

Make Room for Rubrics

Rubrics are receiving high marks from teachers around the country. Here's how to use these scoring devices for authentic assessment.

By [Mary Rose](#)

Exactly how can teachers determine whether a **student's piece of writing** meets the standard of exceptional versus that of good? How can a child's **science project**, which may involve drawing, writing, dioramas, oral presentation, and other elements, be accurately evaluated? How can subtle, gradual progress in the development of speaking skills be measured? One increasingly popular method is rubrics — a type of scoring guide **used to assess more complex, subjective criteria**. Rubrics enable an evaluation of student performance in situations that more closely replicate the challenges of real life than isolated tests. As such, they **support the mandate for authentic assessment** stated in national standards across the curriculum.

A rubric is **a device for organizing and interpreting data gathered from observations of student performance**. More precisely, it is a scoring guide that **differentiates between levels of development** in a specific area of performance or behavior. Conventional rubrics, such as the Science Journal Rubric pictured opposite, use a range of three or more levels to assess performance — for example, from beginning to developing to proficient. Each of the levels contains specific, measurable performance characteristics, such as “makes few/occasional/frequent spelling errors.” Checklists, which provide specific steps for completing tasks to the highest level of quality, such as the Student Rubric, opposite, are another form of rubric. Combination rubrics incorporate aspects of both.

Why Use Rubrics?

Rubrics differ from traditional methods of assessment in that they **examine students in the actual process of learning, clearly showing them how their work is being evaluated**. Rubrics communicate **detailed explanations** of what constitutes excellence throughout a project and provide a clear teaching directive. “Rubrics **help teachers clarify exactly what students need to achieve** in content and performance standards,” says Rob Southworth, education consultant for District 2, in New York City. Because rubrics set forth precise criteria, teachers are **better able to assess skills that may fall outside the scope of traditional testing**. Consistent scores attached to each level of a rubric, such as 1 through 4, can provide an objective basis for assigning grades.

When shared with children before a project or an assignment, rubrics **can be powerful motivational tools**. If students are given the chance to contribute to the content of a rubric, “then it is much easier to hold them to its standards,” says Charlotte Sassman, a kindergarten teacher at the Alice Carlson Applied Learning Center in Fort Worth, Texas. Rather than directing youngsters toward past performance (“Why did I get a B instead of an A?”), rubrics can teach them to focus on current and future performance (“What steps can I take to progress to the next level?”).

Student rubrics used for self-assessment, such as “How Good Is My Book Cover?,” pictured above, encourage learners to participate in the grading process. Kara Staunton, a seventh-grade teacher at Reid Middle School in Pittsfield, Massachusetts, finds that “the students like feeling as if they have a voice. And surprisingly, they are harder graders than I am!”

Finding and Building Rubrics

Creative, ready-made rubrics are widely available. Among the many excellent resources: *35 Rubrics and Checklists to Assess Reading and Writing*, by Adele Fiderer (Scholastic Professional Books, 1998. Grades K–2); *Rubrics, Checklists & Other Assessments for the Science You Teach*, by Ann Flagg (Scholastic, 1998. Grades 1–3); and *The Rubrics Way*, by David Lazear (Zephyr Press, 1998. Grades K–12).

For best results, design your own all-purpose rubric template, or frame, that can be adapted for different projects. Base your template on the Teacher Rubric above, filling in criteria that apply to your subject area and the elements you need to evaluate. Know the specific skills that you want students to develop

throughout the activity. Describe the criteria that reflect the highest level of performance. And vary your descriptions of accuracy, completion, consistency, quality, and other factors to signify performance levels. Good resources for building your own rubric can be found online at Servtech (www.servtech.com/~germaine/rubric.html) and Ask Dr. Rubric (www.classnj.org/IDEA/).

Putting Rubrics to the Test

A rubric's strength is its specificity, which means that individual students can fall between levels, attaining some but not all standards in a higher level. Some teachers find that attaching a plus or minus sign to the level, thereby creating more levels between beginning and proficient, or making checklists within a rubric can help to define progress. And while rubric scores can be translated into final grades, it is important that we remind children that not every rubric score "counts." Rubrics serve, above all, to inform and improve teachers' instruction while giving students the feedback they need to learn and grow.

Mary Rose, M.A., is a fourth-grade teacher at Lake Sybelia Elementary School, Orange County, Florida.

The rubric should organize and clarify the scoring criteria well enough so that two teachers who apply the rubric to a student's work will generally arrive at the same score. The degree of agreement between the scores assigned by two independent scorers is a measure of the reliability of an assessment. This type of consistency is needed for a performance assessment to yield good data that can be meaningfully combined across classrooms and used to develop school improvement plans.

A good scoring rubric will:

- Help teachers define excellence and plan how to help students achieve it.
- Communicate to students what constitutes excellence and how to evaluate their own work.
- Communicate goals and results to parents and others.
- Help teachers or other raters be accurate, unbiased and consistent in scoring.
- Document the procedures used in making important judgments about students.

-- Herman, Aschbacher, and Winters (1992)

A scoring rubric has several components, each of which contributes to its usefulness. These components include one or more dimensions on which performance is rated, definitions and examples that illustrate the attribute(s) being measured and a rating scale for each dimension. Ideally, there should also be examples of student work that fall at each level of the rating scale.

Elements of a scoring rubric

- One or more traits or **dimensions** that serve as the basis for judging the student response
- **Definitions and examples** to clarify the meaning of each trait or dimension
- A **scale** of values on which to rate each dimension
- **Standards** of excellence for specified performance levels accompanied by models or examples of each level

-- Herman, Aschbacher, and Winters (1992)

The scoring rubric rating scales may be numerical, qualitative, or a combination of the two. **Qualitative rubrics may have scale points with labels such as these:**

- Not yet, developing, achieving
- Emerging, developing, achieving
- Novice, apprentice, proficient, distinguished
- No evidence, minimal evidence, partial evidence, complete evidence

California's math rubric uses a **combination of numerical and qualitative scales:**

1. Unable to begin effectively
2. Begins, but fails to complete problem
3. Serious flaws but nearly satisfactory
4. Minor flaws but satisfactory
5. Competent response
6. Exemplary response

Each point on the scale should be clearly labeled and defined. There is no single best number of scale points, although it is best to **avoid scales with more than 6-7 points.** With very long scales, it is often difficult to adequately differentiate between adjacent scale points (e.g., on a 100-point scale, it would be hard to explain why you assigned a score of 81 rather than 80 or 82). It is also harder to get different scorers to agree on ratings when very long scales are used. The rule of thumb is to **have as many scale points as can be well defined** and that **adequately cover the range** from very poor to excellent performance.

How many points should a rating scale have?

There is no one right answer to this question. Consider these as you make your decision:

- Each point on the scale needs to be well defined. This may be difficult to do for longer scales.

- Longer scales make it harder to get agreement among scorers (inter-rater reliability).
- Extremely short scales make it difficult to identify small differences between students.
- Do you simply want to divide students into two or three groups, based on whether they have attained or exceeded the standard for an outcome? If so, then a short scale may be adequate.
- If you are rating a product/performance on several different dimensions, will you want to add up the scores so that each is equally weighted? If so, you may find it easier to have all scales the same length.

Evaluating Rubrics

There is no single best rubric for all purposes and many different rubrics could be applied to the same task. The following criteria for evaluating scoring rubrics are adapted from Herman, Aschbacher and Winters (1992), Arter (1990) and ISBE (1994).

Does the rubric relate to the outcome(s) being measured? The rubric should address all aspects of the outcome(s) being measured and it should *not* address anything extraneous. For example, spelling and grammar might be considered extraneous on a science assessment, unless it is measuring an outcome that deals specifically with communication. If the assessment is to address critical thinking skills, does the rubric reflect that?

Does it cover important dimensions of student performance? Do the criteria reflect current conceptions of excellence in the field? Does the rubric reflect what you emphasize in your teaching? Does the highest scale point represent a truly exemplary performance or product? When you evaluate rubrics, you need not be concerned about having a certain number of students score at every point on the scale. It may be that no student will attain the highest scale point. Nonetheless, it still may be worthwhile to have that point on the scale as a standard of excellence for which students should strive.

Are the dimensions or scales well-defined? Is it clear to everyone what each scale measures?

Is there a clear basis for assigning scores at each scale point? Is it clear exactly what a student needs to do to get a score at each scale point? Can you easily differentiate between scale points? An easily understood rubric with clear definitions of each score point is the ideal. Conversely, it is usually best to avoid scales that are labeled only at the highest and lowest points.

Can the rubric be applied consistently by different scorers? Inter-rater reliability depends on how well the scales and scale points are defined and the extent to which you

and your colleagues can arrive at consensus about how performance should be measured and what constitutes good performance.

Can the rubric be understood by students and parents? Can it be explained without technical jargon and in terms that even young students can understand?

Is the rubric developmentally appropriate? Does the rubric specify some score points appropriate for the age of the student?

Is the rubric fair and free from bias? Does it reflect teachable skills or does it address variables over which students and educators have no control, such as the student's culture, gender or home resources? Does the rubric reward or penalize students based on skills unrelated to the outcome being measured? Have all students had an equal opportunity to learn the content and skills addressed in the rubric? Is the rubric appropriate for the conditions under which the task was completed?

Is the rubric useful, feasible, manageable and practical? Does the rubric make sense to you? Will it provide the kind of information you need and can use effectively? Does the rubric have a reasonable number of scales and score points? It is difficult to rate student work on many scales at once and it may be hard for students to assimilate data from many score scales. Using long scales is often more time-consuming than using shorter ones.