

Scoring Rubrics: What, When and How?

Barbara M. Moskal

Associate Director of the Center for Engineering Education
Assistant Professor of Mathematical and Computer Sciences
Colorado School of Mines

Scoring rubrics have become a common method for evaluating student work in both the K-12 and the college classrooms. The purpose of this paper is to describe the different types of scoring rubrics, explain why scoring rubrics are useful and provide a process for developing scoring rubrics. This paper concludes with a description of resources that contain examples of the different types of scoring rubrics and further guidance in the development process.

What is a scoring rubric?

Scoring rubrics are descriptive scoring schemes that are developed by teachers or other evaluators to guide the analysis of the products or processes of students' efforts (Brookhart, 1999). Scoring rubrics are typically employed when a judgement of quality is required and may be used to evaluate a broad range of subjects and activities. One common use of scoring rubrics is to guide the evaluation of writing samples. Judgements concerning the quality of a given writing sample may vary depending upon the criteria established by the individual evaluator. One evaluator may heavily weigh the evaluation process upon the linguistic structure, while another evaluator may be more interested in the persuasiveness of the argument. A high quality essay is likely to have a combination of these and other factors. By developing a pre-defined scheme for the evaluation process, the subjectivity involved in evaluating an essay becomes more objective.

Figure 1 displays a scoring rubric that was developed to guide the evaluation of student writing samples in a college classroom (based loosely on Leydens & Thompson, 1997). This is an example of a holistic scoring rubric with four score levels. Holistic rubrics will be discussed in detail later in this document. As the example illustrates, each score category describes the characteristics of a response that would receive the respective score. By having a description of the characteristics of responses within each score category, the likelihood that two independent evaluators would assign the same score to a given response is increased. This concept of examining the extent to which two independent evaluators assign the same score to a given response is referred to as "rater reliability."

Figure 1.

Example of a scoring rubric designed to evaluate college writing samples.

-3-

Meets Expectations for a first Draft of a Professional Report

- The document can be easily followed. A combination of the following are apparent in the document:
 1. Effective transitions are used throughout,
 2. A professional format is used,
 3. The graphics are descriptive and clearly support the document's purpose.
- The document is clear and concise and appropriate grammar is used throughout.

-2-

Adequate

- The document can be easily followed. A combination of the following are apparent in the document:
 1. Basic transitions are used,
 2. A structured format is used,
 3. Some supporting graphics are provided, but are not clearly explained.
- The document contains minimal distractions that appear in a combination of the following forms:
 1. Flow in thought
 2. Graphical presentations
 3. Grammar/mechanics

-1-

Needs Improvement

- Organization of document is difficult to follow due to a combination of following:
 1. Inadequate transitions
 2. Rambling format
 3. Insufficient or irrelevant information
 4. Ambiguous graphics
- The document contains numerous distractions that appear in the a combination of the following forms:
 1. Flow in thought
 2. Graphical presentations
 3. Grammar/mechanics

-0-

Inadequate

- There appears to be no organization of the document's contents.
- Sentences are difficult to read and understand.

When are scoring rubrics an appropriate evaluation technique?

Writing samples are just one example of performances that may be evaluated using scoring rubrics. Scoring rubrics have also been used to evaluate group activities, extended projects and oral

presentations (e.g., Chicago Public Schools, 1999; Danielson, 1997a; 1997b; Schrock, 2000; Moskal, 2000). They are equally appropriate to the English, Mathematics and Science classrooms (e.g., Chicago Public Schools, 1999; State of Colorado, 1999; Danielson, 1997a; 1997b; Danielson & Marquez, 1998; Schrock, 2000). Both pre-college and college instructors use scoring rubrics for classroom evaluation purposes (e.g., State of Colorado, 1999; Schrock, 2000; Moskal, 2000; Knecht, Moskal & Pavelich, 2000). Where and when a scoring rubric is used does not depend on the grade level or subject, but rather on the purpose of the assessment.

Scoring rubrics are one of many alternatives available for evaluating student work. For example, checklists may be used rather than scoring rubrics in the evaluation of writing samples. Checklists are an appropriate choice for evaluation when the information that is sought is limited to the determination of whether specific criteria have been met. Scoring rubrics are based on descriptive scales and support the evaluation of the extent to which criteria has been met.

The assignment of numerical weights to sub-skills within a process is another evaluation technique that may be used to determine the extent to which given criteria has been met. Numerical values, however, do not provide students with an indication as to how to improve their performance. A student who receives a "70" out of "100", may not know how to improve his or her performance on the next assignment. Scoring rubrics respond to this concern by providing descriptions at each level as to what is expected. These descriptions assist the students in understanding why they received the score that they did and what they need to do to improve their future performances.

Whether a scoring rubric is an appropriate evaluation technique is dependent upon the purpose of the assessment. Scoring rubrics provide at least two benefits in the evaluation process. First, they support the examination of the extent to which the specified criteria has been reached. Second, they provide feedback to students concerning how to improve their performances. If these benefits are consistent with the purpose of the assessment, than a scoring rubric is likely to be an appropriate evaluation technique.

What are the different types of scoring rubrics?

Several different types of scoring rubrics are available. Which variation of the scoring rubric should be used in a given evaluation is also dependent upon the purpose of the evaluation. This section describes the differences between analytic and holistic scoring rubrics and between task specific and general scoring rubrics.

Analytic verses Holistic

In the initial phases of developing a scoring rubric, the evaluator needs to determine what will be the evaluation criteria. For example, two factors that may be considered in the evaluation of a writing sample are whether appropriate grammar is used and the extent to which the given argument is persuasive. An analytic scoring rubric, much like the checklist, allows for the separate evaluation of each of these factors. Each criterion is scored on a different descriptive scale (Brookhart, 1999).

The rubric that is displayed in Figure 1 could be extended to include a separate set of criteria for the evaluation of the persuasiveness of the argument. This extension would result in an analytic scoring rubric with two factors, quality of written expression and persuasiveness of the argument. Each factor would receive a separate score. Occasionally, numerical weights are assigned to the evaluation of each criterion. As discussed earlier, the benefit of using a scoring rubric rather than weighted scores is that scoring rubrics provide a description of what is expected at each score level. Students may use this information to improve their future performance.

Occasionally, it is not possible to separate an evaluation into independent factors. When there is an overlap between the criteria set for the evaluation of the different factors, a holistic scoring rubric

may be preferable to an analytic scoring rubric. In a holistic scoring rubric, the criteria is considered in combination on a single descriptive scale (Brookhart, 1999). Holistic scoring rubrics support broader judgements concerning the quality of the process or the product.

Selecting to use an analytic scoring rubric does not eliminate the possibility of a holistic factor. A holistic judgement may be built into an analytic scoring rubric as one of the score categories. One difficulty with this approach is that overlap between the criteria that is set for the holistic judgement and the other evaluated factors cannot be avoided. When one of the purposes of the evaluation is to assign a grade, this overlap should be carefully considered and controlled. The evaluator should determine whether the overlap is resulting in certain criteria are being weighted more than was originally intended. In other words, the evaluator needs to be careful that the student is not unintentionally severely penalized for a given mistake.

General versus Task Specific

Scoring rubrics may be designed for the evaluation of a specific task or the evaluation of a broader category of tasks. If the purpose of a given course is to develop a student's oral communication skills, a general scoring rubric may be developed and used to evaluate each of the oral presentations given by that student. This approach would allow the students to use the feedback that they acquired from the last presentation to improve their performance on the next presentation.

If each oral presentation focuses upon a different historical event and the purpose of the assessment is to evaluate the students' knowledge of the given event, a general scoring rubric for evaluating a sequence of presentations may not be adequate. Historical events differ in both influencing factors and outcomes. In order to evaluate the students' factual and conceptual knowledge of these events, it may be necessary to develop separate scoring rubrics for each presentation. A "Task Specific" scoring rubric is designed to evaluate student performances on a single assessment event.

Scoring rubrics may be designed to contain both general and task specific components. If the purpose of a presentation is to evaluate students' oral presentation skills and their knowledge of the historical event that is being discussed, an analytic rubric could be used that contains both a general component and a task specific component. The oral component of the rubric may consist of a general set of criteria developed for the evaluation of oral presentations; the task specific component of the rubric may contain a set of criteria developed with the specific historical event in mind.

How are scoring rubrics developed?

The first step in developing a scoring rubric is to clearly identify the qualities that need to be displayed in a student's work to demonstrate proficient performance (Brookhart, 1999). The identified qualities will form the top level or levels of scoring criteria for the scoring rubric. The decision can then be made as to whether the information that is desired from the evaluation can best be acquired through the use of an analytic or holistic scoring rubric. If an analytic scoring rubric is created, then each criterion is considered separately as the descriptions of the different score levels are developed. This process results in separate descriptive scoring schemes for each evaluation factor. For holistic scoring rubrics, the collection of criteria is considered throughout the construction of each level of the scoring rubric and the result is a single descriptive scoring scheme.

After defining the criteria for the top level of performance, the evaluator's attention may be turned to defining the criteria for lowest level of performance. What type of performance would suggest a very limited understanding of the concepts that are being assessed? The contrast between the criteria for top level performance and bottom level performance is likely to suggest appropriate criteria for middle level of performance. This approach would result in three score levels.

If greater distinctions are desired, then comparisons can be made between the criteria for each

existing score level. The contrast between levels is likely to suggest criteria that may be used to create score levels that fall between the existing score levels. This comparison process can be used until the desired number of score levels is reached or until no further distinctions can be made. If meaningful distinctions between the score categories cannot be made, then additional score categories should not be created (Brookhart, 1999). It is better to have a few meaningful score categories than to have many score categories that are difficult or impossible to distinguish.

Each score category should be defined using descriptions of the work rather than judgements about the work (Brookhart, 1999). For example, "Student's mathematical calculations contain no errors," is preferable over, "Student's calculations are good." The phrase "are good" requires the evaluator to make a judgement whereas the phrase "no errors" is quantifiable. In order to determine whether a rubric provides adequate descriptions, another teacher may be asked to use the scoring rubric to evaluate a sub-set of student responses. Differences between the scores assigned by the original rubric developer and the second scorer will suggest how the rubric may be further clarified.

Resources

Currently, there is a broad range of resources available to teachers who wish to use scoring rubrics in their classrooms. These resources differ both in the subject that they cover and the level that they are designed to assess. The examples provided below are only a small sample of the information that is available.

For K-12 teachers, the State of Colorado (1998) has developed an on-line set of general, holistic scoring rubrics that are designed for the evaluation of various writing assessments. The Chicago Public Schools (1999) maintain an extensive electronic list of analytic and holistic scoring rubrics that span the broad array of subjects represented throughout K-12 education. For mathematics teachers, Danielson has developed a collection of reference books that contain scoring rubrics that are appropriate to the elementary, middle school and high school mathematics classrooms (1997a, 1997b; Danielson & Marquez, 1998).

Resources are also available to assist college instructors who are interested in developing and using scoring rubrics in their classrooms. *Kathy Schrock's Guide for Educators* (2000) contains electronic materials for both the pre-college and the college classroom. In *The Art and Science of Classroom Assessment: The Missing Part of Pedagogy*, Brookhart (1999) provides a brief, but comprehensive review of the literature on assessment in the college classroom. This includes a description of scoring rubrics and why their use is increasing in the college classroom. Moskal (1999) has developed a web site that contains links to a variety of college assessment resources, including scoring rubrics.

The resources described above represent only a fraction of those that are available. The ERIC Clearinghouse on Assessment and Evaluation [ERIC/AE] provides several additional useful web sites. One of these, *Scoring Rubrics - Definitions & Constructions* (2000b), specifically addresses questions that are frequently asked with regard to scoring rubrics. This site also provides electronic links to web resources and bibliographic references to books and articles that discuss scoring rubrics. For more recent developments within assessment and evaluation, a search can be completed on the abstracts of papers that will soon be available through ERIC/AE (2000a). This site also contains a direct link to ERIC/AE abstracts that are specific to scoring rubrics.

Search engines that are available on the web may be used to locate additional electronic resources. When using this approach, the search criteria should be as specific as possible. Generic searches that use the terms "rubrics" or "scoring rubrics" will yield a large volume of references. When seeking information on scoring rubrics from the web, it is advisable to use an advanced search and specify the grade level, subject area and topic of interest. If more resources are desired than result from this conservative approach, the search criteria can be expanded.

References

- Brookhart, S. M. (1999). *The Art and Science of Classroom Assessment: The Missing Part of Pedagogy*. ASHE-ERIC Higher Education Report (Vol. 27, No.1). Washington, DC: The George Washington University, Graduate School of Education and Human Development.
- Chicago Public Schools (1999). *Rubric Bank*. [Available online at:http://intranet.cps.k12.il.us/Assessments/Ideas_and_Rubrics/Rubric_Bank/rubric_bank.html].
- Danielson, C. (1997a). *A Collection of Performance Tasks and Rubrics: Middle School Mathematics*. Larchmont, NY: Eye on Education Inc.
- Danielson, C. (1997b). *A Collection of Performance Tasks and Rubrics: Upper Elementary School Mathematics*. Larchmont, NY: Eye on Education Inc.
- Danielson, C. & Marquez, E. (1998). *A Collection of Performance Tasks and Rubrics: High School Mathematics*. Larchmont, NY: Eye on Education Inc.
- ERIC/AE (2000a). *Search ERIC/AE draft abstracts*. [Available online at:<http://ericae.net/sinprog.htm>].
- ERIC/AE (2000b). *Scoring Rubrics - Definitions & Construction* [Available online at:http://ericae.net/faqs/rubrics/scoring_rubrics.htm].
- Knecht, R., Moskal, B. & Pavelich, M. (2000). *The Design Report Rubric: Measuring and Tracking Growth through Success*, Paper to be presented at the annual meeting of the American Society for Engineering Education.
- Leydens, J. & Thompson, D. (August, 1997), *Writing Rubrics Design (EPICS) I*, Internal Communication, Design (EPICS) Program, Colorado School of Mines.
- Moskal, B. (2000). *Assessment Resource Page*. [Available online at:<http://www.mines.edu/Academic/assess/Resource.htm>].
- Schrock, K. (2000). *Kathy Schrock's Guide for Educators*. [Available online at:<http://school.discovery.com/schrockguide/assess.html>].
- State of Colorado (1998). *The Rubric*. [Available online at:<http://www.cde.state.co.us/cdedepcom/asrubric.htm#writing>].