The Journal of Effective Teaching
an online journal devoted to teaching excellence

The Role of Rubrics in Advancing and Assessing Student Learning

Kenneth Wolf[†] and Ellen Stevens
University of Colorado at Denver and Health Sciences Center

Abstract

A rubric is a multi-purpose scoring guide for assessing student products and performances. This tool works in a number of different ways to advance student learning, and has great potential in particular for non-traditional, first generation, and minority students. In addition, rubrics improve teaching, contribute to sound assessment, and are an important source of information for program improvement. In this article, we discuss key features of a quality rubric, present an example of a rubric for assessing a social science research study, and describe three basic steps in designing an effective rubric.

Keywords: Rubrics, assessment, planning, instructional design.

While schoolteachers and their students have long seen the value of assessment rubrics, our experience in working with faculty is that rubrics have been largely ignored in higher education contexts (with the exception of Schools of Education). These multi-purpose scoring guides for assessing student products and performances work in a number of different ways to advance the goals of an educational program. Not only do rubrics contribute to student learning, they have great potential for non-traditional, first generation, and minority students. As well, rubrics improve teaching, provide feedback to students, contribute to sound assessment, and are an important source of information for program improvement.

So, what exactly are rubrics? How are they developed? What are their key features? Why are they useful? What are their limitations? What role can they play in program improvement? These questions, and more, will be addressed in this article.

Before we define and describe rubrics, here are a couple of scenarios to help set the stage (modified from Arter & McTighe, 2001, pp. x-xi):

> An undergraduate student in an American History course spent many hours work-

> are 2cD -0.er o6sa2aduate st7 o6sa2aduate st70.0347 Tc 1.1f1raduate studen 769 prograeivenE-3

> an example, the professor replied, "Well, you could have presented it from the point of view of the Native Americans affected by the Gold Rush."

What's the problem here…? There are no explicit performance criteria to inform students in creating their projects or to guide the professor in assessing them. A rubric here could help address this situation.

How do you think this student felt? Probably the same way that students in any course feel when the criteria for an assignment are ambiguous and the assessment seems arbitrary. When the curriculum is "hidden," students who can't guess what the expectations are will be more at risk than those who know how to "play the game" (Jackson, 1990). A good rubric can take the mystery out of assignments for *all* students. As Eisner notes: "More than what educators say, more than what they write in curriculum guides, evaluation practices tell both students and teachers what counts. How these practices are employed, what they address and what they neglect, and the form in which they occur speak forcefully to students about what adults believe is important" (Eisner, 1991, p. 81).

Now, let's look at another scenario:

> In an English department class, a professor introduced her students to the qualities of an effective oral presentation by showing them videotaped examples of excellent, as well as poor, speeches and presentations. Guided by the teacher, the students identified four key criteria (traits) that they agreed were important for an effective speech—content, organization, delivery, and language. They defined each of these and what would constitute strong, middle, and weak performance on each trait. They then referred to these performance criteria when preparing their own speeches, and the teacher used the same criteria when providing feedback on, and grading, their presentations.

What's going on in this scenario? Not only are there criteria that define the features of a speech, but the professor has shown strong and weak examples of oral presentations and even invited the students to generate evaluation criteria based on these examples and their own experiences. Clearly, both students and professor use the criteria in talking about and giving feedback on the speeches. In other words, the learning process is anchored by a rubric--a scoring tool used to evaluate a performance in a given outcome area based on a list of criteria describing the characteristics of products or performances at varying levels of accomplishment.

A Rubric for Springboard Diving

We always have criteria in mind when we evaluate something–whether it's a piece of art or a dive off a springboard. It's just that these criteria aren't always explicit, sometimes even to ourselves. When we judge a springboard diver's performance as good or bad, for example, we are basing that judgment on something. We have some criteria in mind. Maybe it's the number of body rotations or the splash the diver makes on entry. Maybe it's something that really has nothing to do with the performance itself such as the diver's

smile or nationality.

As we become more informed about springboard diving, though, we may begin to draw on the five criteria used by the professional association (Federation Internationale de Natation, 2006): Starting Position, Take Off, Approach, Flight, and Entry. These criteria are then elaborated in a rubric that describes what we mean by each. "Entry," for example, is based on a number of considerations about body position. "The entry into the water shall in all cases be vertical, or nearly so, with the body straight, the feet together and the toes pointed. When the entry is short or over, the judge shall deduct according to his opinion (p. x)." Each of these criteria is then described on six levels of performance from "complete failure" to "very good" (see Table 1).

A rubric in springboard diving makes it more clear to the judges how to rate the performance, though these judges still need to draw on their extensive professional knowledge in applying these criteria. As well, coaches study the criteria so that they can provide effective instruction to their athletes. And the athletes themselves examine the criteria to guide them in planning and perfecting their dives. In the same fashion, for an assignment in a course or for other types of learning experience, such as studios or internships, learning is best achieved if all participants are clear about the criteria for the performance and the levels at which it will be assessed.

Table 1. Springboard Diving Rubric

	Complete Failure	*Unsatisfactory*	*Deficient*	*Satisfactory*	*Good*	*Very Good*
Starting						
Take-off						
Approach						
Flight						

Three to six criteria seem to work best. It is not so many that it overwhelms the memory and not so few that meaningful distinctions in the performance can't be made. Sometimes these criteria can be weighted as well. There may be one or two criteria that are valued more than the others and they could be given a higher value when calculating the overall score for the performance or product.

Another important consideration is that the performance to be assessed should be observable and measurable. Some descriptions of learning outcomes or performance criteria are so vague that accurate measurement is difficult. For example, if the criterion is that "Students will know the states of the union," it may not be clear what "know" means. Does 'knowing" mean that students need only to be able to list the states, or be able to fill in the names on a map, or draw a map of the United States, or discuss the history of the state, or ….? The measurement problem can be lessen if the performance to be assessed is described with more specific action verbs where possible, such as list, identify draw, discuss, explain, compare, critique, predict, and so on.

Often the performance criteria are determined ahead of time by the instructor or a professional organization, but sometimes they can be created by the students in a course, especially if the assignment is new to the instructor. Having students generate the criteria for assessing the performance can serve several purposes. Engaging students in a discussion about "What makes for a good speech" (or essay or model or dance or…) can help them deepen and internalize their understanding of the criteria for a quality performance in that particular area. As well, involving students in this conversation before they begin the assignment or project can help them make more informed choices as they begin to identify the topic for their laboratory study, the medium for their performance, or the design for their model. Another benefit is that students can sometimes offer insights into the performance that the instructor may not have envisioned. When a student asks if their oral presentations can be a video of themselves before a live audience rather than a live in person in class presentation, it can open possibilities the instructor hadn't considered. An additional pedagogical benefit is that the students' comments can reveal to the instructor misconceptions that students may have about the topic, and the instructor can adjust his or her teaching of these concepts accordingly. A valuable activity can be to make a list of the assessment criteria that students identify as the project is introduced and another list again after they have completed the project, and then have them compare their pre-and-post lists to see if their understanding of the key concepts have changed or deepened. Even if the rubric has already been developed in advance however, asking students to engage in a discussion about the assessment criteria before the rubric is handed out can still be a valuable activity for many of these same reasons.

Setting Performance Levels. The second step in the process is to decide how many levels of performance are appropriate for the assessment. Typically, rubrics have from three to six rating levels. What drives the choice of the number of levels is the purpose for the assessment. If the main purpose is to make summative decisions, such as whether someone will pass or fail a course or an exam for example, then fewer levels are better. The fewer the levels of performance for the rater to consider, the greater the reliability and effi-

Table 3. Speech Rubric with Performance Statements for the "Delivery" Criterion

	Below Proficient (1)	*Proficient (2)*	*Beyond Proficient (3)*
Delivery • Volume • Pacing • Rapport	It is difficult to hear the speaker, and the pace is either too slow or too fast. Speaker has little connection with audience.	Speaker is easy to hear and pace keeps audience's attention.	Speaker varies volume to fit the message, with a pace that is appropriate to the rhythms of the topic. Audience is clearly engaged.

When using the rubric in making an overall decision about a performance, the final rating can be based on an analytic process of adding up the scores for each of the four criteria (i.e., content, delivery, language, physicality) and calculating an average, or, alternatively, by looking over the ratings for the four criteria and making a holistic judgment that considers each of the scores but blends them in an overall judgment-based rating process. For example, if the scores were delivery = 2, content = 3, organization = 2, and

Table 5. Rubric for Research Project in Education

	Below Proficient	*Proficient*	*Above Proficient*
Abstract	The abstract is missing, incomplete, or inaccurate.	The abstract summarizes the study in 50-150 words (essentially drawing a sentence from each of the main sections of the completed research report).	The abstract concisely summarizes the study in 50-150 words.
Introduction	The introduction section may be incomplete or unclear. Potential problems may include a vague problem statement, research question(s) may not be measurable, or constructs may not be clearly defined.	The introduction section includes a rationale, problem statement, literature references and research question(s). The rationale and problem statement are clear and credible. Three or more literature references are cited. The research question is stated and can be addressed with empirical evidence. Constructs are defined and variables explained.	The introduction section is complete and clear. Additionally, the rationale and problem statement are compelling (and may be linked to a conceptual framework) and the research question(s) insightful.
Methods	The methods section may be incomplete or unclear. Possible problems may include insufficient information about subjects/informants, instruments not fully described in terms of their conceptualization or aligned with the research questions, or procedures not accurately reported.	The methods section provides essential information about the subjects, data collection procedures, and, if appropriate, treatment. The research question has been translated into appropriate choices at the design level. Subjects are described in terms of number and important characteristics. Data sources and collection procedures are described in terms of underlying conceptualizations. If appropriate, scales are described, and examples of items given. Data collection protocols (e.g., questionnaires, interview questions, structured observation protocols) are included in the appendix.	The methods section provides essential information about the subjects, data collection procedures, instruments, procedures, and, if appropriate, treatment. In addition, the instrument or procedures, for example, might represent a novel and insightful approach to the research problem.

While not a panacea, the benefits of rubrics are many—they can advance student learning, support instruction, strengthen assessment, and improve program quality.

References

Andrade, H., & Ying, D. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation, 10*(3), 1-11.

Arter, J. & McTighe, J. (2001). *Scoring Rubrics in the Classroom*. Thousand Oaks, CA: Corwin Press.

Delpit. L. (1988). The Silenced Dialogue: Power and Pedagogy in Educating Other People's Children. *Harvard Educational Review, 58*(3), 280-298.

Eisner, E. (1991). *The Enlightened Eye: Qualitative Inquiry and the Enhancement of Educational Practice*, New York: Macmillan.

Federation Internationale de Natation. (2006). *Rules and Regulations: FINA Diving Rules 2005-2009*. Retrieved January 27, 2006 from http://www.fina.org.

Hafner, J. C., & Hafner, P. M. (2004). Quantitative analysis of the rubric as an assessment tool: An empirical study of student peer-group rating. *International Journal of Science Education, 25*(12), 1509-1528.

Heath, S. B. (1983). *Ways with Words.* Cambridge: Cambridge University Press.

Jackson, P. (1990*). Life in Classrooms.* New York: Teachers College Press.

Stiggins, R. (2001). *Student-Involved Classroom Assessment* (3rd ed.). New York: Merrill.