**Assessment Design Tips and Guidelines**

**Stage 2: Determine acceptable evidence.**

What evidence will show that students understand and can meet unit goals?

**A Performance Task-**

* Requires *transfer—a* repertoire of knowledge and skill to be used wisely and effectively in a new situation (used with understanding).
* Asks students to “do” the subject, not just recall and plug in discrete learning, out of context.
* Is set in a novel situation, with little or no scaffolding or cues provided; the student has to think through what the task demands as part of the assessment (the “game” versus the scaffolded and simplified “drills”).
* Is as realistic as possible, requiring students to confront the same kinds of challenges, constraints, and options found in the real world.
* A task, like any other aspect of Stage 2, is meant to yield evidence of key elements in Stage 1. The  pri*mary* concern is validity, not whether the task is interesting or fun.
* A performance task may actually involve a  variety of situations, performances, and products (it can be a complex task, with related subtasks).
* Give the student a goal, a role, an audience, a setting, performance/product demands, and a set of standards and criteria by which work will be judged and it is likely the task will involve authentic transfer.
* Build in as much differentiation via options and alternatives in the situations as is feasible-without *corrupting the validity of the assessment.* (The various options should be relatively equal in what they demand and reveal about a student’s understanding.)
* The goal is sufficient evidence for  each student. Any group component to a task should be matched by evidence about the individual’s understanding. This can be accomplished by having separate parts to the task (such as a separate subtask in which roles and perspective change—for example, from group design team to solo reviewer) or by quizzes and prompts given to each student (and put in Other Evidence) that assess for the same goals.
* Don’t unwittingly assess for evidence unrelated to your goals. Keep in mind the “two-question validity test” and its implications: *Can the task be done well without understanding? Can the task be done poorly by someone with deep understanding?* If the answer is yes, then the task will not yield valid evidence, by definition. Be especially careful about demanding a mode or method of assessment that favors one student ability over others in ways that are not at the heart of your goals, so that, for example, you end up assessing—unfairly—writing ability or multimedia facility instead of understanding of the subject.

 **Rubrics**

* Clarify the criteria by which constructed-response work should be judged, and develop rubrics for each continuum of quality. It doesn’t matter whether you have a single rubric in a matrix form for the various traits or separate pages for each rubric related to each criterion. What matters is that you assess all the independent variables central to success.
* Use as many distinct criteria as needed to ensure excellent feedback. A good rule of thumb is to identify the fewest independent variables, but “accuracy” is independent of “well developed” and “creative.”
* Make sure that you identify valid criteria for scoring that suit the transfer goals and understandings, not just the particulars of the performance tasks (e.g., if the standard involves causes and effects of the Civil War, score “causal reasoning” and “insight of historical analysis,” not just “high-quality museum display”).

**Other Evidence (Quizzes, Tests, Prompts, Observations, Dialogues)**

* Identify the  specific *questions* related to key knowledge and skill goals from Stage 1 that you expect students to be able to answer upon completion of this unit.
* Typical tests, quizzes, and homework belong in this box; discrete and uncomplicated assessment of skills and factual knowledge that isn’t otherwise assessed in the performance tasks.
* Supplement all your performance tasks, as needed, to get more reliable and varied evidence of understanding, knowledge, and skill for *each individual student.* This is especially important if you claim that this unit addresses a standard in an in-depth way, and the performance tasks are basically group projects. Ultimately you need evidence for each student.
* Don’t confuse “assessment evidence” with “giving grades.” Just because you plan to assess it doesn’t mean you will grade it. Nor does the assessment “score” need to translate mechanically into a “grade.” If the task is difficult and new for students, then grade accordingly; if the point of the assessment is more for feedback, then don’t make it a grade for achievement, but only for process and effort, for example. This does not mean “don’t give grades.” It does caution not to confuse “feedback to students” with the separate act of “giving grades.” (Local grading policy may need discussion and revision as a separate issue.)

The Understanding by Design Guide to Advanced Concepts in Creating and Reviewing Units Module M: Authentic Assessment and Validity