Program Assessment

Assessment is an on-going process designed to monitor and improve student learning. Faculty:
1. develop explicit statements of what students should learn (SLOs).
2. verify that the program is designed to foster this learning (curriculum mapping).
3. develop a meaningful, manageable, sustainable assessment plan.
4. collect empirical evidence that indicate student attainment.
5. assess the evidence and reach a conclusion (faculty are satisfied or disappointed with the extent of student learning).
6. use these data to improve student learning (close the loop).

Common GE Learning Outcomes

1. Writing: Students can write effectively.
2. Speaking: Students can make effective oral presentations.
3. Quantitative Reasoning: Students can use quantitative reasoning to analyze or solve problems.
4. Information Competence: Students can identify, analyze, integrate, and present information ethically.
5. Critical Thinking: Students can think critically about issues, perspectives, or problems.

GE Outcomes

1. Outcomes have active verbs--how students can demonstrate their learning.
2. Outcomes can be stated in simple language; the details are in the rubrics.
3. Outcomes should be real, not aspirational.
4. Avoid compound outcomes that require multiple lines of evidence.
5. Sometimes an "or" helps us create GE outcomes that fit the variety of courses that students might take to meet requirements. (Example: Students can create or analyze a work of art.)

Assessment Evidence

Well-articulated learning outcomes clarify what faculty want students to learn and how the assessment should be conducted. For example, how would you assess each of the following outcomes?

1. Students can list major events in American history.
2. Students can describe major events and trends in American history.
3. Students can apply their knowledge of American history to examine contemporary American issues.
College of Charleston GE Learning Outcomes

Academic Writing Student Learning Outcomes:
1. Students construct persuasive arguments, analyze the arguments of others, and incorporate research material that includes persuasive evidence from experts.
2. Students analyze and evaluate numerous examples of academic writing, conduct library research, and incorporate academic research materials in their papers.

Pre-Modern History Student Learning Outcomes
3. Students demonstrate knowledge of the history of human civilization, societies, and cultures and an awareness of the historical experience of the pre-modern era.
4. Students demonstrate that they can situate primary historical records of the pre-modern era in their contexts and use these sources to construct historical arguments.

Modern History Student Learning Outcomes
5. Students demonstrate knowledge of the history of human civilization, societies, and cultures and an awareness of the historical experience of the modern era.
6. Students demonstrate that they can situate primary historical records of the modern era in their contexts and use these sources to construct historical arguments.

Natural Science Student Learning Outcomes
7. Students apply physical/natural principles to analyze and solve problems.
8. Students demonstrate an understanding of the impact that science has on society.

Math/Logic Student Learning Outcomes
10. Students apply the models to answer questions and establish results.

Foreign Languages, Classical or Modern Student Learning Outcomes
11. Students read, write, and understand languages other than English.
12. Students use their knowledge of languages other than English to analyze the perspectives of historical and/or modern cultures that can be obtained only through reading and/or listening to that language.

Social Science Student Learning Outcomes
13. Students demonstrate an understanding of how empirical evidence and the varieties of social scientific methods of inquiry provide explanations for human behavior, social interactions, and/or social institutions.
14. Students demonstrate an understanding of the impact the social sciences have on society.

Humanities Student Learning Outcomes
15. Students analyze how ideas are represented, interpreted, or valued in various expressions of human culture.
16. Students examine relevant primary source materials as understood by the discipline and interpret the material in writing assignments (or alternatives that require equally coherent and sustained analysis).
Properties of Good Assessment Data

- Valid—directly reflects the learning outcome being assessed
- Reliable—including inter-rater reliability when subjective judgments are made
- Actionable—results point reviewers toward challenges that can be approached

Embedded Assessment/Signature Assignments

- Exams or parts of exams
- Class activities or labs
- Homework assignments
- Group projects
- In-class presentations
- Community-service learning, practica, internships, and other fieldwork activities
- Culminating projects, such as papers in capstone courses

Assignments and activities are purposefully created to collect information relevant to specific learning outcomes. Results are pooled across courses and instructors to indicate program accomplishments, not just the learning of students in specific courses.

Consider systematically integrating signature assignments into the curriculum, i.e., assignments designed to assess specific learning outcomes. Signature assignments must require students to perform the outcome.

Drafting Your Own Signature Assignment(s)

- Should be given near the end of the learning experience.
- Must require students to perform the outcomes.
- What characteristics must your signature assignment(s) have (e.g., length, format, citations, proof, elaboration)?
- Draft an example or two of an assignment that meets your signature assignment requirements.
Signature Assignment Example

Outcome: Students who complete the GE program can make effective oral presentations.

Evidence: Students must plan and deliver a 4-5 minute formal presentation on a topic that meets course expectations and that allows listeners to learn something. The assignment must require students to organize the presentation effectively, provide accurate information, and effectively deliver the presentation to the audience. The audience may be fellow students, but the instructor may decide to give the audience a different role, such as university administrators. The presentation should be delivered during the last two weeks of the semester.

Course Example: FYE Course on Urban Planning
Identify something that would increase the well-being of Charleston citizens. You have four to five minutes to address the Charleston City Council to offer your idea, explain why they should implement it, and suggest implementation strategies. Be sure that your presentation is well-organized, accurate, and effectively delivered. Your idea should be something that the city actually could do to benefit its citizens.

Course Example: Introduction to Psychology
Select a major psychological concept or theory. Your task is to teach this concept or theory to an audience of high school students who have never studied psychology, and your presentation should take between four and five minutes. You should carefully define your concept or theory, and you should give three real-world examples that illustrate it in ways that will help students understand it. Be sure that your presentation is well-organized, accurate, and effectively delivered.

Assessing the Evidence

- usually involves subjective judgments concerning complex products or behaviors
- rubrics provide the criteria to guide these subjective judgments

Rubrics

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 clarify expectations to students, to provide formative feedback to students, to grade students, and/or to assess courses and programs.

There are two major types of rubrics:
- Holistic rubric — one global, holistic score for a product or behavior
- Analytic rubric — separate, holistic scoring of specified characteristics of a product or behavior
Rubric Examples

- Campus Examples
- VALUE Rubrics (Valid Assessment of Learning in Undergraduate Education; AAC&U)

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.

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

- Does Not Meet Expectations, Needs Improvement, Meets Expectations, Exceeds Expectations
- 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

Rubric Criteria

Say you are drafting a holistic rubric to assess students' command of basic grammar and punctuation. Look at these four examples. Which do you prefer? Why?

**Version 1**

<table>
<thead>
<tr>
<th>Does Not Meet Expectations</th>
<th>Needs Improvement</th>
<th>Meets Expectations</th>
<th>Exceeds Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 or more errors</td>
<td>5-12 errors</td>
<td>2-7 errors</td>
<td>0-1 error</td>
</tr>
</tbody>
</table>

**Version 2**

<table>
<thead>
<tr>
<th>Does Not Meet Expectations</th>
<th>Needs Improvement</th>
<th>Meets Expectations</th>
<th>Exceeds Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 or more errors</td>
<td>5-9 errors</td>
<td>2-4 errors</td>
<td>0-1 error</td>
</tr>
</tbody>
</table>

**Version 3**

<table>
<thead>
<tr>
<th>Does Not Meet Expectations</th>
<th>Needs Improvement</th>
<th>Meets Expectations</th>
<th>Exceeds Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td>lowest 10%; more errors than 90% of the other students</td>
<td>bottom half; more errors than the average student, but not in the bottom 10%</td>
<td>top half; fewer errors than the average student, but not in the top 10%</td>
<td>top 10%; fewer errors than 90% of the students</td>
</tr>
</tbody>
</table>

**Version 4**

<table>
<thead>
<tr>
<th>Does Not Meet Expectations</th>
<th>Needs Improvement</th>
<th>Meets Expectations</th>
<th>Exceeds Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td>errors frequently interfere with reader's ability to understand meaning</td>
<td>many errors or occasional errors that interfere with reader's ability to understand meaning</td>
<td>contains a few errors, but they do not interfere with reader's ability to understand meaning</td>
<td>virtually free of errors</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Rubric for Assessing Oral Presentations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below Expectation</td>
</tr>
<tr>
<td><strong>Organization</strong></td>
</tr>
<tr>
<td><strong>Content</strong></td>
</tr>
<tr>
<td><strong>Delivery</strong></td>
</tr>
</tbody>
</table>
Alternative Format 1.
Points are assigned and used for grading, as shown below, and the categories (Below Expectation, Needs Improvement, 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.

<table>
<thead>
<tr>
<th></th>
<th>Below Expectation</th>
<th>Needs Improvement</th>
<th>Satisfactory</th>
<th>Exceeds Expectations</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization</td>
<td>No apparent organization. Evidence is not used to support assertions.</td>
<td>There is some organization, but the speaker occasionally goes off topic. Evidence used to support conclusions is weak.</td>
<td>The presentation has a focus and provides some reasonable evidence to support conclusions.</td>
<td>The presentation is carefully organized and provides convincing evidence to support conclusions.</td>
<td>(0-6)</td>
</tr>
<tr>
<td>Content</td>
<td>The content is inaccurate or overly general. Listeners are unlikely to learn anything or may be misled.</td>
<td>The content is sometimes inaccurate or incomplete. Listeners may learn some isolated facts, but they are unlikely to gain new insights about the topic.</td>
<td>The content is generally accurate and reasonably complete. Listeners may develop a few insights about the topic.</td>
<td>The content is accurate and comprehensive. Listeners are likely to gain new insights about the topic.</td>
<td>(0-4)</td>
</tr>
<tr>
<td>Delivery</td>
<td>The speaker appears anxious and uncomfortable and reads notes, rather than speaks. Listeners are ignored.</td>
<td>The speaker occasionally appears anxious or uncomfortable, and may occasionally read notes, rather than speak. Listeners are often ignored or misunderstood.</td>
<td>The speaker is generally relaxed and comfortable. Listeners are generally recognized and understood.</td>
<td>The speaker is professional, relaxed, and comfortable and interacts effectively with listeners.</td>
<td>(0-9)</td>
</tr>
</tbody>
</table>

Total Score
Alternative Format 2.
Weights are used for grading; categories (Below Expectation, Needs Improvement, 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.

<table>
<thead>
<tr>
<th>Rubric for Grading Oral Presentations</th>
<th>Below Expectation</th>
<th>Needs Improvement</th>
<th>Satisfactory</th>
<th>Exceeds Expectations</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization</td>
<td>No apparent organization. Evidence is not used to support assertions.</td>
<td>There is some organization, but the speaker occasionally goes off topic. Evidence used to support conclusions is weak.</td>
<td>The presentation has a focus and provides some reasonable evidence to support conclusions.</td>
<td>The presentation is carefully organized and provides convincing evidence to support conclusions.</td>
<td>30%</td>
</tr>
<tr>
<td>Content</td>
<td>The content is inaccurate or overly general. Listeners are unlikely to learn anything or may be misled.</td>
<td>The content is sometimes inaccurate or incomplete. Listeners may learn some isolated facts, but they are unlikely to gain new insights about the topic.</td>
<td>The content is generally accurate and reasonably complete. Listeners may develop a few insights about the topic.</td>
<td>The content is accurate and comprehensive. Listeners are likely to gain new insights about the topic.</td>
<td>20%</td>
</tr>
<tr>
<td>Delivery</td>
<td>The speaker appears anxious and uncomfortable and reads notes, rather than speaks. Listeners are ignored.</td>
<td>The speaker occasionally appears anxious or uncomfortable, and may occasionally read notes, rather than speak. Listeners are often ignored or misunderstood.</td>
<td>The speaker is generally relaxed and comfortable. Listeners are generally recognized and understood.</td>
<td>The speaker is professional, relaxed, and comfortable and interacts effectively with listeners.</td>
<td>50%</td>
</tr>
<tr>
<td>Comments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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, Needs Improvement, Satisfactory, Exemplary) can be used for assessment.

<table>
<thead>
<tr>
<th>Rubric for Assessing Oral Presentations</th>
<th>Below Expectation</th>
<th>Needs Improvement</th>
<th>Satisfactory</th>
<th>Exceeds Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization</td>
<td>☐ No apparent organization. ☐ Evidence is not used to support assertions.</td>
<td>☐ There is some organization, but the speaker occasionally goes off topic. ☐ Evidence used to support conclusions is weak.</td>
<td>☐ The presentation has a focus. ☐ Student provides some reasonable evidence to support conclusions.</td>
<td>☐ The presentation is carefully organized. ☐ Speaker provides convincing evidence to support conclusions.</td>
</tr>
<tr>
<td>Content</td>
<td>☐ The content is inaccurate or overly general. ☐ Listeners are unlikely to learn anything or may be misled.</td>
<td>☐ The content is sometimes inaccurate or incomplete. ☐ Listeners may learn some isolated facts, but they are unlikely to gain new insights about the topic.</td>
<td>☐ The content is generally accurate and reasonably complete. ☐ Listeners may develop a few insights about the topic.</td>
<td>☐ The content is accurate and comprehensive ☐ Listeners are likely to gain new insights about the topic.</td>
</tr>
<tr>
<td>Delivery</td>
<td>☐ The speaker appears anxious and uncomfortable and reads notes, rather than speaks. ☐ Listeners are ignored.</td>
<td>☐ The speaker occasionally appears anxious or uncomfortable, and may occasionally read notes, rather than speak. ☐ Listeners are often ignored or misunderstood.</td>
<td>☐ The speaker is generally relaxed and comfortable. ☐ Listeners are generally recognized and understood.</td>
<td>☐ The speaker is professional, relaxed, and comfortable. ☐ The speaker interacts effectively with listeners.</td>
</tr>
<tr>
<td>Comments</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

10
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.

<table>
<thead>
<tr>
<th>Rubric for Assessing Oral Presentations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Below Expectation</strong></td>
</tr>
<tr>
<td><strong>Organization</strong></td>
</tr>
<tr>
<td>☐ Evidence is not used to support assertions.</td>
</tr>
<tr>
<td><strong>Content</strong></td>
</tr>
<tr>
<td>☐ Listeners are unlikely to learn anything or may be misled.</td>
</tr>
<tr>
<td><strong>Delivery</strong></td>
</tr>
<tr>
<td>☐ Listeners are ignored.</td>
</tr>
<tr>
<td><strong>References</strong></td>
</tr>
</tbody>
</table>
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.
- Grading results in one score or grade; analytic rubrics result in multiple scores that are separately examined.
- 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
- Improve the reliability and validity of assessments and grades
- 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.

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.

**Inter-Rater Reliability**

- Correlation Between Paired Readers
- Discrepancy Index

**Calculation of Reliability Coefficients**

Say you use an analytic rubric with 3 dimensions (organization, accuracy, and writing style) to assess eight artifacts. Here are the scores for the two raters who scored each artifact:

<table>
<thead>
<tr>
<th>Artifact</th>
<th>Organization</th>
<th>Accuracy</th>
<th>Writing Style</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rater 1</td>
<td>Rater 2</td>
<td>Rater 1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

**Inter-rater reliability**

- Correlation=0.34
- Correlation=0.83
- Correlation=0.95

**Discrepancy Index**

- Discrepancy of 0: 38% 75% 88%
- Discrepancy of 1: 38% 25% 12%
- Discrepancy of 2: 25% 0% 0%
- Discrepancy of 3: 0% 0% 0%

What do you conclude about the reliability of these ratings?
One Way to Calculate the Reliability Estimates: Use Excel.

In an Excel spreadsheet rows are numbered and columns are lettered. Below are the data for the eight artifacts summarized on the previous page. Each cell is identified by its column letter and row number. For example, the score in Cell A1 is a 1 and the score in Cell E7 is a 3.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

To correlate columns A and B select Statistical Functions from the Formula Menu and in the boxes type A1:A8 and B1:B8 to identify the cells in the two columns that will be correlated. Excel will give you the correlation. Round it to two decimal places.

The "Diff" columns contain the absolute value of the size of the difference (i.e., the difference, ignoring sign). Entries in these columns are used to calculate the discrepancies.

---

**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 nor 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 (to avoid the "halo effect"), and they should apply the criteria without concern for how often each category is used (to avoid "grading on a curve" and norm-referenced judgments).
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.
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. Distribute the products and begin the data collection using either independent or paired readers.
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.

---

**Develop/Refine Your Own Rubric**

**Step 1.**
**State your outcome.** Outcomes are expressed in behavioral terms—how students can demonstrate their learning, e.g.,

- Students can explain the ethical standards for conducting sociological research.
- Students can describe the major factors that influenced the development of the American political system.
- Students can collaborate with others effectively.
- Students can define basic biological concepts and theories.
- Students can apply their knowledge of child development to parenting, education, and public policy issues related to children and families.
- Students can analyze data and interpret the results of statistical tests.
- Students can locate appropriate sources by searching electronic databases.

**Step 2.**
**Consider your signature assignment(s).** Each assignment should align with the outcome and rubric. It should require the student to perform the outcome (e.g., explain, describe, collaborate, define, apply, etc.) so the rubric can be used to assess how well the students have mastered the outcome. Be sure that the assignment tells students to do all the things you will be assessing in the rubric, and consider handing out the rubric with it.

**Step 3.**
**Draft the rubric.** Do you want a holistic or analytic rubric? How many rating categories do you want? What are their labels? If analytic, what are the major dimensions of the outcome that you want to assess? Be sure that the criteria in the cells can be applied to the evidence you plan to collect.

**Step 4.**
**Consider the standard.** What level of performance would be satisfactory?

**Step 5.**
**Pilot test.** Be sure to pilot test the rubric before using it in a real assessment study.
Example: Writing Rubric

Outcome: Students who graduate from our campus can write effectively.

When the faculty discussed this outcome, they decided that students should be able to prepare papers that are focused and well-organized, they provide evidence or rationale to support each conclusion, and their writing has appropriate spelling, punctuation, and sentence complexity.

Evidence: Seniors in capstone courses write a 3-5 page paper in which they take and defend a position relative to a controversy in their field. Instructors may identify a single controversy that all students explore, or they may choose to identify several from which students can choose.

Standard: We'd be satisfied if at least 80% of the seniors score at level 3 or higher for each of the three dimensions.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Rating Category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 below expectations</td>
</tr>
<tr>
<td></td>
<td>focus and organization</td>
</tr>
<tr>
<td></td>
<td>support for conclusions</td>
</tr>
<tr>
<td></td>
<td>spelling, punctuation, and sentence complexity</td>
</tr>
</tbody>
</table>
Holistic Rubric Template

Outcome:

Evidence:

**Standard**: We will be satisfied if:

<table>
<thead>
<tr>
<th>Category</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
Analytic Rubric Template

Outcome:

Evidence:

Standard: We will be satisfied if:

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Rating Category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

|           | 1               | 2               | 3               | 4               |

|           | 1               | 2               | 3               | 4               |

|           | 1               | 2               | 3               | 4               |

|           | 1               | 2               | 3               | 4               |