
Authentic Tasks and Rubrics: Going Beyond Traditional Assessments in College Teaching

Kathleen Montgomery

The intent of this article is to encourage the use of authentic tasks and assessment rubrics to round out the assessment of student learning in many college courses. The method of doing so is to provide a solid rationale for using authentic tasks and rubrics as well as offer several rubrics that could be used in a variety of courses.

Background

Traditional assessment practices in college courses consist primarily of pencil-and-paper activities. In a typical undergraduate course, students write two or three short papers or one large library-based research paper and take a mid-term and final examination. They expect to "sit" for one or more hourly exams often consisting of short answer questions (multiple-choice, true-false, fill-in-the-blanks, matching) and/or short essays. Some instructors give homework assignments; others require research projects instead of library research. Students in laboratory courses usually submit lab reports. For each of these products, the instructor uses some kind of personally predetermined weighting to assign a grade. The criteria for grades may or may not be shared with students.

Traditionally, many college instructors write comments on tests, papers, and

projects. These comments may range from terse to elaborate, from incomprehensible to clear. They may be directed toward the quality of the student's performance (Good point! Excellent conclusion!) or may be about the material itself. They may include proofreader's marks scattered throughout a paper by the instructor attempting to identify writing errors. Interpretation of such written comments is frequently left up to each student. The student must decide whether the comments are informative, instructive, critical, praiseworthy, or a mixture of these. More significant, the instructor's comments on papers and tests are done after rather than before the writing, so they cannot serve as guidelines, compromising the value of writing comments at all.

The goal of undergraduate programs all over the United States is that students become lifelong learners, a goal reached by enhancing the problem-solving and critical thinking abilities of students. For example, referring to its general education program, the University of New Hampshire's *2000 Undergraduate Catalog* states that the

program is designed to emphasize the acquisition and improvement of those fundamental skills essential to advanced college work, especially the abilities to think critically, to read with discernment, to write effectively, and to understand quantitative data. General education . . . aims to go beyond the mastery of job-related skills and educate students so that they learn how to

learn. The program . . . stresses intellectual adaptability and the development of those problem-solving abilities, cognitive skills, and learning techniques vital to lifelong learning. (3)

At the same time, the development of the skills mentioned above is not easily assessed by traditional methods of assessment such as the use of standard, objective forms of testing.

Moving Toward Authentic Assessment

There has been a renewed interest in examining traditional assessment practices, as evidenced by an increase in professional literature on the subject (Beaman 1998, Riley 1994). Numerous articles are devoted to such topics as the relationships among the areas of curriculum, instruction, and assessment; the understanding and examination of relationships between traditional assessment practices and their worth in measuring learning; the effects of grading on teaching and learning (Edwards and Edwards 1999; Busching 1998; Eison, Janzow, and Pollio 1993); reporting performance in a course with a single letter grade; and the purposes for placing grades on students' papers and projects (Miller 1999; Ekstrom and Villegas 1994; Goulden and Griffin 1995). Birenbaum and Douchy (1996) offer the following concerning the directions that should be taken in the area of achievement assessment in the college curriculum:

Kathleen Montgomery is the director of basic undergraduate programs at the University of Scranton, in Pennsylvania.

A shift has taken place from what some call a “culture of testing” to a “culture of assessment.” A strong emphasis is put on integrating assessment and instruction, on assessing process rather than just products and on evaluating individual progress relative to each student’s starting point. (47)

The question common to many authors examining assessment reform is when traditional assessments should be used, not if they should be used. Many instructors of college courses would probably say that traditional, standardized assessment would be most helpful when the need arose to measure the attainment of content knowledge. Such tests are easy to administer and to correct, not a small outcome considering the size of some college classes. At the same time, authentic tasks and rubrics help to measure the application of products and process to the real world where problem-solving and critical thinking abilities are often used.

Cognitive learning theory and its constructivist approach to knowledge acquisition support the need to use assessment methods that move away from passive responses by students to active construction of meaning. Students now are being asked to demonstrate, in a meaningful way, what they know and are able to do. Rather than measuring discrete, isolated skills, authentic assessment emphasizes the application and use of knowledge. Authentic assessment includes the holistic performance of meaningful, complex tasks in challenging environments that involve contextualized problems. Authentic tasks are often multidimensional and require higher levels of cognitive thinking such as problem solving and critical thinking.

Authentic assessment can involve the examination of the process as well as the product of learning. Progress in both process and product can be measured by assessing parallel performances over time. In assessing the process and product of a student’s writing, for example, expository essays should be compared to expository essays and not research papers. Position papers should be compared to position papers and not reviews of articles. Obviously, progress in writing dialogue does not necessarily mean progress in creating poetry. To evaluate a student’s progress successfully, then, the instructor needs to collect similar prod-

ucts and grade them after applying the same criteria. In art, for example, a series of small sculptures could be compared using one set of criteria while two-dimensional drawings would need another set.

Both process and product assessment methods can be accurately and fairly assessed through the use of specific evaluation criteria. The criteria must be known in advance by students so that they can apply them as they work through the process to arrive at the desired product. Students as well as the instructor can make formative assessments and continue or modify the work in progress.

According to Wiggins (1989), authentic assessment should involve real-life tasks, performances, or challenges that mirror those faced by experts in the par-

ticular field. The students should be asked to demonstrate their control over the essential knowledge being taught by actually using the information in a way that reveals their level of understanding. The criteria for judgment should be understood by the students from the start so that they can self-assess their work by applying the criteria. This latter part is where rubrics are useful. A *rubric* is a scaling process developed for and in conjunction with assessments that contains proficiency levels with clearly specified criteria. The criteria, known in advance by the students, provide descriptions of each level of performance in terms of what students are able to do. Process, progress, and product may all be evaluated by means of rubrics.

1. Letter is written in a professional manner and is error-free.		
Yes	Somewhat	No
Feedback:		

2. The reading program described contains at least 3 correct examples of how the program would ensure the continuing development of students' interests and attitudes.		
Yes	Somewhat	No
Feedback:		

3. The program's relationship to the theory of multiple intelligence is appropriate and clear.		
Yes	Somewhat	No
Feedback:		

4. Your viewpoints of cultural diversity are reflective of multicultural research and the relationship of your viewpoints to the proposed program is clearly established.		
Yes	Somewhat	No
Feedback:		

5. At least 5 examples of research information are present to support and emphasize your philosophical stance.		
Yes	Somewhat	No
Feedback:		

FIGURE 1. Evaluation Rubric for Letter to Curriculum Director

Constraints of Authentic Assessments

Instructors who use authentic tasks and rubrics recognize that designing and implementing such methods is time consuming. These concerns are well founded and are also well documented in the literature. For example, reported constraints to authentic assessments include the following:

- Strategies are often costly in terms of time, especially in general education classes where class size can be in the hundreds (Meadows, Dyal, and Wright 1998; Bracey 1993)
- Students often lack experience and skill in responding to open-ended, interpretive, applied, and higher-order essay questions (Liftig, Liftig, and Eaker 1992)
- The linguistic demands of many authentic assessment approaches may result in some equity issues (Spinelli 1998 and Rudner 1993)
- The increased validity of authentic assessment is often associated with a decrease in reliability (Riley and Stern 1998; Rudner 1993 and Bowers 1989)

Although these concerns are real, they are certainly not insurmountable. For example, even larger classes can appear

smaller by using some of the teaching techniques of smaller classes (Johnson 1995). The use of authentic tasks and rubrics that clearly state evaluation criteria for students can effectively break larger, lecture-oriented classes into smaller learning units of students. Rubrics can be used in peer review tasks as well as becoming a self-assessment tool.

Authentic assessment is not meant to replace traditional testing, but rather to add to the multifaceted approach of contemporary assessment. Given that the purpose of a college education program is to give students a liberal foundation by which good decisions can be made, the effort to work out the problems of any kind of assessment is a critical task in realizing this purpose.

Using Rubrics

Instructors who create authentic tasks often ask questions such as the following when they consider how to evaluate student work:

- What makes a quality process or product?
- Are the expectations for exceptional work clear to me? Are they clear to the students?

- How do I know what the students have learned as a result of completing a complex task?

Rubrics can provide the structure to answer these questions in a clear manner. A carefully designed rubric helps both the instructor and the students to define precise criteria for a successful process and/or product prior to and during the completion of a task. It also offers students specific feedback for future success on a similar task. In the absence of clear, specific criteria explained in advance by the instructor, assessment remains an isolated and incidental activity and the success of the learner is mostly incidental.

Specific evaluation criteria also have a positive impact on instruction. Establishing the criteria before the instruction focuses the instructor on the critical components of the curriculum and increases the likelihood that such components will be emphasized. Thus, there is an integration and alignment of curriculum content, instruction, and assessment that allows instructors and students to engage in meaningful learning.

Authentic assessment must be tied closely to authentic tasks. An example of an authentic assessment for education students in a reading course could be to prepare individual letters about the reading programs they wish to use:

As a first year classroom teacher, write a letter that you might send to the director of curriculum in your school district. The letter should explain why you have decided to design a reading program in your classroom that ensures continuing development of students' attitudes and interests. Include your viewpoint of multiple intelligences and cultural diversity as they relate. Use research information to support and emphasize your philosophical stance. (Jones, et al. 1994)

The use of a rubric with specific criteria and proficiency levels is highly desirable for evaluating such a multidimensional task. The comprehensiveness of the letter would demonstrate each student's depth of understanding of what makes an effective reading program.

Figure 1 represents a rubric that would effectively measure the level of success for the education students who were asked to write a letter to the director of curriculum.

Sobresaliente (16–20 points)

There is evidence of connected, organized discourse with some evidence of fluency of expression. The information presented about Madrid is accurate and plausible. The topic is thoroughly and logically developed with ideas clearly stated. An attempt at more complex structures and vocabulary may be evident, with some errors. There are almost no errors in agreement, word usage, or language mechanics.

Notable (11–15 points)

The letter is limited in topic development, somewhat choppy, and loosely organized. The information presented about Madrid is mostly accurate. Sequencing is logical but may be incomplete with some errors in word choice or usage. The meaning is clear, but there may be errors in agreement or syntax. Occasional errors in spelling. Communication is hampered by interference from English.

Bien (6–10 points)

The letter demonstrates a notable lack of fluency with inadequate development of the topic. Ideas are confused, with a lack of logical sequencing, limited vocabulary, and frequent errors in spelling, agreement, and mechanics of language. The information about Madrid is largely invented and limited. Meaning may not be clear, and there is evidence of interference from English.

Mención Honorífica (1–5 points)

The letter is not substantive, or there may not be enough to evaluate. There is no evidence of organization of thought and almost no information on Madrid is present. Errors dominate, and the letter does not communicate meaning. Lack of control of the language is evident.

FIGURE 2. Rubric for Español I

Giving the rubric to the students in advance of the task increases the chance that a quality product will be produced. The rubric needs to be as clear and specific as possible when a multidimensional task is assigned as the challenge of the task should be in its completion, not in figuring out the task itself.

Any discipline that assesses students via a written assignment can benefit from using rubrics. For example, a foreign language instructor who wants to measure a beginning student's performance with written Spanish may ask students to write a letter from the point of view of a Madrid native to a friend in Malaga, Spain. The letter should describe what weekends in Madrid are like for the letter writer. Figure 2 shows a rubric for this

task (modified from Scott and Rodgers 1993).

The letter-writing task is an authentic one and asks for knowledge of the lifestyle in Madrid as well as a working fluency with the target language, Spanish. The rubric is clear and specific, giving the students excellent guidelines for both process and product. The rubric fairly and accurately assesses student letter writing in Spanish.

Performance tasks such as presentations of any kind can also benefit from assessment using rubrics. Despite the fact that group presentations are commonly assigned throughout college classes, the criteria for a successful group presentation are rarely established in advance or fully communicated to students. The

quality of presentation is often, then, determined by the luck of the draw in the background knowledge of individual group members. As such, group presentations are difficult to assess and are often of mediocre quality. Figure 3 shows a rubric for a group presentation that can easily be modified to suit a particular task or topic across disciplines.

The learning gained from a group presentation, by the group members and the audience, can be considerable if the presentation is an accomplished one. The use of rubrics with specific evaluation criteria increases the chance of a successful presentation, making classroom presentations a good use of class time.

Any discipline can benefit from using well-defined rubrics. Figure 4 represents a rubric that specifies evaluation criteria for both process and product in the sciences (Jensen 1995).

The same rubric used over time would measure student progress toward understanding phase change, energy transformation and the key components in a phase change system.

Figure 5 shows how the process of problem solving in mathematics can be assessed. The following authentic task involves problem-solving reasoning and conceptualization of mathematics that could be assessed via the rubric:

You are a greenskeeper at a golf course. In order to know how much fertilizer to apply to the greens, you need an estimate of their areas. How would you estimate the area of a golf green? Keep practical considerations in mind. Your procedure should be feasible under actual conditions on a golf green and as simple as possible. To simplify the problem, assume that the green is flat. (Wilson 2000)

Figure 5, a rubric from the Baltimore City College faculty (2000), could be used to assess this particular problem as well as many other different problem-solving scenarios.

Changing the criteria or modifying the format of the rubric in one's own discipline should not be a difficult task for an instructor. Comments that used to show up on the finished product can be used to formulate the evaluation criteria presented at the start. An added advantage for the instructor is that the comments, in the form of criteria, are prepared once but used over and over again as the task is

Criteria	3	2	1
Content and Organization	Much research-based information clearly related to thesis; examples all support thesis; good transitions; work well organized.	Sufficient information that relates to thesis; lacks detail and depth in places; some transitions are abrupt and leaves questions about how some ideas are related.	There is a great deal of information that is not clearly connected to the thesis; no apparent transitions; organization has little logical order.
Audiovisual Media	Unique; adds to or develops thesis; use of media is varied and appropriate; enhances the presentation	Material supports the topic but does not enhance the presentation; not particularly well connected to the thesis.	Little or no audiovisual media used or not used effectively; what is used detracts from the thesis
Contributions of Individuals	Each member is equally involved in the presentation and is well informed about the topic.	Most members are actively involved and informed about the topic.	One or two members dominate; some members seem ill-prepared or not informed.
Speaking Skills	Poised, clear articulation, suitable volume; good posture, enthusiasm, and confidence noted.	Clear articulation; volume needs adjustment; enthusiasm not evident.	Some mumbling or inarticulate talk; the point of the presentation is not entirely clear.
Audience	Involved the audience in the presentation and held their attention throughout; conveyed enthusiasm to the audience	Presented thesis and evidence in an engaging manner and held audience's attention most of the time.	Some related facts, but went off the topic or presented facts with no enthusiasm and lost the audience's attention practically from the beginning.

FIGURE 3. Rubric for Group Oral Presentation

Task: This is a three-day activity in which students observe and perform a distillation to demonstrate phase change, explain energy transformation, and identify key components in the system.

Topics	Scores			
	4	3	2	1
Collaborative Worker: Student can take charge of his/her own behavior in a group.	Student stays on task; offers useful ideas and can defend them; can take on various roles; participates without prompting.	Student stays on task; offers useful ideas and can defend them; can take on various roles; rarely requires prompting to participate.	Student does not attend to the lab; accepts group view or considers only his/her own ideas worthwhile. Needs regular prompting to stay on task.	Student does not respond to the group. Student is not involved or may try to undermine the efforts of the group.
Scientific Literacy: Student uses processes and skills of science to conduct investigations.	Student identifies the question, forms a possible solution, writes out steps to test the possible solution, designs a data chart, collects data, and concludes about the validity of the possible solution.	Student identifies the question and forms a possible solution. Procedure and data charts are complete but lack clarity and/or creativity. Student concludes about the validity of the possible solution.	Student identifies the question but does not form a complete solution. Procedure and data charts are incomplete and the conclusion does not speak to the possible solution.	Student does not identify the question. No possible solution is given. Procedure and data charts are incomplete or missing. The conclusion is incomplete or missing.
Systems Analysis: Student describes how a system operates internally and how it interacts with the outside world	Student identifies how parts of the system interact and provides personal insight into the interaction of the parts; relates how the system interacts with the outside world.	Student identifies how parts of the system interact and relates how the system interacts with the outside world.	Student does not identify some parts of the system. Student does not understand how the parts interact and does not relate how the system interacts with the outside world.	Student incorrectly identifies the parts and cannot describe how they interact either within or outside the system.

FIGURE 4. Rubric for Distillation Exercise

being evaluated for each student. In contrast, traditional assessment meant writing some of the same comments over and over again for different students—and far too often, for the same students on the next similar task. As with any assessment tool, rubrics need to be modified every time using direct feedback from the students and indirect feedback from the quality of the students' work.

The Challenge of Assigning Grades

Rubrics can help with the challenging task of assigning grades as well. The letter grades used in the reporting systems at most colleges and universities can be made compatible with the numbers or words used in rubrics to describe the proficiency levels presented in rubrics. For example, the rubric for Español I (figure 2) could be adapted as follows:

Rubric	Grade
Sobresaliente (16–20 points)	A/A-

Notable
(11–15 points) B+/B/B-

Bien
(6–10 points) C+/C/C-

Mencion Honorifica
(1–5 points) D+/D/D-

The practice of holistic grading, a way of assessing the overall effectiveness of a project, paper, or presentation, is also compatible with rubrics. The instructor need only refer to the rubric to provide students with specific feedback. When rubrics have been presented to students as an integral part of the learning process, the assigning of grades becomes a small step in the process. The grade becomes more the vehicle for reporting to others and is no longer made responsible for providing substantive information to students.

Conclusion

Viewing assessment as a way to measure the process of learning, the progress of learning, and the product of learning

(including course content) leads instructors away from many of the traditional problems in evaluating students. Just as traditional pencil-and-paper assessment methods cause instructors and students to emphasize isolated and discrete facts (sometimes thought of as “course content”), authentic assessment methods free them to develop higher level concepts and related understandings. Curriculum, then, becomes more than just an outline of a course of study to be tested. It begins to reflect the complexities of the real world—past, present, and future.

With authentic assessment, emphasis is on process as well as product. This means learning is seen as an ongoing refinement leading toward the achievement of established outcomes. Because the refinement takes place over time, progress can be recorded along the way. Both instructors and students are involved in systematic formative assessments of the developing product. Instructional modifications are a natural outcome as instructors guide stu-

Level	Constructed Response
3	The response indicates application of a reasonable strategy that leads to a correct solution in the context of the problem. The representations are correct. The explanation and/or justification are logically sound, clearly presented, fully developed, support the solution, and do not contain significant mathematical errors.
2	The response indicates application of a reasonable strategy that may be incomplete or undeveloped. It may or may not lead to a correct solution. The representations are essentially correct. The explanation and/or justification support the solution and are plausible, although it may not be well developed or complete. The response demonstrates a conceptual understanding and analysis of the problem.
1	The response indicates little or no attempt to apply a reasonable strategy or applies an inappropriate strategy. It may or may not have the correct answer. The representations are fundamentally correct. The explanation and/or justification reveal serious flaws in reasoning. The explanation and/or justification may be incomplete or missing. The response demonstrates a minimal understanding and analysis of the problem.
0	The response is completely incorrect or irrelevant. There may be no response, or the response may state "I don't know."

Explanation refers to the student using the language of mathematics to communicate how the student arrived at the solution.

Justification refers to the student using mathematical principles to support the reasoning used to solve the problem. This could include the appropriate definitions, postulates, and theorems.

Essentially correct representations may contain minor errors such as missing labels, reversed axes, or scales that are not uniform.

Fundamentally correct representations have the majority of the information correctly represented.

**FIGURE 5. Rubric for Problem Solving Response
Baltimore City College**

dents' learning. Using authentic assessment methods such as rubrics honors the complexities of the teaching-learning paradigm.

REFERENCES

- Baltimore City College. 2000. *BBC mathematics rubrics*. Retrieved November 27, 2000 from <<http://baltimorecitycollege.org/faculty/mathdept/mathrubric.htm>>.
- Beaman, R. 1998. The unquiet . . . even loud, andragogy! Alternative assessments for adult learners. *Innovative Higher Education* 23(1): 47–59.
- Birenbaum, M., and F. Douchy. 1996. *Alternative in assessment of achievements, learning processes and prior knowledge*. Norwell, MA: Kluwer Academic Publishers.
- Bowers, B. 1989. *Alternatives to standardized*

educational assessment. Eugene, OR: ERIC Clearinghouse on Educational Management. ERIC Document Reproduction No. ED 312 773.

- Bracey, G. 1993. Assessing the new assessments. *Principal* 72: 34–36.
- Busching, B. 1998. Grading inquiry projects. *New Directions for Teaching and Learning* (74): 89–96.
- Edwards, C., and L. Edwards. 1999. Let's end the grading game. *Clearing House* 72(5): 260–63.
- Eison, J., F. Janzow, and H. Pollio. 1993. Assessing faculty orientations towards grades and learning: Some initial results. *Psychological Reports* 73(2): 643(14).
- Ekstrom, R., and A. Villegas. 1994. *College grades: An exploratory study of policies and practices*. New York: College Board Publications.

Goulden, N., and C. Griffin. 1995. The meaning of grades based on faculty and student metaphors. *Communication Education* 44(2): 110–25.

Jensen, K. 1995. Effective rubric design: Making the most of this powerful assessment tool. *Science Teacher* 62(5):34–37.

Johnson, P. 1995. *Students' perceptions of foreign language communicative teaching and class size*. ERIC Document (ED 396 560).

Jones, et al. 1994, November. *Project authentic connections: Performance-based assessment in reading/language arts courses or preservice teachers*. Paper presented at the 44th Annual National Reading Conference, San Diego, CA.

Liftig, I, B. Liftig, and B. Eaker. 1992. Making assessment work: What teachers should know before they try it. *Science Scope* 15: 4–6.

Meadows, R., A. Dyal, and J. Wright. 1998. Preparing educational leaders through the use of portfolio assessment: An alternative comprehensive examination. *Journal of Instructional Psychology* 25(2): 94–99.

Miller, J. 1999. What's wrong at E.T.S.? Insider's view of grading A.P. government essays. *College Teaching* 47(1): 2–7.

Riley, H. et al. 1994. *Current trends in grades and grading practices in undergraduate higher education: Results of the 1992 AACRAO survey*. Washington, D.C.: American Association of Collegiate Registrars and Admissions Officers.

Riley, K., and B. Stern. 1998. Using authentic assessment and qualitative methodology to bridge theory and practice. *Educational Forum* 62(2): 178–85.

Rudner, L. 1993. *Issues and concerns*. Gopher Resources <gopher.ed.gov/Educational>, Improvement and Statistics (OREI and NCES)/Educational Resources Information Center (ERIC)/ERIC.

Scott, R., and B. Rodgers. 1993. Assessing communication in writing: The development of a Spanish writing contest. *Foreign Language Annals* 26(3): 383–392.

Spinelli, C. 1998. *Teacher education reform: Promoting interactive teaching strategies and authentic assessment for instructing an increasing diverse population of students*. ERIC Document Reproduction No. ED 418 076.

The University of New Hampshire. 2000. *Undergraduate Catalog 2000*. Durham, NH: University Press.

Wiggins, G. 1989. A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan* 49(8): 35–37.

Wilson, J. 2000. *Problem: Area of golfing greens*. Retrieved November 27, 2000 from <<http://jwilson.coe.uga.edu/emt725/Bob2/Golf/GolfGreen.html>>.

