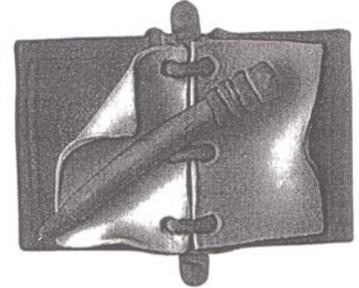


Designing Great Rubrics

By Michael Simkins

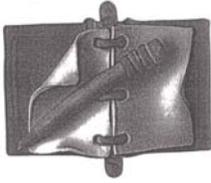
**Rubrics are all the rage these days.
But how do you design ones that work?
Here are ten pointers.**



If you look in the dictionary, you'll find the word *rubric* has several definitions. First and foremost, it means a heading of a part of a book or manuscript set apart in decorative red lettering. Less commonly, it refers to a liturgical direction. It can mean a short commentary on a broad subject. Or it can be simply a synonym for *category*. In education, of course, it means none of these things. We've latched onto the term to refer to the sets of formal guidelines we use to rate examples of student work. These guidelines are usually presented in the form of a matrix with performance levels in the top row and performance dimensions along the left column. The rest of the cells hold the criteria required to earn a given score in a given dimension.

Rubrics can be extremely useful tools, and teachers are often encouraged to use them to score everything from writing samples to multimedia projects. But a rubric is only as useful as it is good. Using a bad rubric is a waste of time, if not downright harmful. Whether you are choosing from among ready-made rubrics or creating your own, here are some pointers to keep in mind.

- 1 Don't make task-specific rubrics.** By task-specific, I mean a rubric crafted for a single assignment or project. It's not immoral; you just don't have the time. To make a really sound rubric takes a lot of thought and work. If you're going to invest enough effort to make a good rubric, be sure you can use it in a range of situations. Creating a more general rubric is not only a more efficient use of your time, it makes it possible for you to track student improvement on successive projects.
- 2 Don't be excessively general.** While you're being careful to avoid making a task-specific rubric, don't let yourself go to the other extreme. For example, you probably don't want to make a "project rubric" and try to use it to judge everything from a digital montage to a *Power-Point* presentation on market economics. Instead, remember the golden mean and aim for a mid-level of generalization in your rubric.
- 3 Avoid dysfunctional detail.** Believe it or not, I recently saw a rubric that was eleven pages long! While it was called a rubric, it was really more like an inventory or checklist. It asked the scorer to rate everything from the quality of the title page to the number of people who helped in producing the project.



Designing Rubrics

4

Focus on a limited number of dimensions. By dimensions, I mean the main areas or parts of the work you care about. Dimensions tell you “where to look” when you are viewing a product or performance. For example, if we were making a “restaurant rubric”, our dimensions might be food, service, atmosphere and price. If we were rating physical fitness, the dimensions might be strength, endurance, and flexibility. By limiting the number of dimensions to no more than four or five, you’ll be able to do a more thorough job of developing each one, and you’ll be forced to set priorities for what really matters in the project.

5

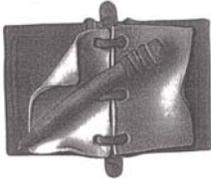
Use key, teachable criteria. Just as dimensions tell you “where to look”, criteria tell you “what to look for.” If you’re rating the quality of food, your criteria might be flavor, appearance, and healthfulness. Your service criteria might have to do with the server’s promptness, efficiency, and attitude. In any case, you want to specify those things that matter the most for each dimension. And as a teacher, you want to focus on things you can teach and students can learn. For example, it’s likely you can teach students how to research a topic, summarize information, and design a legible Web page that uses complementary colors. It’s less certain you can teach your students to be creative, inventive, or imaginative, so you may not want to include such criteria in your rubrics.

6

Use measurable criteria. As much as possible, choose criteria that you can count, or that you can at least mark as “present” or “not present.” Returning to our restaurant analogy, we can measure things such as how long it takes the server to show up to take our order or how much fat and cholesterol is in the food. We can determine whether or not the place is “too noisy to carry on a conversation” or if we spy cockroaches scurrying along the baseboards. In our classroom rubrics, we want to look for the same sorts of concrete, measurable criteria. Is it legible? Do the links work?

7

Select descriptors carefully. As you strive for measurability, you may be tempted to use relative terms; try to avoid them. For example, in elaborating on criteria many rubrics include terms such as *poor*, *fair*, *good*, and *excellent*, or *minimal*, *adequate*, and *fully developed*. This really begs the issue: What’s the difference between fair and good service? How much detail is enough for an idea to be fully developed? When the rubric results will be viewed by students, it’s also important to consider the emotional aspects of value-laden terms such as *boring* or *poorly presented*. To provide specific feedback without discouraging your students, it’s better to identify the actual traits you’re looking for that constitute poor versus good performance. If a good presentation is one that “begins with a hook, ties facts together, and ends with a strong point,” then say so.



Designing Rubrics

8

Aim for four levels. Typically, you'll see rubrics with three to five performance levels. Occasionally, you'll see six or more. For classroom use, I suggest you try using four levels. The problem with using three levels is that you probably won't be able to make fine enough discriminations. When I was in high school, I went to work for Sears as a sales clerk. They taught us that all Sears merchandise fell into three types: "good, better, and best." In the classroom, you need to be prepared to have "not so good," at least in the beginning stages of learning. On the other end of the spectrum, if you try to force student work into five or more levels, you may find yourself splitting hairs.

9

Keep the "distance" between levels equal. Ideally, the difference between a "1" and a "2" on your rubric should represent the same amount of "betterness" as the difference between a "2" and a "3" or a "3" and a "4." Put another way, work that scores "4" on your rubric should be four times better than work that scores "1." Statisticians call this an *interval scale*. In the messy business of classroom assessment, you will not reach such perfection. You're actually creating an *ordinal scale* that merely ranks one product or performance as better than another. Nonetheless, the more "interval-like" your scale, the sounder your rubric will be.

10

Include students in creating or adapting the rubric. Two heads are often better than one, especially in the creative process. By involving your students in the process of constructing or adapting the rubric that will be used to score their work, you accomplish several things. Your students will have a head start on understanding what your expectations are; they'll know where the target is. They're also more likely to "buy in" to the work and strive to do their best. But most importantly, their good ideas will add to the quality of the rubric itself.

To help you see these ideas in action, we have included some examples drawn from rubrics I've run across in my work. As you look at each of the rubrics on pages 28-29, think about how it does or does not embody the pointers in this article. You might even jot your thoughts on a piece of scratch paper. Then, turn to page 30 and compare your comments to mine.

Rubric Examples

Weak
Satisfactory
Strong

0 **1** **2** **3** **4** **5**

Example 1: Earth Science

Here is an example from a rubric developed for use in a middle school science class to evaluate student multimedia presentations. In addition to content, the rubric measures quality, connections, usage, and group work.

Content
<ul style="list-style-type: none"> • Presentation shows strong understanding of major ideas and connects global/ science tasks • Presentation displays evidence of critical thinking with text and slide compositions • Includes references to sources, including Internet/CD-ROM on last slide of presentation (at least two book references and two Internet sources/ addresses plus any computer sources) • Has a text clipping, 10 second personal voice-over, clip on, movie clipping (8 sec. max.) and a sound bite clipping (less than 20 sec.) • Includes a minimum of eight slide stills with at least one moving clipping • Includes at least one Quicktake image at beginning or end with caption • Has two backup (working) disks, one for sound, one for text/slides

Points	1	2	3	4	Total
Originality	The work is a minimal collection or rehash of other people’s ideas, products, and inventions. There is no evidence of new thought.	The work is an extensive collection and rehash of other people’s ideas, products, images and inventions. There is no evidence of new thought or inventiveness.	The product shows evidence of originality and inventiveness. While based on extensive collection of other people’s ideas, products, images and inventions, the work extends beyond that collection to offer new insights.	The product shows significant evidence of originality and inventiveness. The majority of the content and many of the ideas are fresh, original, inventive and based upon logical conclusions and sound research.	
Screen Design	Screens are either confusing and cluttered or barren and stark. Buttons or navigational tools are absent or confusing.	Screens are difficult to navigate, but some buttons and navigational tools work. Users can navigate a few screens.	Screens contain adequate navigational tools and buttons. Users can progress through screens in a logical path to find information.	Screens contain all necessary navigational tools and buttons. Users can progress intuitively through screens in a logical path to find information.	

Example 2: Multimedia Contest

In this example, you’ll find two of ten dimensions of the “Multimedia Mania 99 Rubric” developed to judge entries in a national competition. Some of the other dimensions are curriculum alignment, teamwork, subject knowledge, and mechanics. (The full rubric may be found at www.ncsu.edu/midlink/rub.multi.htm.)

Example 3: Research Project

Here, you have a section from a rubric developed by teacher Kay Evatz of Federal Heights Elementary Media Center in Westminster, Colorado, together with her upper elementary students. It was designed to assess work on an “animal assignment.” Part of the assignment was to develop and then explain some sort of product that related to the chosen animal. In addition to rating the student’s presentation, the rubric looks at the student’s choice of animal, research plan, and note-taking.

Presentation	In Progress	So-So	Good	Excellent
	<ul style="list-style-type: none"> • Boring • Sloppy • Unprepared • Not finished 	<ul style="list-style-type: none"> • Not practiced • Quiet voice • Mistakes in product 	<ul style="list-style-type: none"> • Clear voice • New information • Neat product • Practiced • Accurate, looks real 	<ul style="list-style-type: none"> • Creative • Uses expression • Motion • Real-looking product

Example 4: Poster Project

Students in a primary grade class were assigned to make posters about the sun. This rubric was created for students to use to self-assess their work. Two of five dimensions are shown here. Other dimensions included are “At least four sources of information,” “Sun’s impact,” and “Discussion of the sun: past, present, and future.”

	Incomplete (1)	Incomplete (2)	Incomplete (3)	Satisfactory (4)	Good (5)	Exceptional (6)
1. There is a labeled drawing of the sun	<ul style="list-style-type: none"> ▪ Unacceptable ▪ No drawing done ▪ No labels 	<ul style="list-style-type: none"> ▪ Unacceptable ▪ Drawing carelessly done ▪ Labels unrelated to drawing 	<ul style="list-style-type: none"> ▪ Unacceptable ▪ Drawing poorly done ▪ Labels inadequate 	<ul style="list-style-type: none"> ▪ Minimum acceptance ▪ Minimum effort given to drawing ▪ Labels adequate 	<ul style="list-style-type: none"> ▪ Drawing executed well ▪ Clear labels 	<ul style="list-style-type: none"> ▪ Great effort given to drawing ▪ Labels explain drawing exceptionally well
2. All information needs to be accurate	<ul style="list-style-type: none"> ▪ Unacceptable ▪ Little or no information present 	<ul style="list-style-type: none"> ▪ Unacceptable ▪ Information made up 	<ul style="list-style-type: none"> ▪ Unacceptable ▪ Information does not match poster 	<ul style="list-style-type: none"> ▪ Minimum acceptance ▪ Minimum effort given to research 	<ul style="list-style-type: none"> ▪ Poster researched well ▪ Valuable information present 	<ul style="list-style-type: none"> ▪ Extra effort given to research ▪ Accuracy of information clear

Standard Category	Excellence 4	Expected 3	Improving 2	Inadequate 1
Detail	Uses interesting details that add to the overall clarity and organization	Uses details to support main idea	Uses few details	Details are missing or inaccurate
Opinion	Opinions are balanced and made in a point-by-point manner	Opinions are presented and balanced	Few opinions are given	Opinions are not given

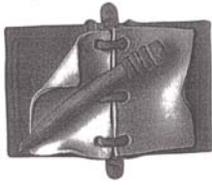
Example 5: Written Presentation Rubric
This excerpt is from a rubric adapted by Don Yackel to score the written portion of integrated presentations at the Thomas O’Brien Academy of Science and Technology in Albany, New York. Other dimensions on the rubric are main idea, vocabulary, voice/style, and global assessment.

Proportion of Text Detail	Misconceptions
<ul style="list-style-type: none"> 0. no response 1. no information from text 2. material from text accounts for about ¼ of the essay 3. material from text accounts for about ½ of the essay 4. material from text accounts for about ¾ of the essay 5. the essay uses or is based on material from the text only 	<ul style="list-style-type: none"> 0. no response 1. one or more serious misconceptions central to the essay 2. at least one serious misconception 3. several minor errors and/or a moderate misconception 4. very minor misconception 5. no misconceptions

Example 6: Essay Scoring Rubric

The example above is from a basic scoring rubric for a complex performance task in social studies. It was developed by staff at the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) at UCLA, and can be found at their Web site (www.cresst96.cse.ucla.edu/CRESST/pages/Rubrics.htm). In addition to the dimensions shown here, the rubric looks at content quality, prior knowledge, argumentation, and the number of principles or concepts included in the essay

For commentary on the six rubric examples, turn to page 30.



Comments on Rubric Examples

Example 1

Content is clearly a key dimension in judging student presentations. This rubric, however, does not explain what constitutes a satisfactory rating in content. How many of the listed features are needed to rate a “3” or “4”? Also, how does the number of slides or the existence of backup disks relate to the content of the project? It would work better to move all but the first two list items into a separate checklist or “project inventory.” Then the rubric designer could elaborate on the criteria used to determine different levels of understanding and critical thinking. An entirely different dimension might be created to address completeness and direction-following, with ratings based on the number of checklist items completed as expected.

Example 2

In spite of the large number of dimensions, this rubric illustrates many of the “do’s” of rubric design. It is general enough to apply to a range of multimedia products. The dimensions are key components of multimedia production. The designers limited themselves to four performance levels. Criteria statements represent a clear and concrete progression. For example, it’s easy to recognize the difference between “collection or rehash of other people’s ideas” and work that “extends beyond that collection to offer new insights.” If this rubric were to be used directly with students, rather than anonymously by a panel of judges, the designers might want to consider rewording such charged phrases as “There is no evidence of new thought or inventiveness.”

Example 3

This is a nice example of involving students in the construction of the rubric. Their influence is clearly evident in the choice of words while the teacher’s professional skill is evident in the structure and clarity of the rubric. Although the rubric was developed for a specific assignment, the wording is general enough that, with minor revisions, the rubric could be applied to other assignments of a similar nature. While it’s not clear what the difference is between “looks real” and “real-looking product,” the general performance levels are differentiated clearly. The teacher might want to push the students to define what they mean by “creative” or “boring”—and substitute such descriptions for the vaguer, more emotion-laden words. It would also make sense to divide this one dimension into two—one focused on oral presentation skills and the other on the appearance/quality of the product created.

Example 4

This is an example of a task-specific rubric; it would only be applied to this particular assignment. While criteria such as “no labels” or “labels unrelated to drawing” are clear and measurable, many of the others are ambiguous. For example, it does not include specific features to differentiate a “carelessly done” drawing from one that is “poorly done.” Similarly, what distinguishes “researched well” from “extra effort given to research?” On another front, the two dimensions shown here are really specifications which would lend themselves well to an assignment checklist. Commensurate dimensions might be “quality of drawing” and “information accuracy.” Also, the teacher might try collapsing the first three performance levels into one.

Example 5

This is a simple, sound rubric that could be used on a range of different assignments. Dimensions represent important aspects of the work and the four performance levels are helpful. While the differences between details that “add to the overall clarity and organization” and those that “support the main idea” could use clarification, most of the other criteria are clear and well differentiated.

Example 6

Both dimensions on this rubric are examples of using cumulative or “countable” features to differentiate performance levels. The scorer simply counts misconceptions or estimates the proportion of the essay that draws directly on material from the assigned text. Assuming the scorer is knowledgeable in the subject matter, it should be easy to distinguish minor from serious misconceptions, the scorer might have trouble choosing between a “1” and “2”, since the distinction seems to hinge on the phrase “central to the essay.”

A former teacher and school principal, Michael Simkins, Ed.D., is now a project director at Joint Venture: Silicon Valley Network in California (www.jointventure.org). His areas of expertise include curriculum design, school reform, staff development and educational technology.