Calling for a "Timeout" on Rubrics and Grading Scales

By: Rick Wormeli

We've been rubricizing everything students create these days: writing, reading comprehension, artwork, math problems, science labs, online practice, physical skills, collaboration, and even their degree of empathy for others.

To quantify the messy business of learning and to provide tools for analysis and accountability, however, we've developed some unhelpful habits, limiting rubrics' positives while enhancing their negatives. Before it gets any worse, let's throw a flag on the play and call a time out.

If we must grade students' work, standards-based grading is better than traditional grading. To use rubrics and grading scales in an evidence-driven way, however, teachers must negotiate with each other about what evidence they will tolerate at each level of performance: What in students' work constitutes a label of emergent, developing, proficient, or mastery? Robust to one teacher is superficial to another; one student's "extended effort" is another student's "barely lifted a finger" whiffle. How do we define adequate, satisfactory, and superior when it comes to knowing the constellations in the night sky or appreciating the poetry and power in ballet?

To be fair, we really can't hold such conversations until we have taught the subject. There's an intimacy and clarity that comes with teaching our disciplines that we can't achieve in undergraduate programs; we must be intimate with a subject before we can argue the merits of its elements. In addition, it's scary to reveal the extent of our personal knowledge base to respected colleagues because we may come up lacking.

The problem here is that students' futures are built or destroyed by the outcome of these deliberations with our colleagues. We can't leave evaluative criteria to chance or teacher indifference. The conversations must happen. They can be awkward and exhausting, but they are ultimately clarifying and liberating.

Basic Do's and Don'ts

As we turn away from percentages and the 100-point scale in the modern education world and instead embrace rubrics and smaller grading scales, let's shed light on some of practices that are hitting or missing the mark. Here are a few basic Do's and Don'ts when it comes to rubrics and grading scales:

Do's

1. Do use fewer levels. Three, four, or five levels is enough. The fewer the levels, the higher the inter-rater reliability (a 3.0 in one teacher's class describes the same level of content mastery as a 3.0 in another teacher's class), especially if teachers have personally vetted and calibrated the evidence for each level.

Imagine the ridiculous nature of writing evidence descriptors for every level of the 100-point scale. There aren't enough words that mean slightly lesser degrees of each other to write descriptors. Instead of increasing objectivity, such reductivity creates subjectivity and arbitrary sorting, claiming a precision that doesn't exist.

2. Reference the same domain all the way through. Rubrics and scales are about clear communication, so let's not muddy the waters. If we describe a student's level of strategic thinking in one descriptor, we refer to different proficiencies in strategic thinking in all levels. It's not helpful for one level of performance to describe whether certain portions of the project were completed while another level describes only the degree to which the student demonstrated strategic thinking.

3. Keep the evaluative criteria for each level authentic to the learner's experience. If we have students
practice one way in class and at home, but we test them a different way at the end of the unit, the report of their learning is invalid. For example, if we don't ask students to make novel applications of content and skills during their learning, but we ask them to do so on the final assessment, it's really not an assessment of what they learned. After wordsmithing our rubric descriptors, let's audit them for how authentic they are to the student's experience during the unit.

4. **Test-drive the rubric on real student work before giving it to students.** This prevents headaches down the road! If descriptors were so generalized students could interpret them in a dozen different ways, far from the true evaluative criteria, and we would still have to accept their responses. In test-driving rubrics, we find elements we forgot to include in the criteria, so we add them to the mix, and we find elements that really aren't that important, so we remove them.

5. **Provide exemplars for each level.** Students and parents need to know what constitutes each level of performance. There should be no surprises. The key here is transparency. Ask students to analyze their final product in light of the standard of excellence cited at the top of the rubric scale and to make a prediction for their final evaluation. Their predictions should come close to what is actually recorded.

6. **Ask students to design the evaluative criteria and rubric themselves.** Let them examine exemplars with a partner, searching for what qualifies them for excellence. Then ask them, as a class, to design the rubric to be used for the project under way. When the class agrees on an acceptable rubric, ask them to apply it to the assessment of another exemplar to see if it holds up. Help them adjust the wording and criteria if it doesn't. This process moves those criteria into students' internal editors, and they reference them in real-time while working on their own efforts.

7. **When providing multiple choices in projects or assessments, create and use only one rubric.** When we differentiate instruction, we often create a menu of three or more options for students to demonstrate final proficiency. Some of us have created eight different scoring rubrics when incorporating multiple intelligences in student assessments—one for each intelligence. This is not necessary, however.

Instead, create a list of criteria that should be expressed regardless of the path or vehicle used to present them. For example, students can create a video, an essay, a speech, a diagram, a 3D model, or a series of metaphors to demonstrate their learning, but no matter which one they choose, they must demonstrate accurate content, a thorough understanding of the topic, attention to craftsmanship, respect for the viewer/reader, at least two pieces of evidence for every general claim, references to the sources of their thinking, a strong voice, and anything specific to the topic.

8. **Reflect on the rubric's use and quality.** Use these mentoring questions with yourself or colleagues:

   - Does the rubric account for everything we want to assess?
   - Is a rubric the best way to assess this product?
   - Is the rubric tiered for this student group's readiness level?
   - Is the rubric clearly written so anyone doing a "cold" reading of it will understand what is expected of the student?
   - Can a student understand the content yet score poorly on the rubric? If so, why, and how can we change the rubric to make sure it doesn't happen?
   - Can a student understand very little content yet score well on the rubric? If so, how can we change that so it doesn't happen?
   - How do the elements of this rubric support differentiated instruction?
   - What should we do differently the next time we create this rubric?

**Don'ts**

1. **Don't use average, above average, or below average for the descriptor at any level.** These all speak to how the student is performing in relation to others. If we're criterion-referenced, we report student performance in relation to the lesson's goals, the standards: *Can he use and interpret a Punnett Square?* It's not helpful to hear that a student's work is "above average" when the average could be anywhere and doesn't identify specific
content and skill targets.

2. **Don't write out every level of descriptors for most assessments.** For some students this is helpful, but in their busy, selectively attentive lives, many students barely register a rubric, let alone think seriously about each level. When they do look at them, they'll settle for the wording of the lower levels: *I don't have to be "exemplary,"* they think, *I just have to be "satisfactory."* If all they see is the fully explained descriptor for excellence, however, they'll know nothing else. They rally around that vision.

3. **Don't let reports of compliance distort reports of learning.** Helpful rubrics are not reports of what students did; they are reports of what students learned. Double-check that the rubric isn't merely reporting what students completed so much as where they are in relation to learning goals.

4. **Don't use symbols with a natural sequence.** In standards-based grading, using evidence-reporting rubrics, we inappropriately worship at the math altar, thinking the math adds credibility. Actually, it corrupts the original goal: clear and accurate communication. On a 4.0 scale, for example, a 2.0 is usually equated with the letter grade C, but that's not supposed to be, "C out of A," because that would be a 2 out of 4, or 50%, and that's usually an F grade.

The grade, number, or symbol is supposed to be a placeholder for a much longer description of evidence. By itself, it is nonsense, communicating nothing without the evidence associated with it. Yes, you can use letter grades in standards-based grading. You can also rubric numbers as long as they directly reference evidence descriptors.

**Rubric Revelations**

The conversations we conduct on rubrics and grading scales are some of the most liberating and inspiring ones we have. They often lead to serious revelations about instruction and learning, and with each conversation, we find one more reason to get out of bed in the morning and teach. Our students can't help but learn something useful those days.

*Rick Wormeli* is a long-time classroom teacher turned writer and education consultant. He is the author of several books, including *The Collected Writings (So Far) of Rick Wormeli: Crazy Good Stuff I Learned about Teaching Along the Way (AMLE).* He lives in Herndon, Virginia, and is currently working on a new book on homework and his first young adult fiction novel.

Published in *AMLE Magazine*, October 2015.