Writing Effective Rubrics

Office of Institutional Assessment
Dr. Timothy S. Brophy
Original 2019; Updated 2025

Office of the Provost
University of Florida

Continuous Quality
Enhancement Series

UF | UNIVERSITY of FLORIDA

What is a Rubric?

A rubric is a measurement tool that describes the criteria against which a performance, behavior, or product is compared and measured. Rubrics list the criteria established for a particular task and the levels of achievement associated with each criterion. These are often developed in the form of a matrix.

Why use a Rubric?

Here are some primary reasons to use rubrics (Hawaii, 2025).

- A rubric creates a common framework and language for assessment.
- Complex products or behaviors can be examined efficiently.
- Well-trained reviewers apply the same criteria and standards.
- Rubrics are criterion-referenced, rather than norm-referenced. Raters ask, "Did the student meet the criteria for level 5 of the rubric?" rather than "How well did this student do compared to other students?"
- Using rubrics can lead to substantive conversations among faculty.
- When faculty members collaborate to develop a rubric, it promotes shared expectations and grading practices.

Faculty members often use rubrics for assessment. Here are two examples.

The University Writing Program collected essays from students in all sections of ENC1101. A team of faculty members evaluated the essays by applying an analytic scoring rubric. Before applying the rubric, they calibrated the rubric by agreeing on how to apply the rubric by scoring the same set of essays and discussing them until they reached consensus.

Biology laboratory instructors agreed to use a "Biology Lab Report Rubric" to grade students' lab reports in all Biology lab sections. The intended outcome is “Students write biology laboratory reports accurately an appropriately.” At the beginning of each semester, instructors meet and discuss sample lab reports and agreed on how to apply the rubric and their expectations for each level. Every other year, they select a random sample of students' lab reports. A Biology professor scores each of those reports. The score given by the course instructor is compared to the score given by the Biology professor. In addition, the scores are reported as part of the program's assessment data.

Types of Rubrics

Here are the descriptions of analytic and holistic rubrics.

Analytic Rubric: An analytic rubric presents a description of each level of achievement for each criterion and provides a separate score for each criterion. For analytic rubrics, the criteria are usually listed down the left column with the descriptions of the levels of achievement across the rows for each criterion.

- Advantages: provides more detailed feedback on student performance; scoring more consistent across students and raters
- Disadvantages: more time consuming than applying a holistic rubric

Use when:

- You want to see strengths and weaknesses.
- You want detailed feedback about student performance.

Holistic Rubric: A holistic rubric presents a description of each level of achievement and provides a single score based on an overall impression of a student's performance on a task (Carriveau, 2010). For holistic rubrics, the levels of achievement are listed down the first column, and the descriptions of each level of achievement for all criteria are listed in a second column.

- Advantages: quick scoring, provides an overview of student achievement, efficient for large group scoring
- Disadvantages: does not provided detailed information; not diagnostic; may be difficult for scorers to decide on one overall score

Use when:

- You want a quick snapshot of achievement.
- A single dimension is adequate to define quality.

The Parts of a Rubric

Rubrics are composed of four basic parts (Hawaii, 2025). In its simplest form, the rubric includes:

1. A task description. The outcome being assessed or instructions students received for an assignment.

2. The characteristics to be rated (rows). The skills, knowledge, and/or behavior to be demonstrated.

3. Levels of mastery/scale (columns). Labels used to describe the levels of mastery should be tactful but clear. Commonly used labels include:

- Exceeds expectations, meets expectations, near expectations, Below expectations
- Exemplary, proficient, marginal, unacceptable
- Mastery, proficient, developing, novice
- 4, 3, 2, 1

4. The description of each characteristic at each level of mastery/scale (cells).

How to Develop a Rubric

Here are some steps to develop a rubric.

Tip: Generative AI is an excellent tool to develop a rubric. The more specific you can be in your prompt, the better the response.

Step 1: Determine the type of rubric you wish to use – holistic or analytic (Carriveau, 2010).

Step 2: Identify what you want to assess. These form the criteria for the assessment. These are usually part of the description of the assignment or task.

Step 3: Identify the characteristics to be rated (rows).

- Specify the skills, knowledge, and/or behaviors that you will be looking for.
- Limit the characteristics to those that are most important to the assessment.

Step 4: Identify the levels of mastery/scale (columns).

Tip: Aim for an even number (we recommend 4) because when an odd number is used, the middle tends to become the "catch-all" category.

Step 5: Describe each level of mastery for each characteristic and write it in the cells.

- Describe the best work you could expect using these characteristics. This describes the top category.
- Describe an unacceptable product. This describes the lowest category.

- Develop descriptions of intermediate-level products for intermediate categories. Important: Each description and each category should be mutually exclusive.
- Focus your descriptions on the presence of the quantity and quality that you expect, rather than on the absence of them. However, at the lowest level, it would be appropriate to state that an element is “lacking” or “absent” (Carriveau, 2010).
- Keep the elements of the description parallel from performance level to performance level. In other words, if your descriptors include quantity, clarity, and details, make sure that each of these outcome expectations is included in each performance level descriptor.

Step 6: Try out the rubric.

- Apply the rubric to an assignment.
- Share with colleagues.

Tip: Faculty members often find it useful to establish the minimum score needed for the student work to be deemed passable. For example, faculty members may decide that a "1" or "2" on a 4-point scale (4=exemplary, 3=proficient, 2=marginal, 1=unacceptable), does not meet the minimum quality expectations. They may set their criteria for success as 90% of the students must score 3 or higher. If assessment study results fall short, action will need to be taken.

Step 7: Discuss with colleagues. Review feedback and revise.

Tip: When developing a rubric for program assessment, enlist the help of colleagues. Rubrics promote shared expectations and grading practices which benefit faculty members and students in the program.

Rubric Variations

There are two variations of rubrics that can be used successfully, if well calibrated by the users.

- Point system rubrics provide a range of points for each level of achievement; points are given at the scorer’s discretion. Each level receives the same number of points.

- Weighted point system rubrics are a variation of the point system rubric, where different criteria are “weighted” by assigning different point ranges to the criteria.

These rubrics convert levels descriptors into points, which creates scores that are compatible with the score ranges used in common grading scales.

Scoring Rubric Group Orientation and Calibration

When using a rubric for program assessment purposes, faculty members apply the rubric to pieces of student work (e.g., reports, oral presentations, design projects). To produce dependable scores, each faculty member needs to interpret the rubric in the same way. The process of training faculty members to apply the rubric is called "calibration." It's a way to calibrate the faculty members so that scores are accurate and reliable. Reliability here means that the scorers apply the rubric consistently, not only to each piece of student work (called intrarater reliability), but among themselves (called interrater reliability).

Directions for Rubric Calibration

Below are directions for the rubric calibration process (Hawaii, 2025).

Suggested materials for a scoring session:

- Copies of the rubric
- Copies of the "anchors": pieces of student work that illustrate each level of mastery. Suggestion: have 6 anchor pieces (2 low, 2 middle, 2 high)
- Score sheets
- Extra pens, tape, sticky notes, paper clips, stapler, rubber bands, etc. Hold the scoring session in a room that:
- Allows the scorers to spread out as they rate the student pieces
- Has a chalk or white board

Process:

1. Describe the purpose of the activity, stressing how it fits into program assessment plans. Explain that the purpose is to assess the program, not individual students or faculty, and describe ethical guidelines, including respect for confidentiality and privacy.

2. Describe the nature of the products that will be reviewed, briefly summarizing how they were obtained.

3. Describe the scoring rubric and its categories. Explain how it was developed.

4. Analytic: Explain that readers should rate each dimension of an analytic rubric separately, and they should apply the criteria without concern for how often each score (level of mastery) is used. Holistic: Explain that readers should assign the score or level of mastery that best describes the whole piece; some aspects of the piece may not appear in that score and that is okay. They should apply the criteria without concern for how often each score is used.

5. Give each scorer a copy of several student products that are exemplars of different levels of performance. Ask each scorer to independently apply the rubric to each of these products, writing their ratings on a scrap sheet of paper.

6. Once everyone is done, collect everyone's ratings and display them so everyone can see the degree of agreement. Alternatively, the facilitator could ask raters to raise their hands when their rating category is announced, making the extent of agreement very clear to everyone and making it very easy to identify raters who routinely give unusually high or low ratings.

7. Guide the group in a discussion of their ratings, and expect differences. This discussion is important to establish standards. Attempt to reach consensus on the most appropriate rating for each of the products being examined by inviting people who gave different ratings to explain their judgments. Raters should be encouraged to explain by making explicit references to the rubric. Usually consensus is possible, but sometimes a split decision is developed, e.g., the group may agree that a product is a "3-4" split because it has elements of both categories. You might allow the group to revise the rubric to clarify its use but avoid allowing the group to drift away from the rubric and learning outcome(s) being assessed.

8. Once the group is comfortable with how the rubric is applied, the rating begins. Explain how to record ratings using the score sheet and explain the procedures. Reviewers begin scoring.

9. If you can quickly summarize the scores, present a summary to the group at the end of the reading. You might end the meeting with a discussion of five questions:

- Are results sufficiently reliable?

- What do the results mean? Are we satisfied with the extent of students' learning?
- Who needs to know the results?
- What are the implications of the results for curriculum, pedagogy, or student support services?
- How might the assessment process, itself, be improved?

Additional Tips for Developing a Rubric

- Find and adapt an existing rubric! It is rare to find a rubric that is exactly right for your situation, but you can adapt an already existing rubric that has worked well for others and save a great deal of time. A faculty member in your program may already have a good one.
- Evaluate the rubric. Ask yourself:
  - Does the rubric relate to the outcome(s) being assessed?
  - Does it address anything extraneous? (If yes, delete.)
  - Is the rubric useful, feasible, manageable, and practical? (If yes, find multiple ways to use the rubric, such as for program assessment, assignment grading, peer review, student self-assessment)
  - Benchmarking - collect samples of student work that exemplify each point on the scale or level. A rubric will not be meaningful to students or colleagues until the anchors/benchmarks/exemplars are available.
  - Anticipate that you will be revising the rubric.
  - Share effective rubrics with your colleagues.

Deeper Dive

UF generally uses two categories of assessments to inform program effectiveness – direct and indirect. Direct assessments of student learning are those that provide for direct examination or observation of student knowledge or skills against measurable performance indicators. Indirect assessments are those that ascertain the opinion or self-report of the extent or value of learning experiences (Rogers, 2011). Direct assessments are either norm-referenced or criterion-referenced.

Norm-referenced assessments are based on a set of assumptions that permit comparison of one individual's performance to others who have completed the same assessment. This allows interpretations of scores relative to the performance of others. (e.g., "this student has performed above the average"). Norm-referenced

assessments generally consist of dichotomous items – those with one clear, correct answer, such as the selected-response questions that are common on tests, quizzes, and examinations. Generally, in a norm-referenced test the individual test-taker earns a certain number of points for each correct answer, and the scorer totals the number of points earned for the correct answers to create a score. The assumption, then, is that the individual's score represents the individual's knowledge of the subject matter being tested. So, the higher the score, the more knowledge the individual possesses. Based on this assumption, scores can be compared among individuals. For instance, on a test with a score range from 0-100 points, we assume that an individual who scores 96 knows more that an individual who scores 80. This scoring system and its assumptions are familiar to anyone who has ever been in school, and the field of psychometrics emerged to describe the study of these types of assessments.

Criterion-referenced assessments are very different. They are designed to compare a student's performance to a particular standard or criterion. This allows interpretations of scores in relation to the body of knowledge. (e.g., "this student has met the specified performance standard".) Test takers are given a task, and their response - performance, behavior, or a final product – is assessed for the degree to which it meets certain levels of quality. Measurement of these types of assessments is done largely through expert judgment by individuals qualified to review the response, usually a teacher, professor, or other disciplinary expert. The resulting measurements are not intended to be used to compare achievement among those who complete the assessment, but rather the degree to which an individual meets the criteria established for the task. These assessments are often measured using rubrics. These assessments can also use dichotomous items. However, with dichotomous items, a standard of performance is set and scores are interpreted in terms of whether they met the standard or cutoff. For accreditation, criterion-referenced assessments are more likely to be used rather than norm-referenced assessments since we are often measuring whether students have met some performance standard.

Additional Resources and Sample Rubrics

- *Assessment Rubrics,* UC Berkeley Center for Teaching and Learning: https://teaching.berkeley.edu/teaching-guides/assessing-learning/assessment-rubrics

- *Creating and Using Rubrics*, University of Hawai'i at Mānoa, https://manoa.hawaii.edu/assessment/resources/creating-and-using-rubrics
- Gonzalez, J. (2014, May 1). *Know your terms: Holistic, analytic, and single-point rubrics*. https://www.cultofpedagogy.com/holistic-analytic-single-point-rubrics/
- *Grading and Performance Rubrics,* Eberly Center for Teaching Excellence and Educational Innovation, Carnegie Mellon University. https://www.cmu.edu/teaching/designteach/teach/rubrics.html
- Mertler, C.A. (2001). *Designing scoring rubrics for your classroom*. Practical Assessment, Research & Evaluation, 7(25).
- *Rubric Best Practices, Examples, and Templates,* North Carolina State University. https://teaching-resources.delta.ncsu.edu/rubric_best-practices-examples-templates/
- *Rubric Creation and Use,* Center for Innovative Teaching and Learning, Indiana University Bloomington. https://citl.indiana.edu/teaching-resources/assessing-student-learning/rubric-creation-use/index.html
- *Rubric Library*, Institutional Effectiveness, Syracuse University. https://effectiveness.syr.edu/assessment/assessment-resources/rubric-library/
- Stevens, D.D. & Levi, A.J. (2013). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback, and promote student learning* (2nd ed). Available at UF Library.
- *Using Rubrics*, Cornell University Center for Teaching Innovation: https://teaching.cornell.edu/teaching-resources/assessing-student-learning/using-rubrics

Works Cited

Carriveau, R. (2010). *Connecting the dots*. Denton, TX: Fancy Fox Publications, Inc.

Rogers, G. (2011, July 19). *Best practices in assessing student learning. The institute on quality enhancement and accreditation.* Fort Worth, Texas, USA: Southern Association of Colleges and Schools Commission on Colleges.

University of Hawai'i. (2025). *Assessment and Curriculum Support Center*. https://manoa.hawaii.edu/assessment/