Performance Assessment

Timothy F. Slater
Department of Physics
Montana State University

WHY USE PERFORMANCE ASSESSMENT?

Although facts and concepts are fundamental in any undergraduate SMET course, knowledge of methods, procedures and analysis skills that provide context are equally important. Student growth in these latter facets prove somewhat difficult to evaluate, particularly with conventional multiple-choice examinations. Performance assessments, used in concert with more traditional forms of assessment, are designed to provide a more complete picture of student achievement.

WHAT IS PERFORMANCE ASSESSMENTS?

Performance assessments are designed to judge student abilities to USE specific knowledge and research skills. Most performance assessments require the student to manipulate equipment to solve a problem or make an analysis. Rich performance assessments reveal a variety of problem-solving approaches, thus providing insight into a student's level of conceptual and procedural knowledge.

WHAT IS INVOLVED?

Instructor Preparation Time: Medium.
Preparing Your Students: None.
Class Time: 10 – 40 minutes depending on complexity of task.
Disciplines: Appropriate for laboratory-based sciences.
Class Size: Small for direct applications, unlimited for embedded assessments using student-completed forms.
Special Classroom/Technical Requirements: Varies according to task.
Individual or Group Involvement: Both.
Analyzing Results: Low.
Other Things to Consider: Manipulative materials are often required as well as room monitors

Contents

Description

Performance assessment strategies are composed of three distinct parts: a performance *task*; a *format* in which the student responds; and a predetermined *scoring system*. Tasks are assignments designed to assess a student's ability to manipulate equipment (laboratory equipment, computers, documents, etc.) for a given purpose. Students can either complete the task in front of a panel of judges or use a written response sheet. The student is then scored by comparing the performance against a set of written criteria. When used with students with highly varying abilities, performance tasks can take maximum advantage of judging student abilities by using tasks with multiple correct solutions.

Students are graded on the process of problem solving using a rating scale based on explicit standards. Performance assessments have been validated by English faculty who conduct writing assessments, Olympic judges who score competition divers, jury panels who evaluate musical performances and K-12 science teachers (Shavelson, et al.). In these examples, individuals receive a score based on analyzing and evaluating various required components of a performance individually. This is known as ANALYTIC SCORING.

Figure 1: Sample Authentic Tasks:

- Is this water sample is suitable for drinking?
- Remove these old, unlabeled chemicals from the lab.
- What is the approximate age of this fossil-bearing rock?
- How fast was the car moving before it crashed if it left 15 meter skid marks in front of this building?

Performance assessment strategies are best utilized in concert with other forms of assessment. Similar to driver education or pilot certification, both factual knowledge and procedural knowledge are important components of a complete of SMET education.

Assessment Purposes

The purpose of performance assessment is to evaluate the *actual process of doing* science or mathematics. Performance assessments examine students' actual application of knowledge to solve problems. In some cases, the solution of the problem may imply the application of a specific procedure learned in class; in others, a combination of procedures; still in others it may require a thoughtful adaptation of students' knowledge. The assessment of student's knowledge focuses on the performance and the result.

Limitations

Performance assessments are typically inappropriate for measuring student knowledge of facts.

Teaching Goals

Develop ability to:

- apply systematic procedures
- authentically utilize resource texts, laboratory equipment, and computers.
- use scientific methodology
- apply and evaluate multiple approaches
- solve complex problems

Suggestions for Use

Diagnostic Purposes

Performance assessments may be used for diagnostic purposes. What do students know about how to solve certain types of problems? Do they know how to control variables? how to use intruments? how to evaluate findings? Information provided at the beginning of the course may help decide where to start or what issues of the course need special attention.

Instructional Purposes

A good performance assessment often is indistinguishable from a learning activity, except for standardization and scoring. In this light, a performance task that simulates the authentic tasks of a scientists or mathematician may be used as either an instructional activity or an assessment activity. If the assessment task is used in such a way that the student would normally not know it is an assessment activity, it is called an *embedded* task.

Monitoring Purposes

The goal of a performance assessment is to judge the level of competency students have achieved in *doing* science and mathematics. Accordingly, performance assessment strategies are best used to monitor student process skills and problem solving approaches. The most effective performance assessments are authentic tasks that are open-ended with multiple-correct solution paths.

Step-by-Step Instructions

- Carefully construct the learning goals for the instructional unit
- Decide if performance assessment supports student learning and assessment for these goals
- Clearly define the knowledge and skills students need to apply or demonstrate in solving a problem.
- Determine the criteria (standards) against which students will be judged and define indicators of "levels" of competence.
- Inform students of your expectations that students have every opportunity to clearly demonstrate to that course learning objectives have been mastered
- Design an authentic task that is somewhat undefined, complex, and has multiple entry and exit points.
- Determine which distinct components of the task need to be analyzed.
- Directly observe students or develop a structured student-answer sheet that allows you to evaluate various components of the task.
- Match student performance to criteria (standards) and determine which level most closely matches student performance.
- Provide student feedback in terms of levels of competence, not numerical scores.

Figure 2: Holistic Scoring Example, The Telescope Task

Your task is to set up and align the 8" telescope, find three different sky objects, and accurately describe some aspects of these objects that astronomers consider to be important.

Level 3: Student completes all aspects of task quickly and efficiently and is able to answer questions about the equipment used and objects observed beyond what is obvious. The tasks are:

1. align telescope mount with north celestial pole;
2. align finder telescope with primary telescope;
3. center on target object;
4. select and focus appropriate eyepiece;
5. provide information about the target beyond the literal descriptive level; and
6. answer questions about the target correctly.

Level 2: Student completes all aspects of task and provides descriptive information about the equipment and objects observed.

Level 1: Student is not able to complete all aspects of task or is not able to sufficient provide information about the equipment used or objects observed.

Level 0: No attempt or meaningful effort obvious.

Figure 3: Analytic Scoring Example, Physics Laboratory

Performance Task Evaluation Sheet

Performance Task Title: _____ Date: _____

Student Name: _____ Total Score: _____

Performance Goals:

	No Evidence 0 points	Approaches Goal 1 point	Meets Goal 2 points	Exceeds Goal 3 points
(1) **Method of Research**: Identifies the information and steps needed to solve the problem.	0	1	2	3
(1) **Appropriate Use of Equipment and Apparatus**: Demonstrates the correct application and cautious use of equipment and apparatus to meet this standard.	0	1	2	3
(1) **Accuracy and Precision**: Demonstrates the ability to make accurate measurements to appropriate precision and to judge the reasonableness of the results.	0	1	2	3
(1) **Comprehension**: Properly applies concepts and formulas related to phenomena.	0	1	2	3
(1) **Calculations**: Properly uses mathematics and mathematical conversions (as needed) to solve the problem.	0	1	2	3
(1) **Laboratory Report**: Communicates conclusions in a complete, clear, and organized way using illustrations.	0	1	2	3

Note: Adapted, with permission, from T.F. Slater and J.M. Ryan (1993). Laboratory performance assessment. The Physics Teacher, v. 31, no. 5, pages 306 – 309.

Variations

Checklists for Highly Structured Tasks

In science and mathematics, some tasks require systematic procedures that do not yield multiple-entry points or exit points. In this case, a check list system can be appropriately used by an observer or a highly-structured student-answer sheet in which each aspect of the procedure and result is described in detail. Faculty have often found the highly-structured format useful when working with large-enrollment classes. Highly -structured assessment tasks provide students with step-by-step instructions to follow. In contrast, less structured assessment tasks give students more opportunity to make judgments in determining the procedures needed to solve the problem.

Collaborative Groups

Performance assessment can be administered individually, in pairs, or collaborative groups. If it is administered in pairs or groups, students should write in their own answer/response sheet. It is important to keep in mind that when students solve the problem in pairs or groups, the goal and the composition of the group will affect the student's individual performance. In this context, it should be clear exactly what the purpose of the assessment is (e.g., how well students' ability to interact and collaborate with others).

Panel of Peers

Similar to the professional lives of college and university faculty, peer assessment can plan an important role in improving student learning of both the assessed and the assessors. If criteria (standards) are clearly described to students with examples showing each level of competency, they are often able to judge the performance of peers effectively and reliably.

Analysis

It is important to have predetermined criteria to evaluate the students' performance. Students should not be scored/graded against their peers, but based on the criteria predefined. Ideally, students should be provided with the criteria before the assessment. Accordingly, the grade book and student feedback reflects levels of competency, rather than comparative scores. It is always useful to try to find in students' performance patterns of appropriate and inappropriate responses (e.g., most of students did not control variable "X"). This helps focus on problems observed across many students during instruction.

Pros and Cons

- Performance assessments provide a way of observing the application of procedures.
- Performance assessments emphasize multiple correct answers and creative solutions.
- Performance assessments simulate the real-world tasks that scientists, mathematicians, engineers, and researchers encounter.
- Performance assessments allow faculty to measure overarching course goals of concept application

However:

- Performance assessments address fewer learning objectives than other forms of assessment.
- Students who have been successful at memorizing initially find performance assessments intimidating.
- Development of clear criteria (standards) that indicate competency levels requires multiple iterations.

Theory and Research

The acknowledged weaknesses of conventional paper and pencil assessments have led to the recent development of alternative testing strategies. Already validated and used in many K-12 schools, one of the most widely used of these is called PERFORMANCE ASSESSMENT. The keystone of performance assessment is the use of a graded, authentic task. An AUTHENTIC task is one in which students are required to address problems grounded in real-life contexts. Such tasks are typically complex, somewhat ill-defined, engaging problems that require students to apply, synthesize, and evaluate various problem solving approaches (Shavelson, Baxter, and Pine, 1991; Wiggins, 1989). For example, a team of students could be assigned the task of conducting a cost/benefit feasibility study for a recycling program at a local business. Such tasks are clearly different in nature, form, and length from multiple-choice questions that can usually be responded to in a matter of seconds.

Performance assessments use grading strategies that are commonly used in the performing arts, fine arts, and Olympic competitions. In the context of the science laboratory, students are graded on the performance of manipulating variables, using scientific apparatus, identifying hypotheses, making measurements and calculations, organizing and managing data, and the communication of results (Slater and Ryan, 1993). Graded laboratory performances go far beyond grading a final field report - this strategy considers the processes that become the laboratory report as well. For example in geology, the manipulation of a Brunton compass to make strike and dip measurements can be a graded task as part of a larger group-mapping project. In the evaluation of a performance task, the process of performing the task is emphasized more than the final product itself.

Studies that have looked closely at performance assessments find that, if the criteria is clear and that examples are available to show levels of competency, performance assessments are highly consistent across different evaluators (Kulm and Malcom, 1991; O'Neil, 1992). Moreover, the clear indication of what is expected of students improves student performance. There are, however, some indications at the K-12 levels that students perform inconsistently from one performance task to the next (Shavelson, Baxter, & Pine, 1991). This suggests that student grades will be most reliably determined from a number of performance assessments in concert with other forms of assessment.

Links

Timothy F. Slater, Research Assistant Professor of Physics, Montana State University. Interests: authentic assessment in support of student-centered instruction; physics/astronomy education for both K-16 levels and public awareness; Internet-based scientific investigations involving real-time earth/space science data; and teacher-enhancement.
Web site: solar.physics.montana.edu/tslater

Performance Assessment examples on the Internet (mostly from K-12)

1. http://nces.ed.gov/nationsreportcard/ITMRLS/intro.shtml
2. http://www.exemplars.com/preview/science/page20.html
3. http://www.sltech.com/
4. http://www-ed.fnal.gov/linc/fall95/proj_assess/assessment_other1.html
5. http://www.manhattan.k12.ks.us/dms/index1.html

Sources

- Kulm, Gerald & Malcom, Shirley M. (1991) Science Assessment in the Service of Reform. American Association for the Advancement of Science, Washington, DC.
- Shavelson R.J., Baxter, G.P. & Pine J. (1991) Performance assessment in science. Applied Measurement in Education, 4(4): 347.
- Slater, T.F. & Ryan, J.M. (1993) Laboratory performance assessment. The Physics Teacher, 31(5): 306-309.
- Tobias, S. & Raphael, J. (1995) In-class examinations in college science - new theory, new practice. Journal of College Science Teaching, 24(4): 240-244.
- Wiggins, G. (1989) A true test: Toward a more authentic and equitable assessment. Phi Delta Kappan, 70(9): 703.

Tim Slater

While I was a graduate teaching assistant in astronomy, I sympathized with students who told me that there were two ways of taking college science classes. One was to learn and understand the material and the other was to get an "A." The students well understood that the most productive strategy for getting a high grade in most introductory science courses involved memorizing the notes from lecture, working enough homework problems so that the proper algorithm could be applied to the corresponding problem on exams, and subsequently forget everything they had temporarily memorized.

Eventually, I began to understand for myself that I was not going to take a multiple-choice examination to pass my Ph.D. defense anymore than my tenure and promotion as a faculty member would depend on how many facts I had memorized. I realized that what I loved about doing science was DOING science and focusing on the aspects most interesting to me. About that same time, the National Council for Teachers of Mathematics (NCTM) began to extol the virtues of alternative assessments as a way of moving students beyond memorizing procedures and motivating them to understand concepts. I began exploring ways to adopt the procedures already well understood in the performing arts such as music and athletics such as diving, gymnastics, and skating - competitive ice skaters Tonya Harding and Nancy Kerigan were in the news quite often in those days - to the excitement of scientific inquiry. Over the years, I have used performance assessment at a variety of levels of commitment. I now believe that, in the same way that students must actively construct their own knowledge with considerable mental effort, the use of performance assessments emphasize to students that procedural knowledge and creative problem solving is at least as important as basic factual knowledge.