

Performance Assessment

In its simplest terms, a performance assessment is one which requires students to demonstrate that they have mastered specific skills and competencies by performing or producing something.³ Advocates of performance assessment call for assessments of the following kind: designing and carrying out experiments; writing essays which require students to rethink, to integrate, or to apply information; working with other students to accomplish tasks; demonstrating proficiency in using a piece of equipment or a technique; building models; developing, interpreting, and using maps; making collections; writing term papers, critiques, poems, or short stories; giving speeches; playing musical instruments; participating in oral examinations; developing portfolios; developing athletic skills or routines, etc.

Authentic Assessment

Similar to performance assessment is the concept of authentic assessment. Meyer (1992) notes that performance and authentic assessments are not the same, and that a performance is “authentic” to the extent it is based on challenging and engaging tasks which resemble the context in which adults do their work. In practical terms, this means that an authentic task or assessment is one in which students are allowed adequate time to plan, to complete the work, to self-assess, to revise, and to consult with others. Meyer also contends that authentic assessments must be judged by the same kinds of criteria (standards) which are used to judge adult performance on similar tasks.

A more elaborate definition of authenticity is offered by Wiggins (1990, CLASS), who suggests that three factors determine the authenticity of an assessment: the task, the context, and the evaluation criteria. An authentic task is one which requires the student to use knowledge or skills to produce a product or complete a performance. Based on this definition, memorizing a formula would not be an authentic task; however, using the formula to solve a practical problem would be.

As for context, Wiggins suggests that there be as much realism as is possible. He maintains that the setting (including the time allowed to complete the task) should mimic or duplicate the context faced by professionals, citizens, and consumers. An examination in which the student has almost no prior knowledge of what will be asked, little time to complete the activity, and no opportunity to reflect or consult appropriate resources would not be authentic.

Finally, Wiggins states that an authentic assessment should be judged using criteria which are similar to those used to judge adults who perform or produce. As an example, authentic criteria used to evaluate a written paper would give primary consideration to the paper’s organization and ideas; mechanical errors (such as spelling, punctuation, grammar) would not be the primary focus.

What is to be made of the distinction between performance and authentic assessments? Fortier (1993) notes that authenticity is always a relative concept and that it is unrealistic to expect that an assessment will be completely authentic. For example, he points out that a driving test, even though most would define it as authentic when compared with a paper and pencil test, can never be completely such because drivers do not ordinarily have a law officer seated next to them while they drive.

In short, as the term is used in the literature, an authentic performance assessment requires students to demonstrate skills and competencies which realistically represent those needed for success in the daily lives of adults. Authentic tasks are worth repeating and practicing. They require students to apply what they know, not merely to recall or recognize information. Finally, authentic tasks are those which are judged by criteria or standards similar to those used to evaluate the efforts of adults.

Authentic Instruction and Learning

Similar to performance or authentic assessment is the term authentic learning and instruction. Although this term refers to instruction and learning, it is appropriate to discuss it within the framework of assessment because those who call for changes in either assessment or instruction maintain that assessment and instruction must be integrated. In a 1993 article in *Educational Leadership*, Newmann and Wehlage use the concept “authentic instruction” to describe instruction which results in significant and meaningful student achievement, in contrast with that which is trivial and useless.⁴

In particular, Newmann and Wehlage maintain that instruction is authentic if it helps students achieve three broad goals:

1. construct meaning and produce knowledge,
2. use disciplined inquiry to construct meaning, and
3. aim work toward production of discourse, products, and performances that have value or meaning beyond success in school

To help the reader understand the concept of authentic instruction, the authors offer five standards or criteria, each based on a five-point scale, which can be used to evaluate the extent to which a lesson is authentic. These criteria, with explanations in parentheses, are as follow

1. Students employ higher-order thinking skills (students apply knowledge and skills to solve problems, to synthesize, to explain, etc.)
2. Depth of Knowledge (understanding of a concept, topic, or skill is not superficial)
3. Connectedness to the World (problems/topics are ones which occur in the larger society/ world)
4. Substantive Conversation (teacher-student conversation is two-way and meaningful)
5. Social Support for Student Achievement (the teacher, school and community expect all students to achieve)

Performance Criteria ⁵

Advocates of performance assessments maintain that every task must have performance criteria for at least two reasons: (1) the criteria define for students and others the type of behavior or attributes of a product which are expected, and (2) a well-defined scoring system allows the teacher, the students, and others to evaluate a performance or product as objectively as possible. If performance criteria are well defined, another person acting independently will award a student essentially the same score. Furthermore, well-written performance criteria will allow the teacher to be consistent in scoring over time.

Stiggins (1991) notes that if a teacher fails to have a clear sense of the full dimensions of performance, ranging from poor or unacceptable to exemplary, he or she will not be able to teach students to perform at the highest levels or help students to evaluate their own performance.

In developing performance criteria, Stiggins maintains that one must both define the attribute(s) being evaluated and also develop a performance continuum. For example, one attribute in the evaluation of writing might be writing mechanics, defined as the extent to which the student correctly uses proper grammar, punctuation, and spelling. As for the performance dimension, it can range from high quality (well-organized, good transitions with few errors) to low quality (so many errors that the paper is difficult to read and understand).

The key to developing performance criteria, asserts Stiggins, is to place oneself in the hypothetical situation of having to give feedback to a student who has performed poorly on a task. Stiggins suggests that a teacher should be able to tell the student exactly what must be done to receive a higher score. If

performance criteria are well defined (with examples provided whenever possible), the student then will understand what he or she must do to improve.

It is possible, of course, to develop performance criteria for almost any of the characteristics or attributes of a performance or product. However, experts in developing performance criteria warn against evaluating those aspects of a performance or product which are easily measured (such as counting mechanical errors) or failing to distinguish between quality and quantity. Ultimately, it is asserted, performances and products must be judged on those attributes which are most crucial.

Portfolios

Invariably, proponents of performance assessment also advocate the use of student portfolios. In doing so, they also remind us that a portfolio is more than a folder stuffed with student papers, video tapes, progress reports, or related materials. It must be a purposeful collection of student work that tells the story of a student's efforts, progress, or achievement in a given area over a period of time. If it is to be useful, specific design criteria also must be used to create and maintain a portfolio system.

Typically, proponents of portfolios suggest two reasons for their use. The first reason reflects dissatisfaction with the kind of information typically provided to students, parents, teachers, and members of the community about what students have learned or are able to do. As examples, we are reminded that traditional grading systems ("A's", "B's", etc.) or test scores (percentile scores or percent correct) tell us almost nothing about what a student has learned or is able to do.

Second, it is argued that a well-designed portfolio system, which requires students to participate in the selection process and to think about their work, can accomplish several important purposes: it can motivate students; it can provide explicit examples to parents, teachers, and others of what students know and are able to do; it allows students to chart their growth over time and to self-assess their progress; and, it encourages students to engage in self-reflection.

Frazier and Paulson (1992) argue that the primary worth of portfolios is that they allow students the opportunity to evaluate their work. Further, "... portfolio assessment offers students a way to take charge of their learning; it also encourages ownership, pride, and high self-esteem" (p. 64).

Vavrus (1990) notes that several decisions must be addressed prior to establishing a portfolio system. The decisions, with some of her recommendations, follow.

1. What will it look like? There must be a physical and a conceptual structure.
"The physical structure refers to the actual arrangement of documents used to demonstrate student progress... The conceptual structure refers to your underlying goals for student learning" (p. 50).
2. What goes in? In order to make this decision, numerous other questions need to be addressed: Who is the intended audience for the portfolios? Parents? Administrators? Other teachers? What will this audience want to know about student learning? Will the selected documents show aspects of student growth that test scores don't capture? ...What kinds of evidence will best show student progress toward your identified learning goals? Will the portfolio contain best work only, a progressive record of student growth, or both? Will the portfolio include more than finished pieces: for example, ideas, sketches, and revisions? (p. 50).
3. How and when to select? Decisions need to be made as to when documents go in and come out during the school year. It is recommended that specific times during the year be identified for selecting student work. In addition, student participation in the selection process is critical, for this allows students to reflect on their work and monitor their progress. It also is suggested that materials which are included be dated and include an explanation for their inclusion.

4. Evaluating Portfolios.⁶ If portfolios are to be evaluated, the evaluation standards should be established before the portfolio system is established. As for the evaluation itself, "...portfolios can be evaluated in terms of standards of excellence or on growth demonstrated within an individual portfolio, rather than on comparisons made among different students' work" (p. 53).
5. Passing Portfolios on. The final decision item has to do with what is done with portfolios at the end of a semester or school year. They could, of course, be turned over to students. However, there are advantages to keeping portfolios over a long period of time and sharing them with other teachers. "Portfolios give you opportunities to promote continuity in your students' educations and to collaborate with other teachers and your students in the process. By passing a portfolio on, you can share important information with the student's next teacher" (p. 53). Dennie Palmer Wolf (1993) also feels that portfolios should be kept for long periods of time (several years), and that they should act as a type of "passport" as a student moves from one level of instruction to another.

As an example of a well-designed portfolio system, one only need look to the Walden III Alternative High School in Racine, Wisconsin. This school requires that all students complete a portfolio before graduation and also demonstrate mastery of the sixteen topics addressed in the portfolio by giving presentations before a committee consisting of staff, another student, and an outside adult.⁷ Among the purposes of the portfolio is to require graduating seniors to look at themselves in-depth. Although students are encouraged to be creative and unique, the portfolio has several basic requirements regarding its form and content. For example, it must be well written, with a title page, table of contents, and specific headings. Students also are encouraged to include photographs, charts, drawings, and appendices along with samples of their work.

Developing Performance Tasks

Developing performance tasks or performance assessments seems reasonably straightforward, for the process consists of only three steps.⁸ The reality, however, is that quality performance tasks are difficult to develop. With this caveat in mind, the three steps, with a brief discussion of each, follow.

Step 1. List the skills and knowledge you wish to have students learn as a result of completing a task.

As tasks are designed, one should begin by identifying the types of knowledge and skills students are expected to learn and practice. These should be of high value, worth teaching to, and worth learning. In order to be authentic, they should be similar to those which are faced by adults in their daily lives and work.

Herman, Aschbacher, and Winters (1992, pp. 25-26) suggest that educators need to ask themselves five questions as they identify what is to be learned or practiced by completing a performance task. Their questions, with examples, follow:

1. What important cognitive skills or attributes do I want my students to develop? (e.g., to communicate effectively in writing; to analyze issues using primary source and reference materials; to use algebra to solve everyday problems).
2. What social and affective skills or attributes do I want my students to develop? (e.g., to work independently, to work cooperatively with others, to have confidence in their abilities, to be conscientious).
3. What metacognitive skills do I want my students to develop? (to reflect on the writing process they use; to evaluate the effectiveness of their research strategies, to review their progress over time).

4. What types of problems do I want them to be able to solve? (to undertake research, to understand the types of practical problems that geometry will help them solve, to solve problems which have no single, correct answer)
5. What concepts and principles do I want my students to be able to apply? (e.g., to understand cause-and-effect relationships, to apply principles of ecology and conservation in everyday lives).

Step 2. Design a performance task which requires the students to demonstrate these skills and knowledge. The performance tasks should motivate students. They also should be challenging, yet achievable. That is, they must be designed so that students are able to complete them successfully. In addition, one should seek to design tasks with sufficient depth and breadth so that valid generalizations about overall student competence can be made.

Step 2. Design a performance task which requires the students to demonstrate these skills and knowledge. The performance tasks should motivate students. They also should be challenging, yet achievable. That is, they must be designed so that students are able to complete them successfully. In addition, one should seek to design tasks with sufficient depth and breadth so that valid generalizations about overall student competence can be made.

Herman, Aschbacher, and Winters (p. 31) have a list of questions which are helpful in guiding the process of developing performance tasks. Those questions, with their recommendations, follow:

1. How much time will it take students to develop or acquire the skill or accomplishment? The authors recommend that assessment tasks should take at least one week for students to complete. Others recommend that worthwhile tasks require far more time.
2. There are no rules regarding the appropriate length or complexity of a task; however, there are problems associated with developing overly complex and creative performance tasks (Cronin, 1993). To begin with, relatively modest performance tasks are easier to develop. Furthermore, if they are well crafted and reasonably short (a few days rather than a few weeks), they are more likely to hold the interest of students. Finally, if a task fails to accomplish its purposes, it is best if the task is limited in duration.
3. How does the desired skill or accomplishment relate to other complex cognitive, social, and affective skills? Priority should be given to those which apply to a variety of situations.
4. How does the desired skill or accomplishment relate to long-term school and curricular goals? Skills or accomplishments which are integral to long-range goals should receive the most attention.
5. How does the desired skill relate to the school improvement plan? Priority should be given to those which are valued in the plan.
6. What is the intrinsic importance of the desired skills or accomplishment? Emphasis should be given to those which are important, while others should be eliminated.
7. Are the desired skills and accomplishments teachable and attainable for your students? Priority should be given to tasks which represent realistic goals for teaching and learning.

Step 3. Develop explicit performance criteria which measure the extent to which students have mastered the skills and knowledge.

It is recommended that there be a scoring system for each performance task. The performance criteria consist of a set of score points which define in explicit terms the range of student performance. Well-defined performance criteria will indicate to students what sorts of processes and products are required to show mastery and also will provide the teacher with an “objective” scoring guide for evaluating student work. The performance criteria should be based on those attributes of a product or performance which are most critical to attaining mastery. It also is recommended that students be provided with examples of high quality work, so they can see what is expected of them.

Additional Recommendations for Developing Performance Tasks

- * Keep in mind that the concepts of performance /authentic assessments are not new. Teachers always have assigned tasks which require their students to perform or develop products.
- * If possible, groups of educators should work together to design performance tasks. Tasks are more likely to be interdisciplinary. In addition, this process allows for discussion and exchange of ideas.
- * Develop tasks which are fair and free of bias. Tasks should not give particular advantage to certain students.
- * Develop tasks which are interesting , challenging, and achievable. This means that the tasks should be neither too complex and demanding, nor too simple or routine.
- * Develop tasks which are maximally self-sustaining, with clear, step-by-step directions and with the record-keeping responsibilities placed mostly on the students. If this is done, the teacher need not guide activity every step of the way and record massive amounts of information throughout the process.

An Example of a Performance Task

The last part of this paper presents an example of a performance task which requires students to interview adults and develop written and oral reports. Although the task is intended for use at the secondary level, the format is appropriate for younger children. Thus, one could modify the sample task to have elementary students study a community's history or investigate important community issues or problems.

Along with this task are two examples of performance criteria which could be used to evaluate the student's written assignment. Similar performance criteria would have to be developed if the other skills involved in this task were to be evaluated, such as interviewing, speaking, and working cooperatively with others.

The sample performance task, *The Effect of the Great Depression on the Lives of Average People*, was developed for use at the secondary level.⁹ In addition to having students learn more about the 1930's Depression, the task is designed to help students learn and practice the following kinds of skills: developing questionnaires; interviewing, taking notes and transcribing them; working with other students; analyzing data (questionnaire responses); developing conclusions, generalizations, and hypotheses; giving an oral presentation; and writing a report. This task also brings students into contact with members of the community.

The task consists of six steps:

1. As a group, members of the class must develop a common set of procedures, interview questions, and a questionnaire. This questionnaire will be used by each student to interview an adult who lived through the Great Depression (see Step 2). The questions to be asked should be based on the kind of information class members wish to gather. (Examples could include the following: How did the Depression affect your family? You personally? Others you knew or heard about? What were job opportunities like? As you think back, how did the experiences of the Great Depression affect your values and behaviors as an adult?).
2. Each student identifies and interviews one person who lived through the Depression. During the interview process, the student should take careful notes. (Note: students may wish to tape record or video tape their interview). After the interview is completed, the student should summarize the process and responses in written form.
3. After all interviews are completed, students are divided into small groups (of perhaps four or five persons) and asked to discuss (compare and contrast) their experiences and to reach some conclusions about what they learned.
4. Each student must consult the written accounts of at least two historians who wrote about the Depression to determine the extent to which the historians' descriptions of the period match the group's conclusions. Similarities and differences should be identified and explained.
5. Each group develops an oral presentation (about 30 minutes in length) in which each student has a role in the presentation. In this report to the class, students should be encouraged to use video and/or audio materials, overhead transparencies, etc. Time for questions should be allowed.
6. Each student is required to write a brief report (perhaps 5–6 pages in length) in which the student summarizes the entire experience.

Performance Criteria

Performance criteria for the written report, as described in Step 6, follow. The criteria are of two kinds: one for writing mechanics, the other for content.¹⁰

Scoring Criteria (Mechanics)

4 = the paper is easy to read and uses appropriate format. It is carefully proofread to correct spelling, capitalization, punctuation and usage errors. It is written in complete sentences and uses paragraphs correctly.

3 = the paper is generally well proofread and uses appropriate format but has occasional minor lapses.

2 = the paper may lack the appropriate format. It is proofread but may display errors in spelling, capitalization, punctuation and usage. It is written in complete sentences but may not be paragraphed correctly.

1 = The paper is poorly presented, indicating the author is unaware of the requirements of written communications. It will have a significant number of proofreading errors, sentence fragments, and/or flaws in usage.

0 = The student failed to attempt the paper.

Scoring Criteria (Content)

4 = The paper is written in a style appropriate to the genre being assessed. It is well organized, clearly written, and meets the needs of the author and reader. It will contain sufficient details, examples, descriptions and insights to engage the reader. The author will bring closure through a resolution of a problem or a summary of the topic.

3 = The paper is written in an appropriate style and format. It may appear to be well organized and clearly written but may demonstrate minor lapses in the communication to the reader. It may be missing some details and/or examples, and offer incomplete descriptions and fewer insights into the characters and/or topics. The author may not sufficiently close the piece of writing and may leave the reader “hanging” or may offer the reader an inappropriate closing or ending.

2 = The paper may demonstrate an incomplete or inadequate knowledge of the skills assessed. Significant flaws may be evident as the author fails to address the prompt in an appropriate manner, ideas may be conveyed in a random method, and very little is given in proof, details, facts, examples or descriptions. Closure is often missing.

1 = The paper will barely attempt the task. The general idea may be conveyed, but there will be a definite lack of understanding on the part of the author regarding the appropriate format or procedures.

0 = The student failed to attempt the paper.

Conclusion

This paper began with a discussion of the criticisms made of current assessment and instructional practices. It was noted that critics maintain that we often fail to teach and assess the kinds of skills and knowledge which have lasting value for students.

Following this was a discussion of the kinds of weaknesses ascribed to standardized achievement tests. However, it was pointed out that many tests provided by publishers of textbooks and teacher-developed tests have similar weaknesses and limitations. As for instruction, it was noted that too much of instruction remains teacher-dominated (with lectures), and that students all too frequently are taught subject matter content at the expense of important skills.

At the end of this paper an example of a performance task was offered. This task requires students to learn and demonstrate a variety of skills and knowledge, ranging from developing questionnaires and interviewing adults to analyzing and reporting data.

This exercise was presented in order to show how teachers might design performance-based instruction and assessment for use in their classroom. This type of performance task has several positive features. It engages the learner, rather than having the teacher dominate the learning process and tell students what is important. It also illustrates how instruction and assessment can be integrated. In addition, this type of performance task meets the definition of authenticity, for it replicates the kind of work done by many students. Finally, and perhaps most important, it teaches students the kinds of skills and knowledge which we want them to master.