

## An Introduction

Rubrics have become commonplace in 21<sup>st</sup> century high schools. Where educators once were perplexed by the mere mention of these complex, subjective assessment tools, they are now being utilized everywhere from the school assembly to the staff meeting. Although rubrics abound, finding teachers who agree upon what constitutes a strong rubric can be difficult. Therefore, we focused this sub-section around three questions necessary to understanding how rubrics can enhance the curriculum in our small schools in a meaningful and unambiguous way.

### **What are the components of a rubric and why are they important to understand?**

Several resources in this section dissect the anatomy of rubrics because far too many teachers take their existing method of assessment (a grading scale or checklist, for example), transplant it into a chart, call it a rubric, and expect different results. Rubrics have parameters. Rubrics are thorough and substantial. Once the parameters and substance are understood, the challenge of designing the rubric becomes more about aligning the content with the task and less about formatting.

### **What do we need to know and do to create useful rubrics?**

We have established that we need to know the anatomy of a rubric to make it effective, but we must also understand the function of a rubric. When should you use a rubric? What should you measure? How do you determine the effectiveness of a rubric? These questions are addressed throughout the section in practical language meant to demystify the rubric-making process.

The next step is to take this newfound confidence and practice your skills. As you practice, there are tools to evaluate your rubric, tools to help you recognize flaws in rubric design, and tools to help you include your students in the rubric-making process. After honing your ability to recognize flaws in rubric design, using our tool on page 101, you may notice that even well-respected sources succumb to an error or two. It is important to note that these errors do not invalidate the entire assessment, and it is crucial to acknowledge that the ultimate judge of your rubric-making finesse will be how well your students understand it, and how well the rubric assesses the skills that were taught.

Once you have practiced your rubric-making skills you will find that your rubrics can be enhanced by adding student voice to the process. What better way to make sure that your rubrics are understood and on target than to ask students to evaluate and improve upon them? There are resources in this section specific to designing rubrics with students, and examples of some of the impressive, rigorous rubrics students have developed.







### **What role do rubrics play in performance assessment?**

As teachers in small schools incorporate more constructivist, project-based, interdisciplinary teaching methods into their personalized classrooms, assessing students in only traditional pencil and paper ways becomes inadequate. Finding ways to measure and account for student learning in these dynamic classrooms will require a more complex and flexible assessment tool. The rubric succeeds in combining the intricacy of multi-faceted classroom work with the accountability of deliberate, well-planned assessment that truly measures the task at hand.

## **PERFORMANCE ASSESSMENT: RUBRICS**

---

TABLE OF CONTENTS

	<b>The Authentic Assessment Toolbox</b>	<b>89</b>
	<b>Checklist to Evaluate Rubric Designs</b>	<b>100</b>
	<b>Recognizing Flaws in Rubric Design</b>	<b>101</b>
	<b>Creating Rubrics Through Negotiable Contracting</b>	<b>105</b>
	<b>The Necessity of Using Detailed, Descriptive Rubrics</b>	<b>112</b>
	<b>RubiStar Rubric Maker</b>	<b>117</b>

## **PERFORMANCE ASSESSMENT: RUBRICS**

---

**Authentic Assessment Toolbox: Rubrics**

Jon Mueller

<http://jonathan.mueller.faculty.noctrl.edu/toolbox/index.htm>

In clear, easy-to-follow language, Jon Mueller systematically dissects the workings of a rubric and then shows us how to build our own. Filled with practical advice from his years of consulting and teaching about authentic assessment, this piece is a great resource for the mechanics of rubric creation. Because of its length, staff could build in a pre-reading time before a meeting to give this piece a thorough examination. Mr. Mueller also has a wealth of information in his toolbox at his website, which is growing and expanding.

**Rubric:** *A scoring scale used to assess student performance along a task-specific set of criteria*

Authentic assessments typically are criterion-referenced measures. That is, a student's aptitude on a task is determined by matching the student's performance against a set of criteria to determine the degree to which the student's performance meets the criteria for the task. To measure student performance against a pre-determined set of criteria, a rubric, or scoring scale, is typically created which contains the essential criteria for the task and appropriate levels of performance for each criterion. For example, the following rubric (scoring scale) covers the research portion of a project:

**Research Rubric**

Criteria		1	2	3
Number of Sources	x3	1-4	5-9	10-12
Historical Accuracy	x1	Lots of historical inaccuracies	Few inaccuracies	No apparent inaccuracies
Organization	x1	Can not tell from which source information came	Can tell with difficulty where information came from	Can easily tell which sources info was drawn from
Bibliography	x1	Bibliography contains very little information	Bibliography contains most relevant information	All relevant information is included

As in the above example, a rubric is comprised of two components: *criteria* and *levels of performance*. Each rubric has at least two criteria and at least two levels of performance. The criteria, characteristics of good performance on a task, are listed in the left-hand column in the rubric above (number of sources, historical accuracy, organization and bibliography). Actually, as is common in rubrics, the author has used shorthand for each criterion to make

## PERFORMANCE ASSESSMENT: RUBRICS

---

it fit easily into the table. The full criteria are statements of performance such as "include a sufficient number of sources" and "project contains few historical inaccuracies."

For each criterion, the evaluator applying the rubric can determine to what degree the student has met the criterion, i.e., the level of performance. In the above rubric, there are three levels of performance for each criterion. For example, the project can contain lots of historical inaccuracies, few inaccuracies or no inaccuracies.

Finally, the rubric above contains a mechanism for assigning a score to each project. (Assessments and their accompanying rubrics can be used for purposes other than evaluation and, thus, do not have to have points or grades attached to them.) In the second-to-left column a weight is assigned each criterion. Students can receive 1, 2 or 3 points for "number of sources." But historical accuracy, more important in this teacher's mind, is weighted three times (x3) as heavily. So, students can receive 3, 6 or 9 points (i.e., 1, 2 or 3 times 3) for the level of accuracy in their projects.

### Descriptors

The above rubric includes another common, but not a necessary, component of rubrics — *descriptors*. Descriptors spell out what is expected of students at each level of performance for each criterion. In the above example, "lots of historical inaccuracies," "can tell with difficulty where information came from" and "all relevant information is included" are descriptors. A descriptor tells students more precisely what performance looks like at each level and how their work may be distinguished from the work of others for each criterion. Similarly, the descriptors help the teacher more precisely and consistently distinguish between student work.

Many rubrics do not contain descriptors, just the criteria and labels for the different levels of performance. For example, imagine we strip the rubric above of its descriptors and put in labels for each level instead. Here is how it would look:

Criteria		Poor (1)	Good (2)	Excellent (3)
Number of Sources	x1			
Historical Accuracy	x3			
Organization	x1			
Bibliography	x1			

It is not easy to write good descriptors for each level and each criterion. So, when you first construct and use a rubric you might not include descriptors.

That is okay. You might just include the criteria and some type of labels for the levels of performance as in the table above. Once you have used the rubric and identified student work that fits into each level it will become easier to articulate what you mean by "good" or "excellent." Thus, you might add or expand upon descriptors the next time you use the rubric.

### **Why Include Levels of Performance?**

#### **Clearer expectations**

It is very useful for the students and the teacher if the criteria are identified and communicated prior to completion of the task. Students know what is expected of them and teachers know what to look for in student performance. Similarly, students better understand what good (or bad) performance on a task looks like if levels of performance are identified, particularly if descriptors for each level are included.

#### **More consistent and objective assessment**

In addition to better communicating teacher expectations, levels of performance permit the teacher to more consistently and objectively distinguish between good and bad performance, or between superior, mediocre and poor performance, when evaluating student work.

#### **Better feedback**

Furthermore, identifying specific levels of student performance allows the teacher to provide more detailed feedback to students. The teacher and the students can more clearly recognize areas that need improvement.

### **Analytic Versus Holistic Rubrics**

For a particular task you assign students, do you want to be able to assess how well the students perform on each criterion, or do you want to get a more global picture of the students' performance on the entire task? The answer to that question is likely to determine the type of rubric you choose to create or use: Analytic or holistic.

#### **Analytic rubric**

Most rubrics, like the Research rubric above, are *analytic* rubrics. An *analytic rubric* articulates levels of performance for *each* criterion so the teacher can assess student performance on each criterion. Using the Research rubric, a teacher could assess whether a student has done a poor, good or excellent job of "organization" and distinguish that from how well the student did on "historical accuracy."

## PERFORMANCE ASSESSMENT: RUBRICS

---

### Holistic rubric

In contrast, a holistic rubric does *not* list separate levels of performance for each criterion. Instead, a *holistic rubric* assigns a level of performance by assessing performance across multiple criteria as a whole. For example, the analytic research rubric above can be turned into a holistic rubric:

<p>3 - Excellent Researcher</p> <ul style="list-style-type: none"> <li>* included 10-12 sources</li> <li>* no apparent historical inaccuracies</li> <li>* can easily tell which sources information was drawn from</li> <li>* all relevant information is included</li> </ul>
<p>2 - Good Researcher</p> <ul style="list-style-type: none"> <li>* included 5-9 sources</li> <li>* few historical inaccuracies</li> <li>* can tell with difficulty where information came from</li> <li>* bibliography contains most relevant information</li> </ul>
<p>1 - Poor Researcher</p> <ul style="list-style-type: none"> <li>* included 1-4 sources</li> <li>* lots of historical inaccuracies</li> <li>* cannot tell from which source information came</li> <li>* bibliography contains very little information</li> </ul>

In the analytic version of this rubric, 1, 2 or 3 points is awarded for the number of sources the student included. In contrast, number of sources is considered along with historical accuracy and the other criteria in the use of a holistic rubric to arrive at a more global (or holistic) impression of the student work.

### When to choose an analytic rubric

Analytic rubrics are more common because teachers typically want to assess each criterion separately, particularly for assignments that involve a larger number of criteria. It becomes more and more difficult to assign a level of performance in a holistic rubric as the number of criteria increases. For example, what level would you assign a student on the holistic research rubric above if the student included 12 sources, had lots of inaccuracies, did not make it clear from which source information came, and whose bibliography contained most relevant information? As student performance increasingly varies across criteria it becomes more difficult to assign an appropriate holistic category to the performance. Additionally, an analytic rubric better handles weighting of criteria. How would you treat "historical accuracy" as more important a criterion in the holistic rubric? It is not easy. But the analytic rubric handles it well by using a simple multiplier for each criterion.

**When to choose a holistic rubric**

So, when might you use a holistic rubric? Holistic rubrics tend to be used when a quick or gross judgment needs to be made. If the assessment is a minor one, such as a brief homework assignment, it may be sufficient to apply a holistic judgment (e.g., check, check-plus, or no-check) to quickly review student work. But holistic rubrics can also be employed for more substantial assignments. On some tasks it is not easy to evaluate performance on one criterion independently of performance on a different criterion. For example, many writing rubrics are holistic because it is not always easy to disentangle clarity from organization or content from presentation. So, some educators believe a holistic or global assessment of student performance better captures student ability on certain tasks. (Alternatively, if two criteria are nearly inseparable, the combination of the two can be treated as a single criterion in an analytic rubric.)

**How Many Levels of Performance Should I Include in my Rubric?**

There is no specific number of levels a rubric should or should not possess. It will vary depending on the task and your needs. A rubric can have as few as two levels of performance (e.g., a checklist) or as many as ... well, as many as you decide is appropriate. (Some do not consider a checklist a rubric because it only has two levels -- a criterion was met or it wasn't. But because a checklist does contain criteria and at least two levels of performance, I include it under the category of rubrics.) Also, it is *not* true that there must be an even number or an odd number of levels. Again, that will depend on the situation.

To further address how many levels of performance should be included in a rubric, analytic and holistic rubrics will be addressed separately.

**Analytic rubrics**

Generally, it is better to start with a smaller number of levels of performance for a criterion and then expand if necessary. Making distinctions in student performance across two or three broad categories is difficult enough. As the number of levels increases, and those judgments become finer and finer, the likelihood of error increases.

Thus, start small. For example, in an oral presentation rubric, amount of eye contact might be an important criterion. Performance on that criterion could be judged along three levels of performance: **never, sometimes, always.**

<b>makes eye contact with audience</b>	never	sometimes	always
--	-------	-----------	--------

## PERFORMANCE ASSESSMENT: RUBRICS

---

Although these three levels may not capture all the variation in student performance on the criterion, it may be sufficient discrimination for your purposes. Or, at the least, it is a place to start. Upon applying the three levels of performance, you might discover that you can effectively group your students' performance in these three categories. Furthermore, you might discover that the labels of never, sometimes and always sufficiently communicates to your students the degree to which they can improve on making eye contact.

On the other hand, after applying the rubric you might discover that you cannot effectively discriminate among student performance with just three levels of performance. Perhaps, in your view, many students fall in between never and sometimes, or between sometimes and always, and neither label accurately captures their performance. So, at this point, you may decide to expand the number of levels of performance to include never, **rarely**, sometimes, **usually** and always.

<b>makes eye contact</b>	never	rarely	sometimes	usually	always
--------------------------	-------	--------	-----------	---------	--------

There is no "right" answer as to how many levels of performance there should be for a criterion in an analytic rubric; that will depend on the nature of the task assigned, the criteria being evaluated, the students involved and your purposes and preferences. For example, another teacher might decide to leave off the "always" level in the above rubric because "usually" is as much as normally can be expected or even wanted in some instances. Thus, the "makes eye contact" portion of the rubric for that teacher might be

<b>makes eye contact</b>	never	rarely	sometimes	usually
--------------------------	-------	--------	-----------	---------

So, I recommend that you begin with a small number of levels of performance for each criterion, apply the rubric one or more times, and then re-examine the number of levels that best serve your needs. I believe starting small and expanding if necessary is preferable to starting with a larger number of levels and shrinking the number because rubrics with fewer levels of performance are normally

- easier and quicker to administer
- easier to explain to students (and others)
- easier to expand than larger rubrics are to shrink

The fact that rubrics can be modified and can reasonably vary from teacher to teacher again illustrates that rubrics are flexible tools to be shaped to your purposes.

### Holistic rubrics

Much of the advice offered above for analytic rubrics applies to holistic rubrics as well. Start with a small number of categories, particularly since

holistic rubrics often are used for quick judgments on smaller tasks such as homework assignments. For example, you might limit your broad judgments to

- satisfactory
- unsatisfactory
- not attempted

or

- check-plus
- check
- no check

or even just

- satisfactory (check)
- unsatisfactory (no check)

Of course, to aid students in understanding what you mean by "satisfactory" or "unsatisfactory" you would want to include *descriptors* explaining what satisfactory performance on the task looks like.

Even with more elaborate holistic rubrics for more complex tasks I recommend that you begin with a small number of levels of performance. Once you have applied the rubric you can better judge if you need to expand the levels to more effectively capture and communicate variation in student performance.

### **Creating the Rubric**

Once you have identified the criteria you want to look for as indicators of good performance, you next decide whether to consider the criteria analytically or holistically.

### **Creating an Analytic Rubric**

In an *analytic rubric* performance is judged separately for each criterion. Teachers assess how well students meet a criterion on a task, distinguishing between work that effectively meets the criterion and work that does not meet it. The next step in creating a rubric, then, is deciding how fine such a distinction should be made for each criterion. For example, if you are judging the amount of eye contact a presenter made with his/her audience that judgment could be as simple as did or did not make eye contact (two levels of performance), never, sometimes or always made eye contact (three levels), or never, rarely, sometimes, usually, or always made eye contact (five levels).

## PERFORMANCE ASSESSMENT: RUBRICS

---

Generally, it is better to start small with fewer levels because it is usually harder to make more fine distinctions. For eye contact, I might begin with three levels such as never, sometimes and usually. Then if, in applying the rubric, I found that some students seemed to fall in between never and sometimes, and never or sometimes did not adequately describe the students' performance, I could add a fourth (e.g., rarely) and, possibly, a fifth level to the rubric.

In other words, there is some trial and error that must go on to arrive at the most appropriate number of levels for a criterion. (See the Rubric Workshop below to see more detailed decision-making involved in selecting levels of performance for a sample rubric.)

### **Do I need to have the same number of levels of performance for each criterion within a rubric?**

No. You could have five levels of performance for three criteria in a rubric, three levels for two other criteria, and four levels for another criterion, all within the same rubric. Rubrics are very flexible tools. There is no need to force an unnatural judgment of performance just to maintain standardization within the rubric. If one criterion is a simple either/or judgment and another criterion requires finer distinctions, then the rubric can reflect that variation.

### **Do I need to add descriptors to each level of performance?**

No. *Descriptors* are recommended but not required in a rubric. Descriptors are the characteristics of behavior associated with specific levels of performance for specific criteria. For example, in the following portion of an elementary science rubric, the criteria are 1) observations are thorough, 2) predictions are reasonable, and 3) conclusions are based on observations. Labels (limited, acceptable, proficient) for the different levels of performance are also included. Under each label, for each criterion, a descriptor is included to further explain what performance *at that level* looks like.

Criteria	Limited	Acceptable	Proficient
<b>made good observations</b>	observations are absent or vague	most observations are clear and detailed	all observations are clear and detailed
<b>made good predictions</b>	predictions are absent or irrelevant	most predictions are reasonable	all predictions are reasonable
<b>appropriate conclusion</b>	conclusion is absent or inconsistent with observations	conclusion is consistent with most observations	conclusion is consistent with observations

As you can imagine, students will be more certain what is expected to reach each level of performance on the rubric if descriptors are provided. Furthermore, the more detail a teacher provides about what good

performance looks like on a task the better a student can approach the task. Teachers benefit as well when descriptors are included. A teacher is likely to be more objective and consistent when applying a descriptor such as "*most observations are clear and detailed*" than when applying a simple label such as "*acceptable*." Similarly, if more than one teacher is using the same rubric, the specificity of the descriptors increases the chances that multiple teachers will apply the rubric in a similar manner. When a rubric is applied more consistently and objectively it will lead to greater reliability and validity in the results.

**Creating a Holistic Rubric**

In a *holistic rubric*, a judgment of how well someone has performed on a task considers all the criteria together, or holistically, instead of separately as in an analytic rubric. Thus, each level of performance in a holistic rubric reflects behavior across all the criteria. For example, here is a holistic version of the oral presentation rubric above.

**Rubric 6: Oral Presentation (Holistic)**

**Oral Presentation Rubric**

<p>Mastery</p> <ul style="list-style-type: none"> <li>* usually makes eye contact</li> <li>* volume is always appropriate</li> <li>* enthusiasm present throughout presentation</li> <li>* summary is completely accurate</li> </ul>
<p>Proficiency</p> <ul style="list-style-type: none"> <li>* usually makes eye contact</li> <li>* volume is usually appropriate</li> <li>* enthusiasm is present in most of presentation</li> <li>* only one or two errors in summary</li> </ul>
<p>Developing</p> <ul style="list-style-type: none"> <li>* sometimes makes eye contact</li> <li>* volume is sometimes appropriate</li> <li>* occasional enthusiasm in presentation</li> <li>* some errors in summary</li> </ul>
<p>Inadequate</p> <ul style="list-style-type: none"> <li>* never or rarely makes eye contact</li> <li>* volume is inappropriate</li> <li>* rarely shows enthusiasm in presentation</li> <li>* many errors in summary</li> </ul>

## PERFORMANCE ASSESSMENT: RUBRICS

---

An obvious, potential problem with applying the above rubric is that performance often does not fall neatly into categories such as mastery or proficiency. A student might always make eye contact, use appropriate volume regularly, occasionally show enthusiasm and include many errors in the summary. Where you put that student in the holistic rubric? Thus, it is recommended that the use of holistic rubrics be limited to situations when the teacher wants to:

- make a quick, holistic judgment that carries little weight in evaluation, or
- evaluate performance in which the criteria cannot be easily separated.

Quick, holistic judgments are often made for homework problems or journal assignments. To allow the judgment to be quick and to reduce the problem illustrated in the above rubric of fitting the best category to the performance, the number of criteria should be limited. For example, here is a possible holistic rubric for grading homework problems.

### Rubric 7: Homework Problems

Homework Problem Rubric
<p>++ (3 pts.)</p> <ul style="list-style-type: none"> <li>▪ most or all answers correct, AND</li> <li>▪ most or all work shown</li> </ul>
<p>+ (1 pt.)</p> <ul style="list-style-type: none"> <li>▪ at least some answers correct, AND</li> <li>▪ at least some but not most work shown</li> </ul>
<p>- (0 pts.)</p> <ul style="list-style-type: none"> <li>▪ few answers correct, OR</li> <li>▪ little or no work shown</li> </ul>

Although this homework problem rubric only has two criteria and three levels of performance, it is not easy to write such a holistic rubric to accurately capture what an evaluator values *and* to cover all the possible combinations of student performance. For example, what if a student got all the answers correct on a problem assignment but did not show any work? The rubric covers that: the student would receive a (-) because "little or no work was shown." What if a student showed all the work but only got some of the answers correct? That student would receive a (+) according to the rubric. All such combinations are covered. But does giving a (+) for such work reflect what the teacher values? The above rubric is designed to give equal weight to correct answers and work shown. If that is not the teacher's intent then the rubric needs to be changed to fit the goals of the teacher.

All of this complexity with just two criteria -- imagine if a third criterion were added to the rubric. So, with holistic rubrics, limit the number of criteria considered, or consider using an analytic rubric.

### **Final Step: Checking Your Rubric**

As a final check on your rubric, you can do any or all of the following before applying it.

- Let a colleague review it.
- Let your students review it -- is it clear to them?
- Check if it aligns or matches up with your standards.
- Check if it is manageable.
- Consider imaginary student performance on the rubric.

By the last suggestion I mean to imagine that a student had met specific levels of performance on each criterion (for an analytic rubric). Then ask yourself if that performance translates into the score that you think is appropriate. For example, on Rubric 3 above, imagine a student scores


- "sometimes" for eye contact (3 pts.)
- "always" for volume (4 pts.)
- "always" for enthusiasm (4 pts.)
- "sometimes" for summary is accurate (4 pts.)

That student would receive a score of 15 points out of a possible 20 points. Does 75% (15 out of 20) capture that performance for you? Perhaps you think a student should not receive that high of a score with only "sometimes" for the summary. You can adjust for that by increasing the weight you assign that criterion. Or, imagine a student apparently put a lot of work into the homework problems but got few of them correct. Do you think that student should receive some credit? Then you would need to adjust the holistic homework problem rubric above. In other words, it can be very helpful to play out a variety of performance combinations before you actually administer the rubric. It helps you see the forest through the trees.

Of course, you will never know if you really have a good rubric until you apply it. So, do not work to perfect the rubric before you administer it. Get it in good shape and then try it. Find out what needs to be modified and make the appropriate changes.

**Checklist to Evaluate Rubric Designs**

The following checklist to evaluate rubric design was developed based on a conversation with Dennie Palmer Wolf, Director of the Rethinking Accountability Initiative at the Annenberg Institute for School Reform at Brown University. Palmer Wolf has been teaching, consulting, and doing research in education since 1971. See *Recognizing Flaws in Rubric Design* (page 101) for a more thorough explanation of checklist criteria.

<b>Checklist to evaluate rubric designs:</b>		
1	DO NOT confuse a higher quality rubric with a higher number of criteria; 3-7 is an optimal amount.	
2	DO NOT include as a central part of a rubric something that translates into a quantifiable term.	
3	DO NOT use the first two levels of the rubric to describe what a student does wrong and the second two levels to describe what she does right.	
4	DO use the first level as a standard base, not as an incomplete or failing level.	
5	DO create levels of performance that advance in equal increments, representing realistic increases in students' skill level.	
6	DO align rubric language with everything else students are doing in their educational life, creating a common language and standard.	
7	DO align rubric criteria with national/state/school standards.	
8	DO involve kids in deciding particular performance tasks.	
9	DO make the rubric understandable to students.	
10	DO use terminology that is understandable to an outside evaluator.	

**Recognizing Flaws in Rubric Design**

Using the ten criteria from the Checklist to Evaluate Rubric Design, the following resource is meant to give teachers examples of common mistakes made when creating rubrics. Some of the criteria are self-explanatory and are not accompanied by examples. The others have pieces of rubrics where the trouble spots occur, followed by an explanation of where the faulty rubric went wrong.

**One:** *DO NOT confuse a higher quality rubric with a higher number of criteria; 3-7 is an optimal amount.*

**First-Draft Essay Rubric**

Criteria:	4	3	2	1
Grammar				
Punctuation				
Neatness				
Source Documentation				
Ideas				
Organization				
Voice				
Word Choice				
Sentence Fluency				
Presentation				
Timeliness				
Length				

**What went wrong?**

This rubric uses twelve criteria to evaluate a first draft of an essay. The suggested number of criteria for a rubric is between 3 and 7. The rubric designer might choose to focus on two or three closely related criteria (ex. voice, word choice, sentence fluency) for the first draft, and then choose another set of criteria for the second draft. The student will understand the set of skills more thoroughly and will not be overwhelmed with the quantity of expectations for the first draft.

**Two:** *DO NOT include as a central part of a rubric something that translates into a quantifiable term.*

**Research Paper Rubric**

Criteria:	4	3	2	1
Source Documentation	Correctly cites 10 or more sources	Correctly cites at least 7 sources	Cites at least 5 sources	Cites less than 5 sources



TOOL

**What went wrong?**

Rubrics are meant to assess criteria that are complex and subjective. If a rubric uses quantities or quantifiable terms (such as rarely, occasionally, frequently, and always) the rubric becomes just as unhelpful as a traditional grading scale at assessing dynamic work. When designing a rubric, it is important to decide what skills are to be assessed and what would be better assessed using a checklist or grading scale.

The example above could be altered to assess the *quality* and *use* of resources in the research paper. For example, a student may have three credible resources that she cites to support her thesis, while another student may have twelve questionable resources that do not support his thesis. A rubric addressing the nuances of individual student work will be far more beneficial to improving skills than categorizing with arbitrary numbers.

**Three:** *DO NOT use the first two levels of the rubric to describe what is wrong and the second two levels to describe what is right.*

**Group Work Rubric**

Criteria:	4	3	2	1
Contribution & Completion of Tasks	Participates actively	Participates in group	<i>Participates inconsistently in group</i>	<i>Does not participate</i>
	Models concern for goals	Shows concern for goals	<i>Shows some concern for goals</i>	<i>Shows no concern for group goals</i>
	Helps lead the group in goal-setting	Participates in goal-setting	<i>Participates sporadically in goal-setting</i>	<i>Impedes goal-setting</i>
	Helps lead the group in meeting goals	Participates in meeting goals	<i>Participates sometimes in meeting goals</i>	<i>Impedes group meeting goals</i>
	Thoroughly completes assigned task	Completes assigned tasks	<i>Completes assigned tasks</i>	<i>Does not complete assigned task</i>

**What went wrong?**

This rubric draws a not-so-subtle distinction with the language used to score a student’s performance as a 1-2 and a 3-4. The rubric draws a line between acceptable and unacceptable work, rather than develop performance levels using 1 as a base of acceptability. Telling a student what she *can’t* do is not motivational. Using phrases such as “can improve upon,” “beginning,” or “building the foundation” gives students a goal to work toward and does not stigmatize students that may score a 1.

**Four:** *DO use first level as a standard base, not as an incomplete or failing level.*

**Math Skills Rubric**

Criteria:	4	3	2	1
Skill Fluency				Unable to carry out basic algebraic and geometric procedures
Modeling and Data Analysis				Unable to accurately formulate the problem or task
				Unable to develop reasonable model
				Does not explain how the model relates to the problem
Problem-Solving				Unable to plan and solve non-routine problem
				Unwilling to complete problem

**What went wrong?**

Every rubric should have a body of student work to draw on so that students recognize quality work and common sticking points. Therefore, the first level of a rubric should be a teaching tool, a base, a springboard representing where many students may be at the beginning of a new school year, a semester, or when learning an unfamiliar skill. It is quite possible that an entire class may strive to move from the first level to the second level over the course of a lesson or unit. If student work falls below the standard base set for the first level, providing a checklist with specific improvements would be more appropriate than a list of failures outlined in a rubric.

**Five:** *DO create levels of performance that advance in equal increments, representing realistic increases in students' skill level.*

**Essay Rubric**

Criteria:	4	3	2	1
<i>Sentence Fluency</i>	Excellent transitions; maturity in sentence structure beyond "the norm"; exceptional "flow."	Highly skilled style, numerous variations in sentence structure; developing "wow factor."	Feels "sticky" at certain points; doesn't peak the reader's interest.	Doesn't use transitions properly (or not at all); juvenile style; no variety in structure.

**What went wrong?**

Unless the class has spent time developing a shared understanding of what constitutes "flow," "the norm," the "wow factor," "sticky," "juvenile," and "maturity," a student would be at a disadvantage for knowing how to improve their skill level, using this rubric as a guide. A student who received a score of 3 would struggle to know if they used transitions adequately or varied their sentence structure appropriately, as there is no mention of either



## PERFORMANCE ASSESSMENT: RUBRICS

in the performance level. Not only are the performance levels not written on an incremental gradient, but also they don't contain the same performance indicators.

**Six:** *DO align rubric language with everything else students are doing in their educational life, creating a common language and standard.*

**Seven:** *DO align rubric criteria with national/state/school standards.*

**Eight:** *DO involve kids in deciding particular performance tasks.*

**Nine:** *DO make rubric understandable to kids.*

**Ten:** *DO use terminology that is understandable to an outside evaluator.*

**Research Paper Presentation**

Criteria:	4	3	2	1
Content	Thoroughly documents type of methodology, data collection methods, and protocols			
Reflection		Adequately uses metacognition and can articulate use of Gardner's multiple intelligences in work		
Connections			Uses at least 3 of Maslow's hierarchy of needs to support thesis	

### What went wrong?

Outside evaluators are a valuable resource because they offer validity to assessment, open the school up to the community, and give students a chance to defend their work for a high-stakes audience. We need to keep our audience in mind when writing rubrics. Without a thorough briefing on research methodology and education theory, a guest evaluator would have little chance of being able to score a student's presentation using this rubric. What may be commonplace to teachers and students could be unfamiliar jargon to the outside evaluator. Be cautious when using acronyms, as well.

**Creating Rubrics Through Negotiable Contracting  
and Assessment**

Andi Stix, Ed.D. (1997)

U.S. Department of Education

Published by ERIC Clearinghouse

With an emphasis being placed on designing rubrics that are understandable to students as well as evaluators, an emerging trend makes a lot of sense—let students assist in creating rubrics and they will be more invested in and accountable for their use. This article demonstrates the critical thinking and imagination that students use when they are given a choice and a voice in assessment development.

What would happen if students were invited to help decide how their work should be evaluated? Would they exploit the opportunity, designing standards so ridiculously low as to guarantee a glut of effortless good grades?

Surprisingly, the answer is no. Experience at Robert Wagner Middle School in Manhattan shows that students who are given a role in the assessment process can and do rise to the occasion. Given the appropriate direction by their teachers, youngsters are able to accurately evaluate their strengths and weaknesses and pinpoint where to focus their efforts to get the most out of what they're learning. As a result, students view assessment not as an arbitrary form of reward or humiliation (a common perception of middle school students), but as a positive tool for personal growth.

This article examines “negotiable contracting,” a new approach to involving students in the assessment process that currently is being implemented in some schools in the New York City area. Negotiable contracting is adaptable to both arts and science curriculum and is flexible enough to accommodate multi-modal forms of learning. Like any assessment, it ensures that the teacher remains squarely in charge of the classroom and, ultimately, responsible for assuring that grading is appropriate.

**Empowering Students**

The art of negotiable contracting consists of giving students shared ownership in their own learning (Wiggins 1993). Although he is ultimately responsible for grading, the teacher functions not as an all-powerful judge of students' work but as a facilitator of discussion on the assessment process (Seeley 1994). Before the teacher presents his or her own expectations of the work, (s)he asks students their opinion of what they think would constitute quality work. Across the “negotiating table,” teacher and class arrive at a consensus that is mutually acceptable. The result is that students feel like valued participants in the assessment process. Thus, they are motivated to strive toward those criteria-based standards.

The contract process can be used independently of a formal evaluation and can serve a variety of purposes. Some lessons do not call for formal assessment. However, the teacher still wants to set short-term goals by establishing criteria for high-quality work. Negotiable contracting is ideal for such a lesson. For example, if students are to work together in groups, negotiable contracting is helpful in setting up expectations such as cooperative roles, research materials and formats for charts and graphs.

### **Creating the Rubric**

The rubric is an important element of using negotiable contracting for formal assessment. (Pate, Homestead, and McGinnis 1993). A rubric is a carefully designed ratings chart that is drawn up jointly by teacher and students. Along one side of the rubric are listed the criteria that the teacher and students decide are the most important ideas to be mastered in the lesson. Across the top of the rubric are listed the rankings that will be used to assess how well students understand each of those criterion. The rubric also indicates how much importance should be given to each criterion, based on its importance to the overall lesson. Within each ranking, there also may be numerical gradations, depending on whether a student performs on the higher or lower level of that category. Unlike a traditionally assigned, generalized number or letter grade, the rubric serves as an in-depth “report card” for a lesson, unit, or project.

Let’s take as an example a social studies teacher at Robert Wagner Middle School. Mrs. Martha Polin, who assigned her students the task of creating a mural for a geography lesson. Before they began any work on the murals, she arranged the class in cooperative learning groups and asked them to consider, “If you were me, what qualities would you look for in deciding how to grade each mural? Come up with six criteria that you would look for.” After allowing time for discussion, Mrs. Polin asked each group to rank the qualities they had selected in order of importance, from most important to least important.

Next, each group presented its top two criteria to the class. Mrs. Polin listed those criteria on the board and the class was asked to choose which ones were truly most relevant to the lesson. With the teacher’s guidance, they agreed on three qualities: 1) detail and depth; 2) a clear focal point; and 3) high-quality design. They then were asked, “What should be considered ‘poor,’ ‘fair,’ ‘good’ and ‘excellent’ performance for each criterion?” One student suggests that a poor mural would have most of the facts wrong, and the other students readily agree. “What about if only some of the facts are wrong?” Mrs. Polin asks. “That would be a fair grade,” says one boy. “I think having some of the facts wrong should still be a poor grade,” argued another student. Finally, after some more discussion, a consensus is reached among the class that getting only some of the facts wrong would earn a “fair” grade. After more discussion, they also decide that getting all the facts right should earn a “good” grade while getting an exceptional amount of accurate, interesting information from unusual sources would earn a rating of excellent.

As a result of their negotiations, before they've even picked up a pencil or pen, Mrs. Polin's students are perfectly clear about what is expected in their murals. Moreover, they have the satisfaction of having had a voice in setting the objectives for the project and establishing a ratings system that they consider to be fair.

<b>Criteria:</b>
Accurate Detail and Depth
Clear Focal Point
High-Quality Design

The next step in creating a rubric is to negotiate ratings to reflect how well each of these criterion are met. Across the top of the rubric chart are listed the various rankings, in lieu of grades or numbers. Again, those rankings may be decided during negotiations between teacher and class. There is a separate rating for each of the criteria in the rubric, since students naturally will be stronger in some aspects of their work than in others.

Choosing neutral words for each rating avoids the implication of good/bad inherent in a generalized A-F or numerical grade. In addition, the natural temptation of instructors -- as well as students -- to award a middle ranking is avoided by the use of an even number of rankings. For example, in a 1-5 ranking system, 3 tends to be used as a "neutral" grade.

Attempted	Acceptable	Admirable	Awesome
Larvae	Pupa	Cocoon	Butterfly
Peasant	Artisan	Noble	Pharaoh
Page	Squire	Knight	Lord
Sour Milk	Milk	Half and Half	Cream
Jeans	Sport Jacket	Suit	Black Tie
Amateur	College Athlete	Semi-Professional	Professional
Byte	Kilobyte	Megabyte	Gigabyte
Private	Sergeant	Lieutenant	Captain
Pepper	Cinnamon	Nutmeg	Saffron
Ground Round	London Broil	Sirloin	Filet Mignon
Bob Cat	Panther	Tiger	Lion
Minnow	Flounder	Tuna	Shark

## PERFORMANCE ASSESSMENT: RUBRICS

Weed	Daisy	Rose	Orchid
------	-------	------	--------

First Base	Second Base	Third Base	Home Run
------------	-------------	------------	----------

Garden	Rattle	Cobra	Python
--------	--------	-------	--------

The State of Kentucky, which uses a rubric system of assessment, utilizes four non-pejorative ratings in its rubrics. In ascending order of competence, they are:

Novice	Apprentice	Proficient	Distinguished
--------	------------	------------	---------------

or

Novice	Apprentice	Veteran	Master
--------	------------	---------	--------

There is no “overall” rating for the child; the terms are used separately to evaluate students’ performance for each of the criterion in the rubric. For a social studies report, for example, the ratings might be defined as follows:

- “Novice” is a student who has absorbed little of the lesson; it signals insufficient preparation, weak conclusions or organization and incorrect information.
- “Apprentice” implies a beginning conceptual understanding; there is a main idea but it is presented only in broad outline with little detail and some erroneous or unclear information.
- “Proficient” signals a clear conceptual understanding of the lesson; the report was well organized, logical, and focused with few errors.
- “Distinguished” means outstanding work; work rich in depth and precise detail with a consistent, powerful presentation and little to no errors.

It is useful to include numeral gradations within each category. For example, a student may receive an Apprentice rating of three or four, depending on whether he performs on a higher or lower end of that category.

Let’s examine how Mrs. Polin’s class created the rubric for their geography mural.

MURAL	Novice	Apprentice	Veteran	Master
<b>Accurate Detail and Depth</b>	Incorrect or little facts, hardly and detail (1-3pts.)	Some facts are accurate, some detail (4-6pts.)	Substantial amount of facts, good amount of detail (7-9pts.)	Exceptional amount of facts, vivid descriptions (10-12pts.)

<b>Clear Focus</b>	Vague and unclear (1-2 pts.)	Some focus, but not organized enough (3-4 pts.)	Well organized and clearly presented (5-6 pts.)	Highly organized and easy to follow (7-8 pts.)
<b>Design</b>	Little to no layout and design (1-3 pts.)	Simple design, but layout could be more organized (4-6 pts.)	Attractive and invites the viewer (7-9 pts.)	Exceptional design and outstanding visual appeal (10-12 pts.)

### Creative Problem Solving

Rubrics can be especially effective in assessing student’s work in mathematics (Moon 1993). While rote skills such as memorizing the time tables may be best suited to traditional quizzing and grading, the majority of mathematics really involves creative problem solving in which there are several ways to arrive at a solution -- some more succinct, effective, or creative than others.

For a lesson involving word problems in fractions, for example, the “report card” for students’ problem-solving might include an assessment criteria decided upon by teacher and students: Is the solution easy to follow? Does it demonstrate clear conceptual understanding? Would the answer work in real life? Do the diagrams, sentences, and number coordinate?

Similarly, rubrics can be used in any discipline-based or interdisciplinary lesson. The rubric can include opportunities for students to use journal work, projects, research studies, experiments, skits, or other vehicles to demonstrate their competence.

### Good Poetry

Let’s look at how a rubric would be utilized in Mrs. Janine Bartko’s 8th grade Language Arts class, which is studying a unit on poetry. After discussing how poetry differs from prose and looking at various types of poetry, the students are given the assignment of writing a poem of their own. Mrs. Bartko then asks: “How can a poem—a subjective assignment with no ‘correct’ answer—be fairly assessed?”

The students launch into a discussion of what constitutes “good” poetry, as they were asked to write a piece reflecting a time in history. Working in groups, they come up with a rubric, composed of four main criteria that Mrs. Bartko and the students agree are the most appropriate and fair qualities.

**PERFORMANCE ASSESSMENT: RUBRICS**

They decide a poem should portray emotion and/or imagery; captivate the reader; use language clearly; and use punctuation purposefully. Mrs. Bartko and her students then read various examples of how those skills are applied at the various ratings levels. Finally, before filling in the rubrics with her students as a whole group, she asks the youngsters—sitting in cooperative work groups—to try to evaluate the assignment and fill in the rubric on their own:

<b>Poetry</b>	<b>Novice</b>	<b>Apprentice</b>	<b>Veteran</b>	<b>Master</b>
<b>Ability to Captivate the Reader</b>	Unfocused; author seems unsure of direction (1-2 pts.)	Some focus, but lacks continuity (3-4 pts.)	Well-focused and interests reader throughout (5-6 pts.)	Captivates and involves reader deeply (7-8 pts.)
<b>Sensory Images</b>	Difficult to visualize image or emotion (1-3 pts.)	Some use of image, idea, or emotion (4-6 pts.)	Clear use of sensory images to portray ideas or emotions (7-9 pts.)	Vivid, detailed images and intensely felt emotion (10-12 pts.)
<b>Use of Language</b>	Imprecise or inappropriate choice of words (1-2 pts.)	Expresses thoughts marginally (3-4 pts.)	Appropriate choice of language (5-6 pts.)	Uses rich and imaginative language (7-8 pts.)
<b>Punctuation</b>	Arbitrary punctuation (1-2 pts.)	Some meaningful punctuation (3-4 pts.)	Punctuation meaningful throughout (5-6 pts.)	Punctuation enhances conveyance of thoughts and images (7-8 pts.)

In addition to the rubric itself, there is an area included for comments. In this space, Mrs. Bartko can be even more specific about strengths and weaknesses and, accentuating the positive, suggest ways for each student to stretch his or her skills and expand his or her understanding. As a result, the rubric gives the student an overall picture of his or her skill level.

At Robert Wagner Middle School, some teachers have enlarged a blank rubric and laminated it. For each project, they use a dry erase marker and fill in the quadrants with the students. They do likewise with the assessment sheet. Students are each given a blank sheet and asked to fill it out with the teacher. Here, students have their own record of what is expected of them. At the end of the project, they may be asked to assess themselves and/or their peers and hand in the assessment sheet for the teacher to grade.

### Recognizing Achievements

Rubrics thus offer an important way for educators to motivate students through assessment. Giving youngsters a voice in their grading provides them with a clear understanding of what is expected of them and the assurance that their accomplishments will be recognized.

### References

- Aschbacher, P. R.** (1993). *Issues in innovative assessment for classroom practice; Barriers and facilitators* (Tech. Rep. No. 359) Los Angeles: University of California, CRESST; Center for the Study of Evaluation.
- Feuer, M. J., & Fulton, K.** (1993) The many faces of performance assessment. *Phi Delta Kappan*, 74:6, 478.
- Jamentz, Kate.** (3/94). Making sure that assessment improves performance. *Educational Leadership*. 51:6, pp. 57.
- Levine, D. S.** (1992). The four p's of context-based assessment: Evaluating literacy across the curriculum. In C. Hedley, P. Antonacci, and D. Feldman (Eds.) *Literacy across the curriculum* Norwood, NJ: Ablex.
- Meyer, Carol A.** (1992) What's the difference between authentic and performance assessment? *Educational Leadership*, 49:8, 39-40.
- Moon, C. J.** (1993). Connecting learning and teaching through assessment. *Arithmetic Teacher*, Vol. 41, No. 1, 13-15.
- O'Neil, John** (1992) Putting performance assessment to the test. *Educational Leadership*, 49:8, 14-19.
- Pate, P. E., Homestead, E., McGinnis, K.** (1993, November). Designing rubrics for authentic assessment. *Middle School Journal*.
- Powell, Janet.** (1993, November). What does it mean to have authentic assessment? *Middle School Journal*.
- Seeley, Marcia M.** (1994, October). The mismatch between assessment and grading. *Educational Leadership*. 52:2.
- Simmons, W., & Resnick, L.** (Feb. 1993) Assessment as the catalyst of school reform. *Educational Leadership*, 50:5, 11-15.
- Tierney, R. J.** (1992, September). Setting a new agenda for assessment. *Learning*. pp. 62-64.
- Wiggins, Grant,** (1993, November). Assessment: authenticity, context and validity. *Phi Delta Kappan*, 75:3, p. 200-214.

### The Necessity of Using Detailed, Descriptive Rubrics

Excerpted from *Getting the Culminating Project Right*

Michael Katims, Ph.D., and Eeva Reeder



The following is a short description of a team of educators who have taken their expertise in the field and created a professional development opportunity for Washington schools by “helping schools prepare young minds for new challenges.” Their argument for, and examples of detailed, descriptive instructional rubrics will be an asset to every school staff in the process of developing performance assessment tasks and rubrics.

Education Redesign has been working with schools and districts in Washington State for four years to help educators take on the challenge of implementing the new graduation requirements. Education Redesign consults schools on developing approaches for culminating projects; state learning goal competencies, authentic and product/performance-based assessment, high-quality instructional rubrics, project-based learning, and principles/processes of the *Understanding by Design* model.

Current/recent projects of Education Redesign:

- Everett School District *Culminating Exhibition*
- Everett School District *Middle School Competency Project*
- Northshore School District *9<sup>th</sup>-Grade Culminating Exhibition Pilot Project*
- Shorewood High School *Learning Exhibition*
- Shoreline School District *Applied-Learning Standards (K-6 report card & project rubric)*
- Mountlake Terrace High School *Smaller Learning Communities*

#### THE NECESSITY OF USING DETAILED, DESCRIPTIVE RUBRICS

An important criterion in evaluating assessment systems is the accuracy and reliability of the judgments that assessors make about the quality of student work. Do the assessors agree on what constitutes adequate work, can they agree on whether or not student work meets the standard? There must be assessment criteria related to meaningful standards, and these criteria must be specified so clearly that (1) assessors agree on decisions about the quality of the student work compared to the standards, and (2) these decisions can be defended and explained to the satisfaction of those being assessed. This is particularly important in the case of high-stakes decision-making, such as the graduation decisions to be made on the basis of the quality of students' CP work.

Our instructional rubrics achieve these assessment goals. But our experience has shown that most CP rubrics are not adequate to support reliable assessment decisions. The typical scoring guide is too vague in its descriptions of the characteristics of student work being assessed, and the

relative importance of different aspects of the work being considered is unclear. As a result, these scoring guides promote consensus in scoring primarily by lowering the standards.

No one is deliberately lowering the standard of acceptable performance. Rather it is an inevitable outcome of the natural reluctance to label students' work as unacceptable, particularly in light of the high-stakes decision to be made. As a result, judges or assessors are inclined to assess students' work as meeting standard unless the scoring guide explicitly describes work similar to that produced by the student and labels it as unacceptable. Unfortunately, most scoring guides do not provide sufficiently detailed descriptions of work to support negative decisions around a meaningful standard.

Consider an example of a rubric currently being used in a Washington State school district. The rubric consists of two parts, one to assess the project itself and another to use in assessing the student's presentation of the project. This rubric is typical of those we have encountered and illustrates the weaknesses of rubrics designed around concepts of scoring/grading. A comparison between the information provided to assessors by this rubric and the information provided by one of our instructional rubrics reveals the problem and a solution.

To help evaluate a student's presentation skills, the example rubric identifies a number of factors related to successful presentation and defines three levels of performance: Exceptional/pass, Adequate/pass, and Insufficient—to be redone.

- **Exceptional:** Highly poised; consistent eye contact; interesting; appropriate tone/language; effective use of visuals enhances presentation; high "WOW!" factor
- **Adequate:** Generally poised, but without the sophistication of the "exceptional" presentation; consistent eye contact, but does not use it to draw in the audience; student has made an effort to make the presentation interesting; appropriate tone/language; visuals don't enhance the presentation as well as they could; low "Wow!" factor
- **Insufficient:** Inappropriately informal, unrehearsed or disorganized; no, or very little eye contact; student has made little effort to make the presentation interesting; inappropriate tone/language (swearing, slang); no "Wow!" factor

Imagine having to make judgments about students' presentations with this information as a guide. What is the difference between "highly poised," and "poised, but without the sophistication of an 'exceptional' presentation?" Don't we all have to agree on what an exceptional presentation is before this rubric has meaning? How does one use eye contact to draw in an audience? What does this look like, and how is it different from other eye contact? How can we tell if the student has attempted to make a presentation interesting to the audience? Is the use of slang and swear words the only aspect that distinguishes among different levels of tone/language? How could the visuals



be made to enhance the presentation as well as they could? What is the “Wow!” factor, and is it the same for everyone? How do you distinguish between high and low “wow?”



These are some of the dilemmas that judges using this rubric will face when attempting to assess student presentations. But certainly the purpose of a scoring rubric is to resolve dilemmas, not to create them. An effective instructional rubric should provide information to help guide assessors’ perceptions of student performance and help them discriminate among performances of different quality. Absent this information, judges are most likely to resolve this dilemma by being very conservative about giving scores outside of the very broad, comfortable category of “adequate.” A few clearly superior presentations will be classified as “exceptional,” but by far the vast majority of presentations will be judged “adequate.” Likely the only presentations judged “insufficient” will be those of clearly poor quality, as defined in the very lowest level of this rubric.

The danger with adopting such a rubric is that the *acceptable* standard for student presentations will be lowered to the upper edge of unacceptable! When this happens, the students whose work will be judged as unacceptable are typically only the ones who did not care enough to complete or refine their work. Is this what we want “meeting standard” to mean – that a mediocre effort equates to competence, irrespective of the quality or skill revealed in the work? In this particular case, if the CP presentation is evidence of State Learning Goal 1, or competence in communicating “effectively and responsibly,” do we really want to say that students’ oral communication ability needs to be only slightly better than unacceptable to meet graduation standard? Or is it our intent that “meeting standard” signifies the student has the ability to do competent work?

In this case, the assessment tool does not provide the information assessors need to adequately discriminate among student performances, except at a very gross level. Now consider the difference in information provided to assessors about very similar oral presentation criteria in the instructional rubric below. Note that these are the middle three levels of a five-level rubric that also describes truly outstanding performance as well as clearly poor performance. We designed this rubric with a district to help judges make difficult decisions around the middle range of performances — to help judges recognize work that has some elements of quality but that does not yet meet the standard of acceptable performance.

<p><b>4:</b> <b>Above Standard</b></p>	<ul style="list-style-type: none"> <li>• Relaxed but confident posture/movement, consistent and effective use of eye contact and gestures, and enthusiasm, all show that the student is prepared to speak on the topic</li> <li>• Clear articulation, good use of variation in volume/tone, moderate pace/timing with pauses used to focus audience attention, all maintain audience interest/engagement</li> <li>• Well-designed audio-visual aids are coordinated with and enhance the oral presentation; they help the audience both to stay engaged and to better understand</li> </ul>
<p><b>3:</b> <b>At Standard</b></p>	<ul style="list-style-type: none"> <li>• Good posture, movement, and use of gestures, with minimal signs of nervousness; eye contact with minimal note reading, some enthusiasm, all indicate the student is ready to speak on the topic</li> <li>• Calm, comfortable delivery, with good articulation; adequate use of volume/tone (minimal monotone); pace/timing varied to maintain audience interest; occasionally the pace may be a little too slow/fast, or articulation may be difficult to understand</li> <li>• Audio-visual aids are integrated with the oral presentation; they add to the presentation and increase audience attention; overall visual appeal could be improved</li> </ul>
<p><b>2:</b> <b>Below Standard</b></p>	<ul style="list-style-type: none"> <li>• Distracting body language, eye contact broken by much note reading, an absence of enthusiasm, all indicate a discomfort in speaking</li> <li>• Nervous delivery with somewhat monotone articulation, some mumbling; uneven, hurried or slow pace, and/or poor timing limits audience engagement</li> <li>• Audio-visual aids both add to and detract from the presentation; they may be: difficult to understand, too overpowering, or too difficult to see/hear</li> </ul>



## PERFORMANCE ASSESSMENT: RUBRICS

---



It is easy to see at a glance that these statements are lengthier, more elaborated, than the descriptors from the sample rubric, including detailed descriptions of different aspects of student performance (e.g. body language, vocal delivery/pace, audio-visual support). They give judges clear images of what to look for in evaluating the quality of the performance. For example, it is certainly simpler for judges to agree on the student's use of eye contact once they place that behavior in the broader context of body language and when they compare what they observe in the student's performance to the descriptive statements in the rubric.

When a scoring rubric gives judges this kind of more detailed, descriptive information, it helps them know what to look for and enables them to classify what they see in a student product or performance. Then there is a high probability of substantial agreement among judges across the entire range of performances. There is agreement among judges with a weak rubric, but it is because no difficult decisions are made about the quality of student work.

In brief, there will be no meaningful standards for the CP without rubrics that provide the descriptive detail judges need to make and defend decisions about the quality of student work. Either there will be no meaningful standards because there is no agreement about what the standards look like in student work/performances, or because the standards are so low as to be meaningless. High quality, instructional rubrics can solve this problem.

"To repeat what others have said, requires education; to challenge it, requires brains."

—Mary Pettibone Poole, *A Glass Eye at a Keyhole* (1938)

### RubiStar Rubric Maker

<http://rubistar.4teachers.org/index.php>

RubiStar is a tool to help the teacher who wants to use rubrics but does not have the time to develop them from scratch. Although formulaic due to technological constraints, it is a good way for beginning rubric-designers to get a feel for how the pieces of a rubric come together to create the whole.

#### What is RubiStar?

While many teachers want to use rubrics or are experimenting with writing rubrics, they can be quite time-consuming to develop. RubiStar is a tool to help the teacher who wants to use rubrics but does not have the time to develop them from scratch. RubiStar provides generic rubrics that can simply be printed and used for many typical projects and research assignments. The unique thing about RubiStar, however, is that it provides these generic rubrics in a format that can be customized. The teacher can change almost all suggested text in the rubric to make it fit their own project.

For example, if RubiStar suggests for a multimedia presentation that "The student includes at least three slides" for the highest rating, the teacher could change that to read "The student includes at least 5 slides" or "The student includes a title slide with the authors' names; a table of contents with links to all slides that follow, ..."

RubiStar also does away with the tedious typing of similar content across all the various quality ratings. When you choose a category to evaluate, all the quality ratings are filled in and are ready to customize.

RubiStar rubrics allow teachers to store their rubrics in the system, post them online to be accessed by students when needed, or download them using Microsoft Excel.

RubiStar was developed by the High Plains Regional Technology in Education Consortium, part of the Regional Technology in Education Consortia (R\*TEC) program established to help states, local educational agencies, teachers, school library and media personnel, administrators, and other education entities successfully integrate technologies into kindergarten through 12th grade (K-12) classrooms, library media centers, and other educational settings, including adult literacy centers.

In providing such help the Regional Technology in Education Consortia establish and conduct regional activities that address professional development, technical assistance, and information resource dissemination to promote the effective use of technology in education with special emphasis on meeting the documented needs of educators and learners in the region they serve.

## PERFORMANCE ASSESSMENT: RUBRICS

---

The U.S. Department of Education, through its Office of Educational Research and Improvement, issued a competitive Request for Proposals that resulted in funding ten Regional Technology in Education Consortia.

### What rubrics are available?

RubiStar has the following rubrics, categorized by type or subject:

Oral Projects: Class Debate, Historical Role Play, Interview, Making a Map, Oral Presentation, Puppet Show, Story Telling, Video Production

Products: Collection or Display, Making A Brochure, Making A Game, Making A Poster, Newspaper, Public Awareness Campaign, Timeline

Multimedia: Digital Storytelling, HyperStudio/Powerpoint Appearance and Content, Multimedia Project, Digital Storytelling, Storyboard, Web Site Design

Science & Math: Building A Structure, Lab Report, Science Experiment, Graphing, Problem-Solving

Research & Writing: 6 + 1 Trait Writing Model, Group Planning — Research Project, Letter Writing, Research Report, Story Writing

Art: Analysis of A Work of Art, Art History, Stage Design

Work Skills: Collaborative Work Skills

Music: Classical Concert Etiquette, Instrumental Music Performance – Individual

### What does a RubiStar rubric look like?

The rubric on the following page is a sample rubric created using RubiStar. The rubric took less than five minutes to make and download. RubiStar uses a set number of four performance levels, which are labeled 1, 2, 3, 4 on the template, and can be customized by the user. The criteria are set by using a pull-down menu listing all of the options available from that category. For example, for the interview rubric below, the user could have chosen additional criteria such as:

- Report writing
- Labeling
- Videography
- Sound quality
- Notetaking

As soon as a criterion is selected, the descriptors at each performance level are automatically filled in. The user can also add their own criteria and descriptors.

## PERFORMANCE ASSESSMENT: RUBRICS

Rubric Made Using: RubiStar ( <a href="http://rubistar.4teachers.org">http://rubistar.4teachers.org</a> )				
<b>Interview : Elders Interview</b>				
Teacher Name: Student Name:				
CATEGORY	cruising	advancing	emerging	beginning
Setting Up the Interview	The student introduced himself, explained why he wanted to interview the person, and asked permission to set up a time for an interview.	The student introduced himself and asked permission to set up a time for the interview, but needed a reminder to explain why he wanted to do the interview.	The student asked permission to set up a time for the interview, but needed reminders to introduce himself and to tell why he wanted to interview the person.	The student needed assistance in all aspects of setting up the interview.
Preparation	Before the interview, the student prepared several in-depth AND factual questions to ask.	Before the interview, the student prepared a couple of in-depth questions and several factual questions to ask.	Before the interview, the student prepared several factual questions to ask.	The student did not prepare any questions before the interview.
Politeness	Student never interrupted or hurried the person being interviewed and thanked them for being willing to be interviewed.	Student rarely interrupted or hurried the person being interviewed and thanked them for being willing to be interviewed.	Student rarely interrupted or hurried the person being interviewed, but forgot to thank the person.	Several times, the student interrupted or hurried the person being interviewed AND forgot to thank the person.
Follow-up Questions	The student listened carefully to the person being interviewed and asked several relevant follow-up questions based on what the person said.	The student listened carefully to the person being interviewed and asked a couple of relevant follow-up questions based on what the person said.	The student asked a couple of follow-up questions based on what s/he thought the person said.	The student did not ask any follow-up questions based on what the person said.
Formatting & Editing	The student edited and organized the transcript in a way that made the information clear and interesting.	The student edited and organized the transcript in a way that made the information clear.	The student edited and organized the transcript but the information was not as clear or as interesting as it could have been.	The student did NOT edit or organize the transcript.
Knowledge Gained	Student can accurately answer several questions about the person who was interviewed and can tell how this interview relates to the material being studied in class.	Student can accurately answer a few questions about the person who was interviewed and can tell how this interview relates to the material being studied in class.	Student can accurately answer a few questions about the person who was interviewed.	Student cannot accurately answer questions about the person who was interviewed.

## **PERFORMANCE ASSESSMENT: RUBRICS**

---

**NOTES:**