Best Practices in Rubric Creation and Use for Classroom and Accreditation

Amanda K. Gustafson[a] and Robert J. Gustafson[b]
[a]Minneapolis, Minnesota 55418
[b]Ohio State University, Columbus, Ohio 43210
Email: gustafson.4@osu.edu

Abstract

Rubrics are a structured method of grading against criteria. Rubrics have the potential to improve the clarity of instruction, ease of grading, reliability and validity of measure, and usefulness of feedback to students in a broad range of courses including engineering and engineering technology. Well constructed rubrics can provide strong assessment evidence of student learning, and be explicitly matched to student learning outcomes for accreditation. This paper will provide guidance on constructing rubrics as they might be applied in engineering, and highlight some best practices for using them as a tool for student learning and assessment.

Introduction

Rubrics are a widely used tool in educational assessment in both K-12 and colleges[1,2,3] and are being built into current course management systems[4]. Although they can take several different forms, rubrics are pre-defined guidelines that contain well-defined criteria and systematic approaches to scoring and feedback. Rubrics can created advantages for both students and instructors, and may also be designed for research and administrative purposes. Rubrics promote 1) advance planning, 2) communications, 3) consistency and fairness, and 4) objectivity. Rubrics can be used for either summative or formative assessment and by careful design a combination of the two. Within engineering, rubrics have been shown to be particularly useful in evaluation of design projects and for assessment related to accreditation.

Advanced Planning

Rubrics allow/require the instructor, instructional team or curriculum developers to, in advance, make expectations explicit. Rubrics represent the expectations for achievement by making explicit: 1) the main traits of an assignment, a course, or a whole curriculum and 2) level of performance expected.

Communications

Students- Rubrics make the expectations public to student. Sharing in advance helps instructors communicate priorities to students, thereby allowing students understand expectations for the work and criteria for success. Sharing results of the evaluation with the student via a rubric form often brings much more clarity to the communication by having structured and organized feedback.

Instructor - The process of development of a rubric can assist the instructor or curriculum developer in clarifying instructional objectives, success criteria for the objectives, and when used for summative purposes weighting across various elements. The rubric then may be used to communicate not only to students but to others involved in implementing or evaluating.

Grader/Evaluator - Rubrics can be used as a tool to communicate to graders or evaluators who may be supporting the instructor. They can be an effective communication tool when working with multiple graders or evaluators, in particular the untrained, as one may often have with visitors or experts evaluating student work like design projects, and oral and written reports. Use of rubrics can dramatically reduce the time, and effort needed to give productive feedback to the student.
Program Evaluators - Results of evaluation using rubrics can be effectively used as evidence for assessment related to program administration and accreditation.

Consistency and Fairness
Whether it be across a set of assignments within a single class, across multiple offerings of the same course, or across a program, rubrics can promote consistency and fairness by being explicit about expectations and criteria for levels of success. When multiple graders are evaluating, rubrics can create uniformity and consistency across while still allowing for individual feedback and expression. For a single evaluator, it can support consistency across time or multiple evaluations such as in exam grading.

Objectivity
Along with consistency and fairness, rubrics structure assessment to be objective by basing judgments on factual observation. A valid rubric measures what it is intended to measure and increases the objectivity and reliability of scoring.[5]

Formative and Summative Assessment

In developing a rubric if must be clear if the objective is formative or summative assessment. *Formative* assessments provide information to both the instructor and students concerning students' progress to guide the learning process. It often gives guidance to the student or teacher that can be used for performance improvement. A key to formative assessment is feedback which allows students to correct errors and encourages instructors to modify instructional activities. Since formative assessments are designed to guide learning and are not generally utilized as an outcome measure, they are generally considered a low stakes assessment.[6] *Summative* assessments are designed to focus on identifying the level of student mastery and the effectiveness of instruction. As such, summative assessments are outcome measures that emphasize student achievement rather than aptitude or effort. From a student perspective, summative assessments are primarily utilized to determine final course grades; from an instructor and program evaluation perspective, they are a means of accountability.[7] One form of accountability may be their use in documentation for accreditation purposes.

Formative and summative evaluation can be seen to be in conflict as they are designed for different primary objectives. Summative may often constrain formative by demanding relative weighting of various items. Formative may not allow for weighting across various performance measures as desired for grading. However, rubrics can be designed in a way to consistently give weighting across items will still allowing for within item formative feedback. This requires careful pre-planning and may result in a multi-dimensional rubric.

Rubric Shapes and Characteristics with Examples

Rubrics can come in many shapes. A common shape is a list of statements with a number range indicating how well the statement was met. This form has the advantage of describing an ideal result. The disadvantage is that it needs written comments by the grader to clarify how the student can move improve.

Language is clear and professional, with no errors.									
1	2	3	4	5	6	7	8	9	10
Comments:									

A rare format is a pass/fail list of requirements. This format is challenging for assigning grades; it raises questions of whether or how many items it is possible to fail and pass the assignment. One of the authors has seen this form used to assess competency in a medical procedure skill that required exact safety measured to be followed and had a specific order to the steps. Missing or performing a step poorly was grounds for failing, and retaking the skill test. This level of skill practice is rarely needed in engineering, where there is generally more flexibility in the order of operations in a skill and the flexibility to correct work if mistakes are initially made. However, a low stakes version may be worth considering for drilling lab safety practices.

Student turns off circuit breaker. y/n
Student uses voltmeter to check for current. y/n

A format which requires more initial work, but reduces the amount of written commenting needed, is a table format. For this type of rubric, assignment objectives or criteria are listed in the first column. The first row lists the value for each level of success. This can be an adjective, a point level, a Likert type scale or a letter value. The boxes beneath the value describe the characteristics of the objective which will receive that value.

5	**Engineering Analysis:** ***Skillfully synthesizes the results of modeling, simulation, and prototyping to refine the design and/or reformulate the problem.***					
0	1	2	3	4	5	6
ABSENT	EMERGING		COMPETENT		MASTERING	
Dimension 5 *Score*	Modeling, simulation, and prototyping are poorly handled and quantitative and/or symbolic tools have frequently been misapplied. Calculations are frequently inaccurate with little evidence that the underlying mathematical concepts have been grasped. Analysis is inadequate for the target design.		Modeling, simulation, and prototyping have been adequately handled but quantitative and/or symbolic tools may sometimes have been applied incorrectly. Calculations are mostly accurate but may display some misunderstanding about underlying mathematical concepts. Analysis is adequate but does not go beyond routine ideas for refinement or problem reformulation.		Modeling, simulation, and prototyping have been skillfully handled and demonstrate the appropriate application of quantitative and/or symbolic tools. Calculations are complete, accurate, self-generated and show a deep understanding of the mathematical concepts. Careful analysis leads to insightful problem refinement or problem reformulation.	

Sample Rubric[8]

This form of rubric is strong for both formative and summative assessment. Describing the qualities of emerging, competent, and mastering- using the sample's adjectives- allows the benefits of communication described earlier in the paper. Students, faculty, graders and program evaluators have common language from which to begin a discussion of quality.

Selecting criteria

When choosing grading criteria, it is helpful to consider two categories- form and content. Form refers to the structure of the assignment. Academic language and citation in a paper, for example, are form criteria as is the verbal qualities of an oral presentation. University writing and teaching centers can be excellent resources for choosing and describing form criteria. Content refers to the knowledge or skills demonstrated within the assessment. Generally, content criteria are given more weight than form unless learning the form is the main purpose of the assignment, such as learning to write an academic paper in a writing class.

When choosing either type of criteria, it is best to start with a careful read of the assessment description. Then asking probing questions like "What does the student need to demonstrate to show mastery of the material?". When adding a rubric to an existing assignment, it can be helpful to have samples of previous student work of both high and low quality. One can then ask questions like "What distinguishes a strong submission from a weaker one?". Writing the differences with as much detail as possible can assist in determining the range later.

Threshold criteria

Threshold criteria come into play when an assessment is used for the purpose of assigning a grade. Occasionally, there are criteria which must be met at a certain level for the overall assessment to be considered passing. For example, in a research paper, if an excellent review of sources is given, but the section showing analysis or original thought is missing the paper would not likely receive a passing grade. It may be appropriate to state in the scoring instructions that certain criteria need to be met at specific levels in order to receive a passing score, or that a certain number of "missing" scores will result in failure.

Overlapping criteria

Ideally, each criterion represents a largely independent variable of the assessment. In order not to accidentally overweight a variable, it is helpful to ask "Could a student do poorly on this criterion and well on others?" If the answer is no, it might be worthwhile to consider combining the overlapping criteria, or reviewing their definitions. Some instructors initially create an "overall impression" criterion as well as a specific breakdown. This is often removed after review when the faculty member realizes that the same grade results from a summary of the specific criteria as the more broad initial impression.

Determining range

Range. The larger the number of distinctions, the more cognitively difficult the task. Large ranges increase the likelihood of inter- and intra- rater variability, as well as causing challenges in interpretation for students. Generally, 2 to 4 categories, with an option for "not shown" balance the needs for separating work by quality without becoming overwhelmed by options.

The third sample rubric cleverly creates 6 options by starting with three broad descriptors, and then allows a "higher" or "lower" distinction within the category. The use of a shadow "absent" rating is also cognitively simple and clear in meaning.

Converting Rubrics to Grades

One common difficulty is converting descriptive rubrics, those describing levels of success as adjectives or Likert values, into grades that fit in a standard grade book of points or percentages. For example, a Likert score of 2 on a scale of 0-4, if directly changed to a percentage, would be 50%, implying failure, where the rubric description of 2 might indicate an adequate grasp of that criterion. Going from a subjective "A" could translate as anything from 100-91 on the conventional percentage scale. If an assignment has 5 criteria, and a student receives 2 "Excellent", 1 "Good", 1 "Adequate" and 1 "Not Adequate" what can be concluded about the overall quality of the work? Unfortunately, there is no ideal solution for this difficulty. The state of art in assessment has not yet created general practices equivalent to the common conversions of 90-100% ≈ 3.01-4.0 G.P.A ≈ "A."

Example with Score for Grading (Some Rows Omitted)[9]

Oral Presentations						COMMENTS	SCORE
Oral presentation	**5**	**4**	**3**	**2**	**1**		
	• Speaks directly to audience • Well-rehearsed • Keeps audience's attention • Smooth transitions between speakers						
Information Flow	**5**	**4**	**3**	**2**	**1**		
	• Strong intro (good "hook") • Good structure and logical flow of information • Good conclusion • Discussion of next steps						
Question Response	**5**	**4**	**3**	**2**	**1**		
	• Several team members respond to questions • Understand and repeats question for clarification • Give logical and complete answers • Does not require follow-up, clarifying question						
						TOTAL SCORE	**0**

Some common workarounds to this conversion issue exist. One option is to use 6-10 in place of 0-4 as criteria values. For assessments with a range of three, this occasionally is seen grouped as >6, 6, 7-8, 9-10. Another option is to use 0-4 and the institutions published method of converting GPA scores to letter grades. Ad hoc solutions are also possible.

Often it is desirable to give some criteria more weight than others, for example criteria requiring critical thinking count more towards the grade than criteria that are more rote memorization. One option is to place a percentage multiplier on each criterion. Another option is to double (or more) a criterion in comparison to others. Like the conversion issue, there is no current standard method. The most important factor in both cases is the clarity with which the standard is communicated to all stakeholders.

Using Rubrics for Student Learning

If used effectively, rubrics can provide feedback to improve student learning. Focused practice, followed by focused feedback, is a key to learning.[10] One best practice is to hand out the rubric at the same time as the assignment is given. Students can be encouraged or required to provide feedback to peers using the rubric. The rubric can be used to focus feedback on early drafts, reviewed by peers, teaching assistants and faculty.

Maintaining a consistent rubric across multiple assignments and, ideally, multiple courses can increase consistency, fairness, and provide evidence of learning across the program. One example might be using the same form criteria, such as use of language and citation, on every paper assigned, while changing the content criteria to match the specific assignment. Another example is using the same assessment of team participation in all classes with group work. The greater the consistency of rubrics, the less time students need to spend learning the grading mechanics of the course and the more energy they can place on the course material.

Using Rubrics for Program Analysis

For general education learning outcomes, departments may wish to consider "adopting and adapting"[11] the Lumina Foundation's Degree Qualifications Profile (DQP).[12] The DQP provides descriptions of "Broad, Integrative Knowledge" and "Intellectual Skills" which are written in observable, measurable language and which may parallel specific institutional desired outcomes.

Rubrics can be used in quantitative documentation of outcomes for engineering and engineering technology programs as required in ABET[13] continuous quality improvement plans. These can be particularly helpful when trying to standardize the reporting of assessment done by persons outside the program.

Conclusions

Well constructed rubrics require advanced planning, which is rewarded through improved communication, fairness, accuracy and ability to provide student and program feedback. While open questions remain about how to translate rubric scores to grades, best practices in their creation and use are available. Rubrics are a teaching tool, and used with skill can support both student learning and program evaluation.

Bibliography

[1]Bean, J. C. 2001. "Developing and Applying Grading Criteria." In *Engaging Ideas: The Professor's Guide to Integrating Writing, Critical Thinking, and Active Learning in the Classroom*. San Francisco: Jossey-Bass.

[2]Wiggins, g. and J. McTighe. 2005 (2nd Ed.). Understanding by Design. Association for Supervision and Curriculum Development, Alexandria, VA.

[3]Moskal, Barbara M. (2000). Scoring rubrics: what, when and how?. *Practical Assessment, Research & Evaluation*, 7(3). http://PAREonline.net/getvn.asp?v=7&n=3 (Accessed 4/3/13)

4Ondemand Learning Center, Blackboard Inc. http://ondemand.blackboard.com/r91/movies/bb91_course_tools_grade_with_rubrics.htm (Accessed 3/4/13)

[5] University of Michigan. "Rubric Scoring and Grading", Assessment for Curricular Improvement. http://www.engin.umich.edu/teaching/assess_and_improve/handbook/direct/rubric.html (Accessed 1/4/13)

[6] Park University. "Formative Assessment in the Classroom", Faculty Resource Quick Tips. http://www.park.edu/cetl/quicktips/formative.html (Accessed 1/4/13).

[7] Park University, "Summative Classroom Assessment", Faculty Resource Quick Tips. http://www.park.edu/cetl/quicktips/summative.html (Accessed 1/4/13).

[8] Washington State University, EECS Capstone Rubric

https://mysite.wsu.edu/personal/pedrow/eecscapstone/fall2009/Rubrics/Engineering%20Design%20Rubric%202009.pdf (accessed 1/20/13)

[9] Personal Communications. Multi-disciplinary Capstone Design Program, Engineering Education Innovation Center, Ohio State University, Columbus, OH.

[10] Ambrose, S., Bridges, M., DiPietro, M., Lovett, M., and Norman, M. (2010). *How Learning Works: 7 Research-Based Principles for Smart Teachin*g. p.136. San Francisco: Jossey-Bass.

[11] Personal Communication. Dr. Paul Gaston Lumina Foundation Consultant, Trustee Professor Kent State University (Dad, I don't know how to cite him)

[12] Lumina Foundation. (2011). *The Degree Qualifications Profile*. http://www.luminafoundation.org/publications/The_Degree_Qualifications_Profile.pdf

[13] ABET. Accreditation Board for Engineering and Technology, Accreditation-criteria-policies-documents. http://abet.org/accreditation-criteria-policies-documents/ (Accessed 2/1/13)