Scoring Rubrics

[FOR WRITTEN ASSIGNMENTS AND ORAL PRESENTATIONS]

Diane Ebert-May
Lyman Briggs School
Department of Botany and Plant Pathology
Michigan State University

WHY USE RUBRICS?

Has a student ever said to you regarding an assignment, "But, I didn't know what you wanted!" or "Why did her paper get an 'A' and mine a 'C?'" Students must understand the goals we expect them to achieve in course assignments, and importantly, the criteria we use to determine how well they have achieved those goals. Rubrics provide a readily accessible way of communicating and developing our goals with students and the criteria we use to discern how well students have reached them.

WHAT IS A RUBRIC?

Rubrics (or "*scoring tools*") are a way of describing evaluation criteria (or "*grading standards*") based on the expected outcomes and performances of students. Typically, rubrics are used in scoring or grading written assignments or oral presentations; however, they may be used to score any form of student performance. Each rubric consists of a set of scoring criteria and point values associated with these criteria. In most rubrics the criteria are grouped into categories so the instructor and the student can discriminate among the categories by level of performance. In classroom use, the rubric provides an "objective" *external standard* against which student performance may be compared.

WHAT IS INVOLVED?

Instructor Preparation Time: Medium to High
Preparing Your Students: Continuous; but students catch on fairly quickly
Class Time: Variable. As students use rubrics, they become better writers and oral presenters; hence the time instructors spend evaluating students' work is reduced.
Disciplines: All
Class Size: All. Rubrics are easy to use in small classes, and are particularly useful in large classes to facilitate scoring large numbers of written or oral assignments.
Individual/Group Involvement: Both.
Analyzing Results: The level of analysis depends on the instructor's intended goal of the assessment task and the type of data desired about students' performance. For detailed analysis of students' responses, each section of the rubric can be scored independently then totaled. For a holistic analysis of students' responses, all sections of the rubric can be blended and an overall score assigned.
Other Things to Consider: Rubrics must be readily available to students before they begin an assignment or written test. Posting rubrics on the web and including them in the course pack for in-class writing promotes their usefulness.

Contents

Description

Example 1. Scoring Rubric for Quizzes and Homework

Level of Achievement	General Approach	Comprehension
Exemplary (5 pts quizzes)	•Addresses the question. •States a relevant, justifiable answer. •Presents arguments in a logical order. •Uses acceptable style and grammar (no errors).	•Demonstrates an accurate and complete understanding of the question. •Backs conclusions with data and warrants. •Uses 2 or more ideas, examples and/or arguments that support the answer.
Adequate (4 pts quizzes)	•Does not address the question explicitly, although does so tangentially. •States a relevant and justifiable answer. •Presents arguments in a logical order. •Uses acceptable style and grammar (one error).	•Demonstrates accurate but only adequate understanding of question because does not back conclusions with warrants and data. •Uses only one idea to support the answer. •Less thorough than above.
Needs Improvement (3 pts quizzes)	•Does not address the question. •States no relevant answers. •Indicates misconceptions. •Is not clearly or logically organized. •Fails to use acceptable style and grammar (two or more errors).	•Does not demonstrate accurate understanding of the question. •Does not provide evidence to support their answer to the question.
No Answer (0 pts)		

Assessment Purposes

- To improve the reliability of scoring written assignments and oral presentations
- To convey goals and performance expectations of students in an unambiguous way
- To convey "grading standards" or "point values" and relate them to performance goals
- To engage students in critical evaluation of their own performance.

Limitations

- **The problem of criteria**: One challenge (or potential limitation) in developing and using rubrics is that of *describing performance expectations*, and *defining the criteria* that differentiate several levels of performance. Quantitative descriptors may be helpful in differentiating among performance levels, e.g., "provide 2 examples of evidence that support the conclusion...," or "zero grammatical errors". Whereas variables such as *vague* or *concise*, must be described clearly so that students see the differences between a statement that is *vague* and a statement that is *concise*. By carefully describing our performance expectations and defining the criteria we use to differentiate levels of performance, our role as evaluators becomes easier, more informative, and more useful to our students and us.
- **The problem of practice and regular use:** No assessment tool is effective if it is not used on a regular basis. Rubrics are most effective when we practice using them with our students over and over again. Developing effective rubrics requires revision based on feedback from students. The best rubrics are products of an iterative effort.

Example 2. Scoring Rubric for Grant Proposals (35 points possible)

Level of Achievement	General Presentation (10 points possible)	Conceptual Understanding (10 points possible)	Argument Structure (10 points possible)	Use of literature and pertinent resources (5 points possible)
Exemplary	*(10 pts.)* • Provides a clear and thorough introduction and background • States a specific, testable research question • Provides clear explanation of proposed research methods • Presents rationale and significance of proposed research in the form of a well-structured, logical argument. • Uses acceptable style and grammar (0 errors)	*(10 pts.)* • Demonstrates a clear understanding of the LTER site and the proposed research. • Uses a broad range of information to build and support arguments. • Demonstrates a good understanding of the implications of the data and/or information.	*(10 pts.)* • Provides strong, clear, convincing statements (i.e. conclusions) of the reasons the proposed research is important and should be funded. • Provides relevant evidence to support conclusions. • Provides reasons for the legitimacy of the evidence (i.e. warrants) that enable conclusions.	*(5 pts.)* • Follows proper format in providing citations. • Uses data and/or information relevant to the proposed research
Adequate	*(8 pts.)* • Provides an	*(8 pts.)* • Demonstrates a	*(8 pts.)* • Provides statements	*(4 pts.)* • Follows proper

	introduction and background that is only somewhat significant to the experiment. • States a clear, but untestable research question. • Provides an adequate explanation of proposed research methods • Shows some effort to present the rationale and significance of proposed research in the form of a well-structured argument. Uses adequate style and grammar (1-2 errors)	partial understanding of the LTER site and the proposed research. • Uses a information from only 2 or 3 sources to build and support arguments. • Demonstrates a partial understanding of the implications of the data and/or information.	(i.e. conclusions) explaining the reasons the proposed research is important and should be funded, but weak evidence to support conclusions and no warrants.	format in providing citations, but not consistently throughout the proposal. • Uses limited number of sources of data and/or information relevant to the proposed research.
Needs Improvement	• *(6 pts.)* • Provides an introduction and background that is insignificant to the experiment. • States a vague, untestable research question. • Provides an unorganized explanation of proposed research methods • Presents rationale and significance of proposed research in the form of a weak, unstructured argument. • Fails to use acceptable style and grammar (>2 errors)	• *(6 pts.)* • Does not demonstrates an understanding of the LTER site and the proposed research. • Uses less than two sources to build and support arguments. • Does not appear to understand the implications of the data and/or information.	• *(6 pts.)* • Provides statements (i.e. conclusions) explaining the reasons the proposed research is important and should be funded, but no evidence to support conclusions and no warrants.	• *(3 pts.)* • Does not follow proper format in providing citations. •

Teaching and Learning Goals

Students learn to communicate about science in a variety of ways and especially improve their writing skills. The quality of students reasoning and logic increases. Instructors gather a variety of data about students' understanding and performance.

Suggestions for Use

I design rubrics for the multiple forms of assessment I use in my courses: *short writing samples, essays, poster displays, research papers, public hearing papers, oral presentations, weekly homework assignments, and concept maps*. Each rubric stands on its own, but the general criteria in many rubrics are similar. For example, rubrics for written assignments have the same criteria for acceptable style and grammar; responses must address the question, and arguments must be presented in a logical order [compare Figures 1 and 3]. Alternatively, a rubric designed to evaluate the mechanics of a poster display may include a checklist to guide the student in developing all of the sections of the poster. Each component of the poster should then have additional criteria for evaluation. For example: What are the criteria for the title of a poster? Is the title informative? Are specific key words used?

Example 3. Scoring Rubric for Essay Questions

Level of Achievement	General Presentation	Reasoning, Argumentation
Exemplary (10 pts)	• Provides a clear and thorough introduction and background • Addresses the question • Addresses the question • Presents arguments in a logical order • Uses acceptable style and grammar (no errors)	• Demonstrates an accurate and complete understanding of the question • Uses several arguments and backs arguments with examples, data that support the conclusion
Quality (8 pts)	• Combination of above traits, but less consistently represented (1-2 errors) • Same as above but less thorough, still accurate	• Uses only one argument and example that supports conclusion
Adequate (6 pts)	• Does not address the question explicitly, though does so tangentially • States a somewhat relevant argument • Presents some arguments in a logical order • Uses adequate style and grammar (more than 2 errors)	• Demonstrates minimal understanding of question, still accurate • Uses a small subset of possible ideas for support of the argument.
Needs improvement (4 pts)	• Does not address the question • States no relevant arguments • Is not clearly or logically organized • Fails to use acceptable style and grammar	• Does not demonstrate understanding of the question, inaccurate • Does not provide evidence to support response to the question
No Answer (0 pts)		

Step-by-Step Instructions

There are many routes to developing a useful scoring rubric, however, all of them involve the following five steps:

- Develop the goals for your course and daily class meetings
- Select the assessment tasks that provide data aligned with the goals
- Develop performance standards for each of the goals
- Differentiate performances (categories) based on well-described criteria.
- Rate (assign weight or value to) the categories

To provide a useful example of how these steps "play out" in a real world context, I will describe how I developed rubrics for my own introductory biology and ecology courses.

1. I developed the goals for my course and daily class meetings. Keep in mind the assessment tasks must be linked to student learning goals and outcomes. So writing goals is the first step. These are examples of stems and sample goals from introductory ecology or biology courses:

Students will be able to demonstrate their ability to:

- utilize science as a process
- communicate an understanding of and links among biological principles
- write about, criticize and analyze concepts in biology
- use the process of scientific inquiry to think creatively and formulate questions about real-world problems
- apply content knowledge in the resolution of real-world problems
- reason logically and critically to evaluate information
- argue science persuasively (in both written and oral format)
- illustrate the relevance of ecology to your lives by applying ecological knowledge in the resolution of real-world problems

2. I selected the assessment tasks.

What type of assessment will provide me data about students' achievement of each of these goals?

Based on the goals for my courses, I selected different forms of extended responses, both written and oral, and concept maps to gather the data that would convince me that my students achieved the goals. The kinds of questions I asked students and the types of projects I assigned, were designed to promote students' reasoning. For example, for the first three goals I have listed, various types of assessment that could be used to gather the type of data desired.

- *Utilize* science -- performance assessment e.g., students conduct a scientific investigation
- *Communicate an understanding* of and links among biological principles -- e.g., concept maps, Vee diagrams, extended written responses (Novak and Gowin 1984, Novak 1998).
- *Write about, criticize and analyze* concepts in biology -- written critical analysis of articles and papers.

3. I developed a set of performance standards.

The performance standards I used in my introductory biology course on "logical reasoning" and "critically evaluating information" were different than the performance standards I developed for my upper division biology majors. The difference was based on the developmental stages of the students and their experience in college-level science courses (Magolda 1992, King and Kitchener 1994).

4. I differentiated performances based on criteria.

Examine the rubric for Quizzes and Homework. The criteria for responses fall into two major categories: general approach and comprehension. Although these two categories are not discrete as indicated by the dotted line between them, students can see all of the itemized components of an exemplary answer. These categories can be divided further. For example, comprehension could be divided into content knowledge, conceptual understanding, and reasoning and critical thinking skills (Freeman 1994). Freeman (1994) includes communication skills as a category in rubrics. Essentially, my rubrics cover the same categories; the difference is in the number of columns used.

Notice, when it is possible to quantify the categories, I did so. So, for example, the criteria for acceptable style and grammar in an exemplary answer is based on *no errors*.

Our ability to differentiate among criteria is critical to the effectiveness of the scoring rubric. So words like "good" are too subjective. The criteria must be designed so that you and your students can discriminate among the qualities you consider important.

When we evaluate students' extended responses, we tend not to score them point by point, however, by elaborating on the criteria that comprise the different levels of performance, we provide the students substantive guidance about what should be included in their extended responses.

5. I assigned ratings (or weights) to the categories.

- Exemplary (5 pts) - highest category of college-level work
- Adequate (4 or 3 pts) - acceptable college-level work
- Needs Improvement (3 or 1 pts) - not yet college level-work
- No answer: 0 points

Point values: Do you assign points on a 5, 3, 1 scale? or a 5, 4, 3 scale? I have tried both. I chose 3 as the middle or as an adequate score. Most student responses in this category can readily be improved through group work, practice, effort and instruction. Therefore, in an effort to develop students' self-efficacy and to promote their achievement of higher standards, I chose the 5,4,3 point scheme.

On a five-point scale, the data do not enable me to discriminate between two consecutive points, such as 3 and 4, in terms of evaluating the response. Rather, three categories were readily distinguishable by my students and me, therefore, little if any time was spent "arguing" for points. The criteria for evaluation were clear and understood.

Variations

Student roles

- I involve students in a dialogue about criteria that we use in any rubric. Students gain a keen sense of my expectations for assessment by explicitly understanding the criteria and by contributing to the modification of criteria in a rubric to enhance clarity.

- Consider letting students develop class rubrics for certain assessments. When students understand and describe criteria for their own performance, they are often better at attaining those standards. My students developed the rubric for the poster displays of their laboratory research. To accomplish this, they walked around the biology department looking at the variety of posters displayed on the walls and then determined their own criteria for what makes a quality poster presentation. We collated this information and designed a rubric for content as well as format.

- Students use rubrics when completing any assessment task for the course such as writing in class, writing on an exam, designing homework, completing and investigation, preparing a research paper.

Faculty Roles

- The critical factor for faculty to consider is that assessments must be linked to the goals of the course. For example, if the goal is for students to demonstrate their ability to design a testable hypothesis in a particular content area - asking students to actually write a testable hypothesis would provide meaningful feedback. The recurring question we must ask is, "Does this evidence convince us that students understand how to write a testable hypothesis?"

- Include rubrics on your web site and in your course packs. Students should refer to rubrics while they are completing any assessment task.

- Rubrics are dynamic and involve a major up-front investment of time.

- You must provide students repeated instruction on how to use rubrics as well as how to achieve each of the criteria.

  Share with students samples of "exemplary", "adequate", "needs improvement" responses. Ask them to work in cooperative groups to analyze the strengths and weakness of the written responses, using the rubric as a guide. With practice, students learn to recognize and ultimately develop their own exemplary responses.

- The advantage of rubrics is that you and the students have well defined pathways for gathering evidence that they have achieved the goals of the course. If either you or your students are dissatisfied with the evidence or criteria, the rubrics should be revised.

Analysis

Rubrics are scoring tools that enable me to assign points to students' assignments and tests. Students' accumulation of points determines their grade in the course. Each assignment, quiz, or test is weighted in terms of value in the overall course evaluation. For example, daily writing samples (quizzes) are worth 5 points, twice weekly, 15 weeks per semester; hence a student can earn a maximum of 75 points for daily performance. The pattern of students' performance is consistent from semester to semester. At the beginning of each semester, many students' responses are below college-level. As students begin to understand the criteria and practice writing, they attain college-level work or exemplary performance on short, five-point assignments or quizzes. A key strategy in promoting improvement by all students is peer review within their cooperative groups.

The formative assessment I gather by using rubrics to evaluate students' responses during the course is valuable. In-class writing assignments give me feedback about the nature of the task and questions I ask students. The components of a question or task that provide meaningful responses are readily identifiable from the rubric and provide us insight into my students' strengths and weaknesses. I use these data to modify, change directions, or add components to our instructional design and strategies.

Pros and Challenges

- Faculty develop and communicate assessable course goals.
- Faculty and students understand and agree upon the criteria for assessment.
- Rubrics provide ways to evaluate many different types of assessment.

However:

- Time - rubric development requires time up front, but the payoff is in increased performance by the majority of students and which ultimately leads to less instructor time in assessment.
- Criteria - qualitative scales are more difficult to define than quantitative scales.
- Practice - both students and faculty need to practice and refine the use of rubrics for multiple types of assignments. Time to do this in class will affect "coverage."

Theory and Research

What is assessment? Simply, assessment is data collection with a purpose. In each of our courses, we engage in the process of *gathering data* about our students' learning. The type of data we gather depends on the evidence we will accept that students have learned what we want them to learn. Generally, the data we collect are intended to be *measures of students' knowledge, attitudes and performance*. Ideally, these data are also matched or "aligned" with the goals of the course and our daily or weekly course activities.

Four functions of assessment data are described by Hodson (1992):

1. *formative assessment* provides diagnostic feedback to students and instructors at short-term intervals (e.g., during a class or on a weekly basis);
2. *summative assessment* provides a description of students' level of attainment upon completion of an activity, module, or course;
3. *evaluative assessment* provides instructors with curricular feedback (e.g., the value of a field trip or a writing assignment).
4. *educative assessment* develops in students and instructors further insight and understanding about their own learning and teaching. In effect, assessment IS a form of learning (NRC 1996, p. 76).

We need confidence in the quality of the data we gather about our students if we want to justify our subsequent decisions about teaching. Many of us who teach introductory science courses are dissatisfied with the type of evidence we are collecting about our students' learning. We admit that data from multiple choice tests measure inert bits of knowledge and some comprehension but provide us incomplete and inadequate feedback about our students' learning. We would like to use alternative forms of assessment to gather multiple, substantive forms of data about active student learning, such as understanding, analysis, reasoning, and synthesis (Ebert-May et al 1997). These kinds of assessments include short answer items, essays, minute papers, oral communication, poster presentations, laboratory projects and research papers, but because of large class sizes and individual research priorities we have limited time to evaluate extended responses from students.

Assessment is learning. We and our students both benefit from meaningful assessment information about the achievement of the broader course goals. Multiple assessment strategies can be implemented to provide evidence that students have or have not learned, have or have not accomplished the goals of the course. Rubrics help us set well-defined standards for our students, provide students guidelines for achieving those standards and facilitate grading extended written and oral responses. This feedback provides us data to interpret and make informed decisions about our students' learning and our own teaching practice, similar to the process of data evaluation that we use daily in our scientific research.

Links

- Diane Ebert-May, Director, Lyman Briggs School and Professor, Botany and Plant Pathology, Michigan State University, East Lansing, Michigan.
  ebertmay@pilot.msu.edu

Sources

- Ebert-May D, Brewer C, Allred S. 1997. Innovation in large lectures-teachings for active learning. Bioscience 47: 601-607.
- Freeman, RHL. 1994. Open-ended questioning: a handbook for educators. Menlo Park, California; Reading, Massachusetts; New York; Don Mills, Ontario; Wokingham, England; Amsterdam; Bonn; Sydney; Singapore; Tokyo; Madrid; San Juan; Paris; Seoul; Milan; Mexico City; Taipei. Addison-Wesley Publishing Company (The book is published by Innovative Learning™, an imprint of the Addison-Wesley Alternative Publishing Group.)
- King PM, Kitchener KS. 1994. Developing reflective judgement: understanding and promoting intellectual growth and critical thinking in adolescents and adults. San Francisco (CA): Jossey-Bass Publishers.
- MacGregor J. 1993. Student self-evaluation: fostering reflective learning. San Francisco (CA): Jossey-Bass Publishers.
- Magolda, MBB. 1992. Knowing and reasoning in college: gender-related students' intellectual development. San Francisco (CA): Jossey-Bass Publishers.
- National Research Council. 1996. National science education standards. Washington (DC): National Academy Press.
- Novak JD, Gowin DB. 1984. Learning how to learn. New York: Cambridge University Press.

Diane Ebert-May

Freshman in college - my favorites. Why? Freshman are excited, energetic, and a bit wary about the challenge before them. Perhaps the subliminal reason is that every year my freshmen are 18 years old, so that must mean I am staying the same age too, right??

Anyhow, when I began teaching a large introductory biology course (600 students) I knew that my multiple choice tests were not providing me the kinds of data I wanted about my students' thinking, because I also knew that freshman can/do think!! Second, I believed that my students needed to learn how to write and speak to explain themselves in the sciences as well as every other facet of their education, and it was my responsibility to assist all of them in this process. On the other hand, I needed a reality check. How would I find time to evaluate 600 writing samples, especially if I asked students to practice writing/speaking more than once throughout the semester?

So I stumbled upon the term "rubric," I learned what it meant, and I learned how to design rubrics from various sources in the literature - a special acknowledgement to BSCS (Biological Sciences Curriculum Studies) whose rubrics influenced my original thinking. As I developed rubrics for each of my assessments, I forced myself to think more explicitly about the goals I wanted my students to achieve and the criteria I would use to monitor their progress. Then I worked with my students to understand and practice achieving the goals and criteria with rubrics as a guide for communication. I now can manage reading and evaluating large numbers of well-written and reasoned responses. We all won. On another note, because my students understand and use rubrics, I seldom, if ever, have individuals who "argue for points" on any given assignment.

Last year, I was recruited by Michigan State University to become the director of a residential science school within the College of Natural Science, the Lyman Briggs School. Our faculty includes scientists, mathematicians and humanists who value their teaching as much as their research. So I left the sunshine of northern Arizona to engage in this great opportunity to continue my research in undergraduate science education, this time with science majors in a small college within a large university -- stay tuned.