in collaborative, structured discussions. These discussions will focus on accurately and consistently documenting evidence within the student work of achievement of the learning target(s).

Method Components
The Advanced Performance Assessment for Learning Design stack is designed so that, if all three credentials are taken together, they will become more than the sum of their parts. Each micro-credential is intended to be able to stand on its own; however, the ideas and activities of each of these credentials support and expand on the others, allowing a fuller appreciation of performance assessment and its implications. Even more value will be gained by engaging in all three Performance Assessment for Learning stacks together.

What Is Assessment Reliability?
When we talk about assessment reliability, we are talking about the consistency with which different educators (or the same educator over time) administer, score, and analyze work. Is there agreement on what proficiency and proficient work looks like? Is there agreement on what the different levels of performance (as described in the rubric) look like in student work? Is the process used to identify student learning strengths and gaps and determine instructional strategies to be used to address the gaps? Is the process used to identify gaps in the task itself and determine ways to strengthen the task to improve the quality of student work that is produced? This agreement is achieved through a calibration process, conducted within a group of teachers.

Your Task:
The learner will conduct the protocol with the team at least twice—first for a student work sample from the learner’s own performance assessment and, second, for a student work sample from a colleague’s performance assessment. In other words, the first time through the protocol, the learner will be the presenter, while a colleague will be the facilitator. The second time through, the learner will be the facilitator. The presenter is responsible for providing enough copies of the materials for each team member.

Learner Participates in a Collaboration Protocol (Tool 4; see Resources) as outlined here:

1. Setting norms—Facilitator reminds team of the norms.
2. Present—Presenter briefly describes the context within which the assessment was administered.
3. Examination—Teachers look briefly at the task, student work, rubric, and score sheet.
4. **Clarifying questions**—Teachers ask the presenter any factual questions necessary for them to score the work.

5. **Read and score**—Group members independently read/view the student work, score it on the rubric, and take notes to back up their decisions. This should happen silently.

6. **Score sharing**—Each member shares his or her scoring without explanation. One member tallies the results of each member’s score.

7. **Discussion**—Group facilitator facilitates discussion on each rubric category, inviting members to talk about discrepancies between scores and the rationale behind the scoring and determining whether there is any movement in individual scores as a result of the discussion.

8. **Debrief**—Discuss next steps regarding revisions to the task, rubric, and/or instruction. Answer the following questions (as a group), taking notes:
   a. What adjustments can we make to the performance assessment task to better elicit evidence of what students know and can do?
   b. What teaching strategies can we use to address the learning gaps we identified to ensure that the student learns what s/he needs to learn?
   c. What are the specific directions that would elicit student evidence for the intended Depth of Knowledge (DOK) level?
   d. Could the rubric be more tightly aligned to the competency? How?

9. The learner will annotate the students’ work to indicate where the evidence exists for the scores given.

10. Repeat with other PLC members’ student work.

The purpose of the Calibration Protocol is to come to agreement within your PLC about what proficiency looks like for a particular competency/standard on a particular PA. Whatever else one learns from the protocol, the primary goal is to come to agreement. Agreement and clarity around what constitutes proficiency is essential for purposes of equity.

### Supporting Research


### Resources

- Video Module: Calibration [http://tinyurl.com/gzeyh2](http://tinyurl.com/gzeyh2)
- Video: Ten Minutes of Calibration [http://tinyurl.com/hqj9dmp](http://tinyurl.com/hqj9dmp)
Submission Guidelines & Evaluation Criteria

To earn the micro-credential, you must receive a passing evaluation for Parts 1 and 3 and a "Got it" for Part 2.

Part 1. Overview questions
(200-word limit for each response)

- Describe the calibration team and their area of expertise. (The team was a group of ninth grade ELA teachers, 9th grade team, etc.). Describe the history of the team (New ... been a PLC for years, etc.)
- Describe the PA used for calibration and the standards measured.
- Describe the context within which the performance assessment was administered, including the student whose work is being examined (no names, please). Include scaffolding or instruction that led to the task and class population, number of IEPs, etc.
  - **Passing**: The educator provides detailed and specific context for each answer.

Part 2. Evidence/artifacts
To earn the micro-credential for Performance Assessment Reliability, the educator must submit the following:

- Two pieces of student work that have been subjects of the calibration, one from the earner and another from a colleague.
- All materials used during calibration of the earner’s student work sample, including student work (before and after copies), rubrics, PA, meeting minutes, scoring sheets, notes, and worksheets from all participants.
- All materials used during calibration of a colleague’s student work sample, including student work (before and after copies), rubrics, PA, meeting minutes, scoring sheets, notes, etc.
- Ten-minute video or audio of earner leading a calibration meeting with an explanation of the protocol and/or a section of strong conversation.

<table>
<thead>
<tr>
<th>&quot;Got It&quot;</th>
<th>&quot;Not Yet&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video or audio demonstrates adherence to protocol and evidence of accountable discussion.</td>
<td>Discussion wanders away from protocol.</td>
</tr>
<tr>
<td>Notes on student work, rubric, and score sheet are aligned and show evidence of accurate and consistent feedback, are aligned to the learning targets, and reflect the calibration discussion.</td>
<td>Facilitator does not lead discussion but is passive and not in control.</td>
</tr>
<tr>
<td></td>
<td>Notes on student work, rubric, and score sheet are scattered, don’t show evidence of feedback, do not reflect the calibration discussion, or are incomplete OR not relevant.</td>
</tr>
</tbody>
</table>

Part 3. Reflection or other type of additional assessment

Write a reflective essay (1,000-word limit) OR record a five- to ten-minute video or audio addressing the following topics (be sure to use specifics that illustrate why you come to the conclusions you do):

- In general, what did you notice about leading the Calibration process? What were some challenges for you? How did others respond?
- Describe how the calibration process affected the team’s understanding of scoring, criteria, and proficiency (you may want to ask team members directly!).
- Were you able to come to agreement on the meaning of proficiency for the learning targets assessed?
- How did looking at student work and calibrating scoring impact your understanding of your assessment? Did it improve your understanding of whether or not the assessment was doing what it was supposed to do?
- What implications do you draw from the Calibration Protocol for your instructional practices?
  - **Passing**: Answer shows thoughtful consideration of the experience, making reference to specifics. In other words, the answer does not speak simply in generalities but is clearly based in the learner's experience with the reliability process.