

Developing and Applying Rubrics

June 14, 2011 ALA Mary Allen

Rubrics provide the criteria for assessing students' work. They can be used to assess virtually any product or behavior, such as essays, research reports, portfolios, works of art, recitals, oral presentations, performances, and group activities. Judgments can be self-assessments by students; or judgments can be made by others, such as faculty, other students, fieldwork supervisors, and external reviewers. Rubrics can be used to provide formative feedback to students, to grade students, and/or to assess courses and programs.

There are two major types of scoring rubrics:

- Holistic scoring — one global, holistic score for a product or behavior
 - Analytic rubrics — separate, holistic scoring of specified characteristics of a product or behavior
-

Rubric Examples

- Critical Thinking (10, 11, 74, 75)
 - Writing (2, 3, 4, 5, 18, 23, 46, 78)
 - Information Competence (77)
 - GE Social Sciences (19, 44)
 - Fine Arts (20, 31, 62, 64)
 - Natural Sciences (5, 41)
 - Leadership (13)
 - Collaboration (27, 76)
 - Intentional Learning (57)
 - Community Service Learning Outcomes (65)
 - AAC&U Value Rubrics
 - Scripps Rubrics
-

Rubrics have many strengths:

- Complex products or behaviors can be examined efficiently.
 - Developing a rubric helps to precisely define faculty expectations.
 - 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?” This is more compatible with cooperative and collaborative learning environments than competitive grading schemes and is essential when using rubrics for program assessment because you want to learn how well students have met your standards.
 - Ratings can be done by students to assess their own work, or they can be done by others, e.g., peers, fieldwork supervisions, or faculty.
-

Two Common Ways to Assess Learning Outcomes Using Rubrics

1. Assess while grading.
 2. Collect evidence and assess in a group session.
-

Adapting Assessment Rubrics for Assessing and Grading

Here's an assessment rubric—an analytic rubric with three dimensions for assessing oral presentation skills.

| Rubric for Assessing Oral Presentations | | | |
|--|--|---|---|
| | Below Expectation | Satisfactory | Exemplary |
| Organization | No apparent organization. Evidence is not used to support assertions. | The presentation has a focus and provides some evidence which supports conclusions. | The presentation is carefully organized and provides convincing evidence to support conclusions. |
| Content | The content is inaccurate or overly general. Listeners are unlikely to learn anything or may be misled. | The content is generally accurate, but incomplete. Listeners may learn some isolated facts, but they are unlikely to gain new insights about the topic. | The content is accurate and complete. Listeners are likely to gain new insights about the topic. |
| Delivery | The speaker appears anxious and uncomfortable, and reads notes, rather than speaks. Listeners are largely ignored. | The speaker is generally relaxed and comfortable, but too often relies on notes. Listeners are sometimes ignored or misunderstood. | The speaker is relaxed and comfortable, speaks without undue reliance on notes, and interacts effectively with listeners. |

Alternative Format 1.

Points are assigned and used for grading, as shown below, and the categories (Below Expectation, Satisfactory, Exemplary) can be used for assessment. Faculty who share an assessment rubric might:

- assign points in different ways, depending on the nature of their courses
- decide to add more rows for course-specific criteria or comments.

Notice how this rubric allows faculty, who may not be experts on oral presentation skills, to give detailed formative feedback to students. This feedback describes present skills and indicates what students should do to improve. Effective rubrics can help faculty reduce the time they spend grading and eliminate the need to repeatedly write the same comments to multiple students.

| Rubric for Grading Oral Presentations | | | | |
|--|---|---|--|--------------|
| | Below Expectation | Satisfactory | Exemplary | Score |
| Organization | No apparent organization. Evidence is not used to support assertions. (0-4) | The presentation has a focus and provides some evidence which supports conclusions. (5-6) | The presentation is carefully organized and provides convincing evidence to support conclusions. (7-8) | |
| Content | The content is inaccurate or overly general. Listeners are unlikely to learn anything or may be misled. (0-8) | The content is generally accurate, but incomplete. Listeners may learn some isolated facts, but they are unlikely to gain new insights about the topic. (9-11) | The content is accurate and complete. Listeners are likely to gain new insights about the topic. (12-13) | |
| Delivery | The speaker appears anxious and uncomfortable, and reads notes, rather than speaks. Listeners are largely ignored. (0-5) | The speaker is generally relaxed and comfortable, but too often relies on notes. Listeners are sometimes ignored or misunderstood. (6-7) | The speaker is relaxed and comfortable, speaks without undue reliance on notes, and interacts effectively with listeners. (8-9) | |
| Total Score | | | | |

Alternative Format 2.

Weights are used for grading; categories (Below Expectation, Satisfactory, Exemplary) can be used for assessment. Individual faculty determine how to assign weights for their course grading. Faculty may circle or underline material in the cells to emphasize criteria that were particularly important during the assessment/grading, and they may add a section for comments or other grading criteria.

| Rubric for Grading Oral Presentations | | | | |
|--|--|---|---|---------------|
| | Below Expectation | Satisfactory | Exemplary | Weight |
| Organization | No apparent organization. Evidence is not used to support assertions. | The presentation has a focus and provides some evidence which supports conclusions. | The presentation is carefully organized and provides convincing evidence to support conclusions | 30% |
| Content | The content is inaccurate or overly general. Listeners are unlikely to learn anything or may be misled. | The content is generally accurate, but incomplete. Listeners may learn some isolated facts, but they are unlikely to gain new insights about the topic. | The content is accurate and complete. Listeners are likely to gain new insights about the topic. | 50% |
| Delivery | The speaker appears anxious and uncomfortable, and reads notes, rather than speaks. Listeners are largely ignored. | The speaker is generally relaxed and comfortable, but too often relies on notes. Listeners are sometimes ignored or misunderstood. | The speaker is relaxed and comfortable, speaks without undue reliance on notes, and interacts effectively with listeners. | 20% |
| Comments | | | | |

Alternative Format 3.

Some faculty prefer to grade holistically, rather than through assigning numbers. In this example, the faculty member checks off characteristics of the speech and determines the grade based on a holistic judgment. The categories (Below Expectation, Satisfactory, Exemplary) can be used for assessment.

| Rubric for Grading Oral Presentations | | | |
|--|---|---|--|
| | Below Expectation | Satisfactory | Exemplary |
| Organization | <ul style="list-style-type: none"> <input type="checkbox"/> No apparent organization. <input type="checkbox"/> Evidence is not used to support assertions. | <ul style="list-style-type: none"> <input type="checkbox"/> The presentation has a focus. <input type="checkbox"/> Student provides some evidence which supports conclusions. | <ul style="list-style-type: none"> <input type="checkbox"/> The presentation is carefully organized. <input type="checkbox"/> Speaker provides convincing evidence to support conclusions |
| Content | <ul style="list-style-type: none"> <input type="checkbox"/> The content is inaccurate or overly general. <input type="checkbox"/> Listeners are unlikely to learn anything or may be misled. | <ul style="list-style-type: none"> <input type="checkbox"/> The content is generally accurate, but incomplete. <input type="checkbox"/> Listeners may learn some isolated facts, but they are unlikely to gain new insights about the topic. | <ul style="list-style-type: none"> <input type="checkbox"/> The content is accurate and complete. <input type="checkbox"/> Listeners are likely to gain new insights about the topic. |
| Delivery | <ul style="list-style-type: none"> <input type="checkbox"/> The speaker appears anxious and uncomfortable. <input type="checkbox"/> Speaker reads notes, rather than speaks. <input type="checkbox"/> Listeners are largely ignored. | <ul style="list-style-type: none"> <input type="checkbox"/> The speaker is generally relaxed and comfortable. <input type="checkbox"/> Speaker too often relies on notes. <input type="checkbox"/> Listeners are sometimes ignored or misunderstood. | <ul style="list-style-type: none"> <input type="checkbox"/> The speaker is relaxed and comfortable. <input type="checkbox"/> Speaker speaks without undue reliance on notes. <input type="checkbox"/> Speaker interacts effectively with listeners. |
| Comments: | | | |

Alternative Format 4.

Combinations of Various Ideas. As long as the nine assessment cells are used in the same way by all faculty, grading and assessment can be done simultaneously. Additional criteria for grading can be added, as shown below.

| Rubric for Grading Oral Presentations | | | | |
|--|--|--|---|---------------|
| | Below Expectation 1 | Satisfactory 2 | Exemplary 3 | Weight |
| Organization | <ul style="list-style-type: none"> ☐ No apparent organization. ☐ Evidence is not used to support assertions. | <ul style="list-style-type: none"> ☐ The presentation has a focus. ☐ Speaker provides some evidence which supports conclusions. | <ul style="list-style-type: none"> ☐ The presentation is carefully organized. ☐ Speaker provides convincing evidence to support conclusions | 20% |
| Content | <ul style="list-style-type: none"> ☐ The content is inaccurate or overly general. ☐ Listeners are unlikely to learn anything or may be misled. | <ul style="list-style-type: none"> ☐ The content is generally accurate, but incomplete. ☐ Listeners may learn some isolated facts, but they are unlikely to gain new insights about the topic. | <ul style="list-style-type: none"> ☐ The content is accurate and complete. ☐ Listeners are likely to gain new insights about the topic. | 40% |
| Delivery | <ul style="list-style-type: none"> ☐ The speaker appears anxious and uncomfortable. ☐ Speaker reads notes, rather than speaks. ☐ Listeners are largely ignored. | <ul style="list-style-type: none"> ☐ The speaker is generally relaxed and comfortable. ☐ Speaker too often relies on notes. ☐ Listeners are sometimes ignored or misunderstood. | <ul style="list-style-type: none"> ☐ The speaker is relaxed and comfortable. ☐ Speaker speaks without undue reliance on notes. ☐ Speaker interacts effectively with listeners. | 20% |
| References | <ul style="list-style-type: none"> ☐ Speaker fails to integrate journal articles into the speech. | <ul style="list-style-type: none"> ☐ Speaker integrates 1 or 2 journal articles into the speech. | <ul style="list-style-type: none"> ☐ Speaker integrates 3 or more journal articles into the speech. | 20% |

Assessment vs. Grading Concerns

- Grading rubrics may include criteria that are not related to the learning outcome being assessed. These criteria are used for grading, but are ignored for assessment.
 - Grading requires more precision than assessment.
 - If multiple faculty will use the rubric for grading or assessment, consider calibrating them. This is especially important when doing assessment.
-

Rubrics Can:

- Speed up grading
 - Clarify expectations to students
 - Reduce student grade complaints
 - Make grading and assessment more efficient and effective by focusing the faculty member on important dimensions
 - Help faculty create better assignments that ensure that students display what you want them to demonstrate
-

Suggestions for Using Rubrics in Courses

1. Hand out the grading rubric with the assignment so students will know your expectations and how they'll be graded.
 2. Use a rubric for grading student work and return the rubric with the grading on it.
 3. Develop a rubric with your students for an assignment or group project. Students can then monitor themselves and their peers using agreed-upon criteria that they helped develop. Many faculty find that students will create higher standards for themselves than faculty would impose on them.
 4. Have students apply your rubric to some sample products before they create their own. Faculty report that students are quite accurate when doing this, and this process should help them evaluate their own products as they are being developed. The ability to evaluate, edit, and improve draft documents is an important skill.
 5. Have students exchange paper drafts and give peer feedback using the rubric, then give students a few days before the final drafts are turned in to you. You might also require that they turn in the draft and scored rubric with their final paper.
 6. Have students self-assess their products using the grading rubric and hand in the self-assessment with the product; then faculty and students can compare self- and faculty-generated evaluations.
-

Typical Four-Point Rubric Levels

1. Below Expectations. Student's demonstrated level of understanding clearly does not meet our expectations. Major ideas may be missing, inaccurate, or irrelevant to the task.
 2. Needs Improvement. Student needs to demonstrate a deeper understanding to meet our expectations, but does show some understanding; student may not fully develop ideas or may use concepts incorrectly.
 3. Meets Expectations. Student meets our expectations, performs at a level acceptable for graduation, demonstrates good understanding, etc.
 4. Exceeds Expectations. Student exceeds our expectations, performs at a sophisticated level, identifies subtle nuances, develops fresh insights, integrates ideas in creative ways, etc.
-

Rubric Category Labels

- Below Expectations, Developing, Acceptable, Exemplary
 - Novice, Apprentice, Proficient, Expert
 - Emerging, Developing, Proficient, Insightful
 - Below Basic, Basic, Proficient, Advanced (AAC&U Board of Directors, *Our Students Best Work*, 2004)
-

Creating a Rubric

1. Adapt an already-existing rubric.
 2. Analytic Method
-

Drafting the Rubric

I generally find it easier to start at the extremes when drafting the criteria in the rubric's cells, then move up and down to draft the levels in the middle. Starting at the lowest and highest cells, you ask:

- What are the characteristics of an unacceptable product, the worst product you could imagine, a product that results when students are very weak on the outcome being assessed?
- What are the characteristics of a product that would be exemplary, that would exceed your expectations, that would result when the student is an expert on the outcome being assessed?

Some words I find helpful:

(in)complete, (in)accurate, (un)reasonable, detailed, thorough, creative, original, subtle, sophisticated, synthesizes, integrates, analyzes, minor/major conceptual errors, flexibility, adaptability, complexity of thought, clarity, well-documented, well-supported, professional, organized, insightful, relevant

Managing Group Readings

1. One reader/document.
 2. Two independent readers/document.
 3. Paired readers.
-

Before inviting colleagues to a group reading,

1. Collect the assessment evidence and remove identifying information.
 2. Develop and pilot test the rubric.
 3. Select exemplars of weak, medium, and strong student work.
 4. Consider pre-programming a spreadsheet so data can be entered and analyzed during the reading and participants can discuss results immediately.
-

Inter-Rater Reliability

- Correlation Between Paired Readers
 - Discrepancy Index
-

Group Orientation and Calibration

1. Describe the purpose for the review, 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. Explain that readers should rate each dimension of an analytic rubric separately, and they should apply the criteria without concern for how often each category is used.
5. Give each reviewer a copy of several student products that are exemplars of different levels of performance. Ask each volunteer to independently apply the rubric to each of these products, and show them how to record their ratings.
6. Once everyone is done, collect everyone's ratings and display them so everyone can see the degree of agreement. The facilitator generally asks raters to raise their hands when their score is announced, and results are displayed in a simple chart.
7. Guide the group in a discussion of their ratings. There will be differences, and 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. 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.

8. Once the group is comfortable with the recording form and the rubric, distribute the products and begin the data collection.
 9. If you accumulate data as they come in and can easily present a summary to the group at the end of the reading, you might end the meeting with a discussion of five questions:
 - a. Are results sufficiently reliable?
 - b. What do the results mean? Are we satisfied with the extent of student learning?
 - c. Who needs to know the results?
 - d. If we're disappointed with the results, how might we close the loop?
 - e. How might the assessment process, itself, be improved?
-

Assessment Standards: How Good Is Good Enough?

Typical Standard:

We would be satisfied if at least 80% of the students are at level 3 or higher.
