

Protocols for Examining Evidence and Artifacts

Teams can use these protocols to ensure their work toward three purposes—(1) to calibrate scoring so it has consistency, (2) to analyze student errors, and (3) to refine assessments after student evidence clarifies whether they assessed the standards—produces common data. These protocols involve seven steps.

- Universal components (steps 1 and 2)
- Scoring calibration protocol (steps 3–5)
- Error-analysis protocol (steps 3–5)
- Assessment and measurement tools refinement protocol (steps 3–5)
- Universal components (steps 6 and 7)

If a team wants to focus on only one of the three protocols' purposes, it would use universal components steps 1 and 2, then jump to the specific protocol (each of which provides a different set of steps 3–5), and finally conclude with universal components steps 6 and 7. If, however, the team wants to engage in all three purposes, then it would follow the entire process from start to finish, addressing universal components steps 1 and 2, steps 3–5 for each individual protocol, and universal components steps 6 and 7.

Universal Components

Complete these steps every time.

Step 1: Make key decisions. If this is not done before the meeting starts, then quickly address it at the start of the meeting.

- How much time will be available for collaborative scoring? Availability of time will dictate the quantity of work the team can review.
- How large will each student sample be? What item, feature, or specific criteria will the team review? Oftentimes the entire assessment is too big to review, so a team might say, "We're only examining item 6 on the last test," or "We only want to look at academic transitions in the students' writing," or "We will only look at the second paragraph in each paper we review for our criterion of sentence fluency."
- Given how much time is available and the quantity of what the team needs to review, how many samples will we examine per teacher or per classroom? Note: Pragmatically, teams determine sample size by the length of each student project (fifteen-page research paper versus item 2 on the test), the amount of time available for the review, and the number of educators involved in the review process. This process is informal, so statistical reliability is not part of the equation. It is a best-guess situation.
- Teachers should not identify their own student work for review. Instead, team members should randomly select a range of quality work (for example, low-, medium-, and high-quality examples) from each others' student work piles.
- Have team members remove or hide students' names from the set of unscored samples they bring.

Step 2: During the meeting, review the assessment, its context, the standards, and the scoring tools, highlighting any clarifying questions or details that emerged during the assessment process.

- Decide if any changes in the scoring process need to happen at this juncture, but only make changes to the scoring criteria or proficiency levels that will benefit learners.
- Select which scoring criteria the team will use.
- Trade student work samples to start the scoring process. No teacher should begin with his or her own students' work.
- Distribute small sticky notes for easy scoring.

Scoring Calibration Protocol

Use these steps (continued from the Universal Components) to make sure the common assessment data are truly common.

Step 3: The team scores several samples of student work from each classroom collaboratively (as determined in step 1).

- Swap student work with each other so team members begin with someone else's student work.
- Individual team members review each sample before them and then place the preferred score (from the predetermined criteria or attending tool) for that sample on a sticky note. He or she places a sticky note on the student's paper or test in a place where it cannot easily be seen (such as on the back of the paper) and passes the paper to the next person to review and score (preferably without reading the previously assigned scores from other teammates). All papers rotate around the table until every paper is scored by each participating teacher in the process.
- Once all of the papers have gone all the way around the table to be scored, turn them over to reveal the collection of scores.

Step 4: Sort and stack results.

- Sort the work into similar stacks (all 3s, all 1s, and so on), adding a mixed number stack. If a paper has even one mixed number on it (such as five 2s and one 1), it goes in the mixed-number stack.
- During the sorting process, there is little discussion. Simply move the papers as quickly as possible to the appropriate designated stack.

Step 5: Go directly to the mixed-number stack to engage in a calibration discussion. This discussion is the most critical component of this protocol. The conversation is necessary to create shared understanding of common expectations so teachers can return to their classrooms and successfully score the remaining work that their sampling did not include.

- Discuss one sample at a time using the rubric or proficiency scale.
- Come to consensus as a team on what score to give the sample.
- Note the first few samples will take more time to score than the remaining samples.
- Review as many assessments in the mixed stack as possible. Once the team reaches shared understanding of how to interpret quality using the attending tools with a high degree of consistency, teachers can score the rest of the student work from their individual classrooms as they normally would have.

- The more this protocol is used, the faster the process goes and the more consistent the scoring becomes until eventually, there will be few disagreements on what a score should be. Consider using this protocol with frequency to ensure scoring remains consistent.

Note: It can be helpful at this point to identify anchor papers—sometimes identified as exemplars—for each designated level of proficiency. The identified sample of what each level looks like becomes the “anchor” that students can reference to more fully understand what each level of proficiency entails. Anchor papers can help students better understand the attending tools.

Error-Analysis Protocol

Use these steps (continued from the Universal Components) to identify and problem solve types of error in student work.

Step 3: Sort cards into appropriate categories.

- If the results of the assessment are binary, then quickly sort the student work into two piles: a right-answer pile and a wrong-answer pile.
- However, if the results of the assessment involve gradations of quality, then quickly sort the student work into the perceived level of proficiency; try not to get lost in the need for accuracy during the sorting process.
- Go quickly and trust that follow-up discussions might lead to moving student work into different proficiency-level stacks later.

Step 4: Isolate and categorize the errors the team finds.

- In the right-and-wrong sorted piles, take the wrong-answer stack and subdivide the answers into new stacks by common error. During the sorting process, do not get lost in the need for accuracy.
- In the proficiency-sorted piles, sort cards for proficiency levels 1 and 2 by category of error (proficiency levels 3 and 4 should be error free). During the sorting process, do not get lost in the need for accuracy.
- For each stack, name the type of errors found. Examine the samples in the error piles and discuss them as a team until an appropriate label for that type of error is selected. Move cards to alternative stacks if discussion reveals the error is something other than the team previously thought when they placed it in the initial stack.

Step 5: Generate ideas for corrective instruction (this step’s time frame depends on the number of samples involved).

- Discuss one error at a time, identifying the instructional tools and strategies, time frames, and so on that team members will use to support corrective instruction.
- Identify all the learners who will require a specific instruction.
- Review work samples in the *right* pile and the proficiency levels 3 and 4 piles to identify the consistencies that team members can further enhance or refine with extension or enrichment activities.
- Identify all the learners who will benefit from the extension or enrichment activities.

Assessment and Measurement Tools Refinement Protocol

Use these steps (continued from the Universal Components) to improve the quality of the assessment design or the attending tools based on the real-time evidence that emerged in the student work.

Step 3: Identify the trouble spots in the assessment (the places where many students got the answer wrong or misread the directions, or where the errors may have been particularly egregious).

- Name each area of the assessment that requires improvement.
- Isolate the issue that requires resolution.

Step 4: Identify next steps for assessment design.

- Have all team members answer the following questions to offer ideas for the next design steps that will improve alignment to the standards, cognitive demands, and critical competencies and literacies.
 - ▶ Is alignment the issue? Does the question, prompt, or task match the standard expectations?
 - ▶ Is clarity the issue? Are the directions or the expectations of each item easy to understand?
 - ▶ Is terminology the issue? Were students confused by the terms employed or the vocabulary in the text?
 - ▶ Is rigor the issue? Were all of the items created at the appropriate level of rigor?
 - ▶ Is relevance the issue? Did students seem interested? Engaged? Concerned?
- Identify the greatest areas of concern for improving the assessment, and target those areas first.

Step 5: Modify the assessment or attending tools.

- Question by question, prompt by prompt, or task by task, discuss, review, and improve the assessment.
- Criterion by criterion, review and improve the attending tools (scales, rubrics, or other measurement tools) for their clarity, accuracy, and overall impact on student motivation and efficacy.

Universal Components

Complete these steps every time.

Step 6: Reflect on the experience, and identify next steps, if there are any.

- Identify any additional learning needs for the team based on the findings in the meeting. If something more needs to happen, identify who will do which parts and by when, sharing responsibility whenever possible.
- Identify any needed resources for the team based on the findings in the meeting. If something more needs to happen, identify who will do which parts and by when, sharing responsibility whenever possible.
- Schedule the team's next review of student work; identify the purpose, content area or standard, and kind of work.

Step 7: Provide closure to the meeting. Review and debrief the following.

- **Process:** What worked? What could we have done better or differently to streamline the process? Are there things we need to change in our protocols?
- **Discoveries:** What did we learn? How does that learning inform our future work?
- **Decisions:** What decisions did we make? Who will enact them and by when?