Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Volume 8, Number 14, May, 2003 ISSN=1531-7714

Recommendations for Developing Classroom Performance Assessments and Scoring Rubrics

Barbara M. Moskal
Colorado School of Mines

This paper provides a set of recommendations for developing classroom performance assessments and scoring rubrics similar to the sets of recommendations for multiple choice tests provided in this journal by Frary (1995) and Kehoe (1995a, 1995b). The recommendations are divided into five categories: 1) Writing Goals and Objectives, 2) Developing Performance Assessments, 3) Developing Scoring Rubrics, 4) Administering Performance Assessments and 5) Scoring, Interpreting and Using Results. A broad literary base currently exists for each of these categories. This article draws from this base to provide a set of recommendations that guide the classroom teacher through the four phases of the classroom assessment process — planning, gathering, interpreting and using (Moskal, 2000a). Each section concludes with references for further reading.

Writing Goals and Objectives

Before a performance assessment or a scoring rubric is written or selected, the teacher should clearly identify the purpose of the activity. As is the case with any assessment, a clear statement of goals and objectives should be written to guide the development of both the performance assessment and the scoring rubric. "Goals" are broad statements of expected student outcomes and "objectives" divide the goals into observable behaviors (Rogers & Sando, 1996). Questions such as, "What do I hope to learn about my students' knowledge or skills?," "What content, skills and knowledge should the activity be designed to assess?," and "What evidence do I need to evaluate the appropriate skills and knowledge?", can help in the identification of specific goals and objectives.

Recommendations for writing goals and objectives:

1. *The statement of goals and accompanying objectives should provide a clear focus for both instruction and assessment.* Another manner in which to phrase this recommendation is that the stated goals and objectives for the performance assessment should be clearly aligned with the goals and objectives of instruction. Ideally, a statement of goals and objectives is developed *prior* to the instructional activity and is used to guide both instruction and assessment.

2. *Both goals and objectives should reflect knowledge and information that is worthwhile for students to learn.* Both the instruction and the assessment of student learning are intentional acts and should be guided through planning. Goals and objectives provide a framework for the development of this plan. Given the critical relationship between goals and objectives and instruction and assessment, goals and objectives should reflect important learning outcomes.

3. *The relationship between a given goal and the objectives that describe that goal should be apparent.* Objectives lay the framework upon which a given goal is evaluated. Therefore, there should be a clear link between the statement of the goal and the objectives that define that goal.

4. *All of the important aspects of the given goal should be reflected through the objectives.* Once again, goals and objectives provide a framework for evaluating the attainment of a given goal. Therefore, the accompanying set of objectives should reflect the important aspects of the goal.

5. *Objectives should describe measurable student outcomes.* Since objectives provide the framework for evaluation, they need to be phrased in a manner that specifies the student behavior that will demonstrate the attainment of the larger goal.

6. *Goals and objectives should be used to guide the selection of an appropriate assessment activity.* When the goals and objectives are focused upon the recall of factual knowledge, a multiple choice or short response assessment may be more appropriate and efficient than a performance assessment. When the goals and objectives are focused upon complex learning outcomes, such as reasoning, communication, teamwork, etc., a performance assessment is likely to be appropriate (Perlman, 2002).

Writing goals and objectives, at first, appears to be a simple. After all, this process primarily requires clearly defining the desired student outcomes. Many teachers initially have difficulty creating goals and objectives that can be used to guide instruction and that can be measured. An excellent resource that specifically focuses upon the "how to" of writing

measurable objectives is a book by Gronlund (2000). Other authors have also addressed these issues in subsections of larger works (e.g., Airasian, 2000; 2001; Oosterhoff, 1999).

Developing Performance Assessment

As the term suggests, performance assessments require a demonstration of students' skills or knowledge (Airasian, 2000; 2001; Brualdi, 1998; Perlman, 2002). Performance assessments can take on many different forms, which include written and oral demonstrations and activities that can be completed by either a group or an individual. A factor that distinguishes performance assessments from other extended response activities is that they require students to demonstrate the application of knowledge to a particular context (Brualdi, 1998; Wiggins, 1993). Through observation or analysis of a student's response, the teacher can determine what the student knows, what the student does not know and what misconceptions the student holds with respect to the purpose of the assessment.

Recommendations for developing performance assessments:

1. *The selected performance should reflect a valued activity*. According to Wiggins (1990), "The best tests always teach students and teachers alike the kind of work that most matters; they are enabling and forward-looking, not just reflective of prior teaching." He suggests the use of tasks that resemble the type of activities that are known to take place in the workforce (e.g., project reports and presentations, writing legal briefs, collecting, analyzing and using data to make and justify decisions). In other words, performance assessments allow students the opportunity to display their skills and knowledge in response to "real" situations (Airasian, 2000; 2001; Wiggins, 1993).

2. *The completion of performance assessments should provide a valuable learning experience.* Performance assessments require more time to administer than do other forms of assessment. The investment of this classroom time should result in a higher payoff. This payoff should include both an increase in the teacher's understanding of what students know and can do and an increase in the students' knowledge of the intended content and constructs.

3. *The statement of goals and objectives should be clearly aligned with the measurable outcomes of the performance activity*. Once the task has been selected, a list can be made of how the elements of the task map into the desired goals and objectives. If it is not apparent as to how the students' performance will be mapped into the desired goals and objectives, then adjustments may need to be made to the task or a new task may need to be selected.

4. *The task should not examine extraneous or unintended variables.* Examine the task and think about whether there are elements of the task that do not map directly into the goals and objectives. Is knowledge required in the completion of the task that is inconsistent with the purpose? Will lack of this knowledge interfere or prevent the students from completing the task for reasons that are not consistent with the task's purpose? If such factors exist, changes may need to be made to the task or a new task may need to be selected.

5. *Performance assessments should be fair and free from bias.* The phrasing of the task should be carefully constructed in a manner that eliminates gender and ethnic stereotypes. Additionally, the task should not give an unfair advantage to a particular subset of students. For example, a task that is heavily weighted with baseball statistics may give an unfair advantage to the students that are baseball enthusiasts.

The recommendations provided above have been drawn from the broader literary base concerning the construction of performance assessments. The interested reader can acquire further details concerning the development process by consulting other articles that are available through this journal (i.e., Brualdi, 1998; Roeber, 1996; Wiggins, 1990) or books (e.g., Wiggins, 1993; 1998) that address this subject.

Developing Scoring Rubrics

Scoring rubrics are one method that may be used to evaluate students' responses to performance assessments. Two types of performance assessments are frequently discussed in the literature: analytic and holistic. Analytic scoring rubrics divide a performance into separate facets and each facet is evaluated using a separate scale. Holistic scoring rubrics use a single scale to evaluate the larger process. In holistic scoring rubrics, all of the facets that make-up the task are evaluated in combination. The recommendations that follow are appropriate to both analytic and holistic scoring rubrics.

Recommendations for developing scoring rubrics:

1. *The criteria set forth within a scoring rubric should be clearly aligned with the requirements of the task and the stated goals and objectives*. As was discussed earlier, a list can be compiled that describes how the elements of the task map into the goals and objectives. This list can be extended to include how the criteria that is set forth in the scoring rubric maps into both the elements of the task and the goals and objectives. Criteria that cannot be mapped directly back to both the task and the purpose should not be included in the scoring rubric.

2. *The criteria set forth in scoring rubrics should be expressed in terms of observable behaviors or product characteristics.* A teacher cannot evaluate an internal process unless this process is displayed in an external manner. For example, a teacher cannot look into students' heads and see their reasoning process. Instead, examining reasoning

requires that the students explain their reasoning in written or oral form. The scoring criteria should be focused upon evaluating the written or oral display of the reasoning process.

3. *Scoring rubrics should be written in specific and clear language that the students understand.* One benefit of using scoring rubrics is that they provide students with clear description of what is expected *before* they complete the assessment activity. If the language employed in a scoring rubric is too complex for the given students, this benefit is lost. Students should be able to understand the scoring criteria.

4. *The number of points that are used in the scoring rubric should make sense.* The points that are assigned to either an analytic or holistic scoring rubric should clearly reflect the value of the activity. On an analytic scoring rubric, if different facets are weighted differently than other facets of the rubric, there should be a clear reason for these differences.

5. *The separation between score levels should be clear.* The scale used for a scoring rubric should reflect clear differences between the achievement levels. A scale that requires fine distinctions is likely to result in inconsistent scoring. A scoring rubric that has fewer categories and clear distinctions between these categories is preferable over a scoring rubric that has many categories and unclear distinctions between the categories.

6. *The statement of the criteria should be fair and free from bias.* As was the case with the statement of the performance activity, the phrasing used in the description of the performance criteria should be carefully constructed in a manner that eliminates gender and ethnic stereotypes. Additionally, the criteria should not give an unfair advantage to a particular subset of students that is unrelated to the purpose of the task.

Greater detail concerning the development of scoring rubrics, both analytic and holistic, is immediately available through this journal. Mertler (2001) and Moskal (2000b) have both described the differences between analytic and holistic scoring rubrics and how to develop each type of rubric. Books have also been written or compiled (e.g., Arter & McTighe, 2001; Boston, 2002) that provide detailed examinations of the rubric development process and the different types of scoring rubrics.

Administering Performance Assessments

Once a performance assessment and its accompanying scoring rubric are developed, it is time to administer the assessment to students. The recommendations that follow are specifically developed to guide the administration process.

Recommendations for administering performance assessments:

1. *Both written and oral explanations of tasks should be clear and concise and presented in language that the students understand.* If the task is presented in written form, then the reading level of the students should be given careful consideration. Students should be given the opportunity to ask clarification questions before completing the task.

2. *Appropriate tools need to be available to support the completion of the assessment activity.* Depending on the activity, students may need access to library resources, computer programs, laboratories, calculators, or other tools. Before the task is administered, the teacher should determine what tools will be needed and ensure that these tools are available during the task administration.

3. *Scoring rubrics should be discussed with the students before they complete the assessment activity.* This allows the students to adjust their efforts in a manner that maximizes their performance. Teachers are often concerned that by giving the students the criteria in advance, all of the students will perform at the top level. In practice, this rarely (if ever) occurs.

The first two recommendations provided above are appropriate well beyond the use of performance assessments and scoring rubrics. These recommendations are consistent with the Standards of the American Educational Research Association, American Psychological Association & National Council on Measurement in Education (1999) with respect to assessment and evaluation. The final recommendation is consistent with prior articles that concern the development of scoring rubrics (Brualdi, 1998; Moskal & Leydens, 2000).

Scoring, Interpreting and Using Results

As was discussed earlier, a scoring rubric may be used to score student responses to performance assessments. This section provides recommendations for scoring, interpreting and using the results of performance assessments.

Recommendations for scoring, interpreting and using results of performance assessments:

1. *Two independent raters should be able to acquire consistent scores using the categories described in the scoring rubric.* If the categories of the scoring rubric are written clearly and concisely, then two raters should be able to score the same set of papers and acquire similar results.

2. *A given rater should be able to acquire consistent scores across time using the scoring rubric.* Knowledge of who a

student is or the mood of a rater on a given day may impact the scoring process. Raters should frequently refer to the scoring rubric to ensure that they are not informally changing the criteria over time.

3. *A set of anchor papers should be used to assist raters in the scoring process.* Anchor papers are student papers that have been selected as examples of performances at the different levels of the scoring rubric. These papers provide a comparison set for raters as they score the student responses. Raters should frequently refer to these papers to ensure the consistency of scoring over time.

4. *A set of anchor papers with students' names removed can be used to illustrate to both students and parents the different levels of the scoring rubric.* Ambiguities within the rubric can often be clarified through the use of examples. Anchor papers with students names removed can be used to clarify to both students and parents the expectations set forth through the scoring rubric.

5. *The connection between the score or grade and the scoring rubric should be immediately apparent.* If an analytic rubric is used, then the report should contain the scores for each analytic level. If a summary score or grade is provided, than an explanation should be included as to how the summary score or grade was determined. Both students and parents should be able to understand how the final grade or score is linked to the scoring criteria.

6. *The results of the performance assessment should be used to improve instruction and the assessment process.* What did the teacher learn from the student responses? How can this be used to improve future classroom instruction? What did the teacher learn about the performance assessment or the scoring rubric? How can these instruments be improved for future instruction? The information that is acquired through classroom assessment should be actively used to improve future instruction and assessment.

The first three recommendations concern the important concept of "rater reliability" or the consistency between scores. Moskal and Leydens (2000) examine the concept of rater reliability in an article that was previously published in this journal. A more comprehensive source that addresses both validity and reliability of scoring rubrics is a book by Arter and McTighe (2001), *Scoring Rubrics in the Classroom: Using Performance Criteria for Assessing and Improving Student Performance.* The American Educational Research Association, American Psychological Association and National Council of Measurement in Education (1999) also address these issues in their Standards document. For information concerning methods for converting rubric scores to grades, see "Converting Rubric Scores to Letter Grades" (Northwest Regional Educational Laboratory, 2001).

Conclusions

The purpose of this article is to provide a set of recommendations for the development of performance assessments and scoring rubrics. These recommendations can be used to guide a teacher through the four phases of classroom assessment, planning, gathering, interpreting and using. Extensive literature is available on each phase of the assessment process and this article addresses only a small sample of that work. The reader is encouraged to use the previously cited work as a starting place to better understand the use of performance assessments and scoring rubrics in the classroom. Additionally, books by Airasian (2000; 2001), Oosterhof (1999), Rudner and Schafer (2002), and Stiggins (1994) provide a more detailed look at the broader classroom assessment process.

Acknowledgments

This article was originally developed as part of a National Science Foundation (NSF) grant (EEC 0230702), *Engineering Our World*. The opinions and ideas expressed in this article are that of the author and not of the NSF.

References

Airasian, P.W. (2000). *Assessment in the Classroom: A Concise Approach* (2nd ed.). Boston: McGraw-Hill.

Airasian, P.W. (2001). *Classroom Assessment: Concepts and Applications* (4th ed.). Boston: McGraw Hill.

American Educational Research Association, American Psychological Association & National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

Arter, J. & McTighe, J. (2001). *Scoring Rubrics in the Classroom: Using Performance Criteria for Assessing and Improving Student Performance.* Thousand Oaks, California: Corwin Press Inc.

Boston, C. (Eds.) (2002). *Understanding Scoring Rubrics*. University of Maryland, MD: ERIC Clearinghouse on Assessment and Evaluation.

Brualdi, A. (1998). "Implementing performance assessment in the classroom." *Practical Assessment, Research & Evaluation, 6*(2) [On-line]. Available: http://pareonline.net/getvn.asp?v=6&n=2.

Gronlund, N.E. (2000). *How to Write and Use Instructional Objectives* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Frary, R. B. (1995). "More multiple-choice item writing do's and don'ts." *Practical Assessment, Research & Evaluation," 4*(11) [On-line]. Available: http://pareonline.net/getvn.asp?v=4&n=11.

Kehoe, J. (1995a). "Writing multiple-choice test items." *Practical Assessment, Research & Evaluation, 4*(9) [On-line]. Available: http://pareonline.net/getvn.asp?v=4&n=9.

Kehoe, J. (1995b). "Basic item analysis for multiple-choice tests." *Practical Assessment, Research & Evaluation,* 4(10). Available online: http://pareonline.net/getvn.asp?v=4&n=10.

Mertler, C. A. (2001). "Designing scoring rubrics for your classroom." *Practical Assessment, Research & Evaluation, 7*(25). Available online: http://pareonline.net/getvn.asp?v=7&n=25.

Moskal, B. (2000a). "An assessment model for the mathematics classroom." *Mathematics Teaching in the Middle School, 6* (3), 192-194.

Moskal, B. (2000b). "Scoring rubrics: What, when and how?" *Practical Assessment, Research & Evaluation, 7*(3) [On-line]. Available: http://pareonline.net/getvn.asp?v=7&n=3.

Moskal, B. & Leydens, J. (2000). "Scoring rubric development: Validity and reliability." Practical Assessment, Research & Evaluation, 7(10). Available online: http://pareonline.net/getvn.asp?v=7&n=10.

Northwest Regional Educational Laboratory (2002). "Converting rubric scores to letter grades." In C. Boston's (Eds.), *Understanding Scoring Rubrics* (pp. 34-40). University of Maryland, MD: Clearing House on Assessment and Evaluation.

Oosterhof, A. (1999). *Developing and Using Classroom Assessments* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Perlman, C. (2002). "An introduction to performance assessment scoring rubrics". In C. Boston's (Eds.), *Understanding Scoring Rubrics* (pp. 5-13). University of Maryland, MD: ERIC Clearinghouse on Assessment and Evaluation.

Rogers, G. & Sando, J. (1996). *Stepping Ahead: An Assessment Plan Development Guide*. Terra Haute, Indiana: Rose-Hulman Institute of Technology.

Roeber, E.D. (1996). "Guidelines for the development and management of performance assessments." *Practical Assessment, Research & Evaluation*, 5(7). Available online: http://pareonline.net/getvn.asp?v=5&n=7.

Rudner, L.M. & Schafer, W.D. (Eds.) (2002). *What Teachers Need to Know about Assessment*. Washington, DC: National Education Association.

Stiggins, R. (1994). *Student-Centered Classroom Assessment*. New York: Macmillan Publishing Company.

Wiggins, G. (1998). *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*. San Francisco: Jossey-Bass Publishers.

Wiggins, G. (1993). *Assessing Student Performances*. San Francisco: Jossey-Bass Publishers.

Wiggins, G. (1990). "The case for authentic assessment." *Practical Assessment, Research & Evaluation, 2*(2). Available online: http://pareonline.net/getvn.asp?v=2&n=2.

Descriptors: *Rubrics; Scoring; *Student Evaluation; *Test Construction; *Evaluation Methods; Grades; Grading; *Scoring

Citation: Moskal, Barbara M. (2003). Recommendations for developing classroom performance assessments and scoring rubrics. *Practical Assessment, Research & Evaluation*, 8(14). Available online: http://PAREonline.net/getvn.asp?v=8&n=14.