Overview of Authentic Assessment

David Johnston

Educational Psychology 6302

Educators must have some form of assessment. Instructors decide whether students are ready for a new unit, administrators judge a teacher's effectiveness, and universities choose among numerous applicants. Developing assessment methods is not difficult. The challenge lies in discovering methods o judgement which are fair, accurate, and affordable.

Grant Wiggins, a frequent critic of current testing methods, said report cards and transcripts are insufficient measures of performance. "An 'A' in English means only that some adult thought the student's work was excellent. Compared to what or whom? As determined by what criteria? … [A diploma] provides no useful information about what students have studied or what they can actually do with what was studied." (Wiggins, 1989, p.41) The current answer to subjective grades in the standardized test. Though common throughout education, standardized exams are subject to their own criticism.

The Shortcomings of Standardized Tests

Many standardized exams are biased against genders or racial groups. Lower-income students do not perform as well and are often less familiar with the test format. Teachers can (and often do) "teach the test," or help students develop skills less related to the curriculum and more related to the upcoming multiple-choice test. A survey of teachers showed they believe standardized exams use unclear questions, do not provide useful information, and fail to address real educational concerns. Teachers also claimed pressure to teach the test limited student's education. (Perrone, 1991, p.1)

Critics assert these tests are overused. Instead of testing children to gauge their level of factual retention or logic ability, schools may use multiple-choice exams to place students or

measure a program's success. This is often beyond the scope of the exam's original design. "Traditional tests tend to reveal only whether the student can recognize, recall, or 'plug in' wha was learned out of context." (Wiggins, 1990, p.1) Wiggins compares this process to gauging driving or teaching ability with a written exam. Standardized tests cannot provide the information sought. This use can lower students' self-concept or decrease teachers' expectations (which leads to lower performance). (Perrone, 1991, p.2)

Standardized tests are often shallow – looking at answers, not the motive or reasoning behind them. (Wiggins, 1990, p.1) The format of the test can create additional problems. Students who are unfamiliar with the multiple-choice format are at a disadvantage. The exam material may not interest students, meaning the test gauges attention spans rather than performance. The test environment conflicts with students' learning environment. "Children who have been routinely encouraged to be cooperative learners are forbidden to talk during testing. Children who have been taught to work problems out slowly are told speed is essential." (Perrone, 1991, p.2)

> Perrone sums up the environment created by the widespread use of standardized tests: If tests play a significant role in grade advancement or are the primary basis for a school's so-called accountability, teachers feel compelled to spend considerable time preparing children to take the tests. In such cases, the tests become the schoo curriculum. Preparation usually begins many weeks before actual testing. During this period, two to three hours a day are often devoted to practicing tests and related exercises, all alien to the ongoing instruction and the usual student response patterns. Teachers readily acknowledge that questions in the practice exercises, which are similar to those on the real test, are trivial … By the time the three days of real testing are over,

weeks, sometimes months, have passed. Time for real books has been sacrificed for time

spent reading isolated paragraphs and answering multiple-choice questions. (Perrone,

1991, p.2)

The past two decades have seen an increase in the cumber of educa ors calling for an

alternative to traditional testing methods. In 1976, the Association for Childhood Education

International recommended schools stop administering standardized exams in lower grades.

Sentiment has continued to turn against standardized tests, but they are still widely used.

(Perrone, 1991, p.1) The exams are objectively graded, inexpensive to score, and simple to

administer. Any proposed replacement for the traditional test must prove its economic and

practical viability.

Authentic Assessmen

Grant Wiggins proposed a change in testing and assessment in a 1989 article. He

suggested teachers evaluate students' performance in "real-world" tasks. Instead of testing

material tangential to educational goals (evaluating a future accountant's multiplication skills,

perhaps), Wiggins urged educators to create exams approximating the tasks the student will face

outside of school (for example, asking the accounting student to prepare tax returns for sample

clients).

Wiggins described his solution to the assessment problem as "straight but steep." H

proposed teachers "make [students] replicate, within reason, the challenges at the heart of each

academic discipline … and worry about a fair, efficient, and objective method of grading them as

a *secondary* problem." (Wiggins, 1989, p.41) Wiggins suggested a system that integrated

teaching and testing. Just as a sports team trains for a game, Wiggins believes students should train for their future performance – perhaps a presentation, recital, or debate. Though the performance can be evaluated, the primary goal is the performance itself. Wiggins points out teacher will still teach to the test, but since the test is actual performance, curriculum matches the assessment method. (Wiggins, 1989, p.43)

Wiggins authentic test calls for a public performance of some kind. Perhaps a pane interview (similar to a doctoral defense) or a presentation of research. Students would be full aware of the assessment methods. The system would judge their overall performance instead o a score on a single exam. Wiggins emphasized the integration of self-evaluation into the student's overall assessment. Assessment standards would be somewhat subjective, and would emphasize use of essential academic skills. (Wiggins, 1989, p.45)

Wiggins recommended assignments such as research projects, conducting oral histories, or writing journals. Portfolios have also become closely associated with authentic assessment. A portfolio normally consists of collected samples of a student's work. A high schoo implementing authentic assessment, for example, might require graduating seniors to present a portfolio summarizing their performance over the past four years. This might include a resume, a video of a speech, a copy of a research paper, samples of artwork or an essay on aesthetics, or other items representing a student's achievements. (Archbald, 1991, p.19) The portfolio shoul demonstrate student achievement in a range of academic pursuits, and it should show provide depth of assessment than a transcript or test score can convey.

Since authentic assignments are open-ended and call for some level of aesthetic performance, the quality of the results can be very subjective. Developing a fair means o

evaluation can be a challenge. Though the test contains subjective elements, instructors need a simple, effective method for objectively rating a student.

Scores on authentic tests should be based on performance against predetermined standards – not the performance of other students. This contrasts with many standardized tests where scores are evaluated by a distribution or curve. Scoring of an authentic performance should also be multifaceted. For example, a teacher evaluating a student report might evaluate the student's writing style, speaking ability, and research skills. Instructors should develop a se of standards for each area of assessment. These standards should define a range of performances that would deserve a certain grade or point value.

For example, when assessing the written portion of a project, the instructor may decide a "five-point paper" shows consistent style, contains no grammatical or spelling errors, and accurately and completely covers the topic. A four-point paper might show a few grammatica errors. A three-point paper might have serious errors or contain logical inconsistencies and so on. Similar criteria would be established for each facet of the assignment to be measured. sample scoring sheet is shown below.

In the sample below, students' oral presentations would be evaluated in four areas. General guidelines for point values in each area are described in the table. Point values would be summed from all four areas to determine an individual student's overall score.

Table 1

Score levels	Multimedia *The integration of media objects such as text, graphics, video, animation, and sound to represent and convey information. Videotapes that include sound and images fit this definition.*	Collaboration *Working together jointly to accomplish a common intellectual purpose in a manner superior to what might have been accomplished working alone.*	Content *The topics, ideas, concepts, knowledge, and opinions that constitute the substance of the presentation.*
5	Students have used multimedia in creative and effective ways that exploit the particular strengths of the chosen format. All elements make a contribution. There are few technical problems and none of a serious nature.	Students were a very effective team. Division of responsibilities capitalized on the strengths of each team member. The final product was shaped by all members and represents something that would not have been possible to accomplish working alone.	Meets all criteria of the previous level and one or more of the following: reflects broad research and application of critical thinking skills; shows notable insight or understanding of the topic; compels the audience's attention.
4	Presentation blends three or more multimedia elements in an attractive, easy-to-follow format. With minor exceptions, all elements contribute rather than detract from the presentation's overall effectiveness.	Students worked together as a team on all aspects of the project. There was an effort to assign roles based on the skills/talents of individual members. All members strove to fulfill their responsibilities.	The project has a clear goal related to a significant topic or issue. Information included has been compiled from several relevant sources. The project is useful to an audience beyond the students who created it.
3	Presentation uses two or more media. There are some technical problems, but the viewer is able to follow the presentation with few difficulties.	Students worked together on the project as a team with defined roles to play. Most members fulfilled thei responsibilities. Disagreements were resolved or managed productively.	The project presents information in an accurate and organized manner that can be understood by the intended audience. There is a focus maintained throughout the piece.
2	Presentation uses two or more media, but technical difficulties seriously interfere with the viewer's ability to see, hear, or understand the content.	Presentation is the result of a group effort, but only some members of the group contributed. There is evidence of poor communication, unresolved conflict, or failure to collaborate on important aspects of the work.	The project has a focus, but may stray from it at times. There is an organizational structure, though it may not be carried through consistently. There may be minor factual errors or inconsistencies.
1	Multimedia is absent from the presentation.	Presentation was created by one student working more or less alone (though may have received guidance or help from others).	Project seems haphazard, hurried, or unfinished. There are significant factual errors, misconceptions, or misunderstandings.
	Multimedia Score =	Collaboration Score =	Content Score =

Adapted from the Challenge 2000 Multimedia Project.

As interest in authentic assessment has grown, it has become necessary to describe wha

constitutes an authentic assignment. The general description of an authentic project has alread

been given; an open-ended project designed to approximate a real-life problem, which requires students to apply a range of knowledge and skill. These projects may include experiments, computations or writings, but the emphasis is always on student-constructed answers. Still, some assignments may be considered more "authentic" than others.

For Wiggins, an authentic test always yielded an independently valuable product – something with worth beyond student evaluation, such as a report, portfolio, artwork, or recital. (Wiggins, 1990, p.3) The assessment itself should have value. "The learning they measure has value beyond the classroom and is meaningful to the learner." (Kerka, 1995, p.1) This criteri would mean experiments or math computations would normally be acceptable authentic assignments only if they were components of a larger project.

An effective assignment should remove focus from trivial or tangential skills or knowledge to focus on the skills essential to the academic area. (Wiggins, 1990, p.3) An assignment is more authentic the closer it mirrors a problem students will face outside the classroom. Real world problems usually do not require the recitation of rote factual information, but call for the application of knowledge upon a new context. Of course, an authentic assessment should adequately cover the domain and attempt to predict future performance, (Elliot, 1995, p.2) but primarily it should mirror actual experiences students will face in the future.

Authentic assignments should avoid the contrived nature of traditional tests. The test should parallel learning. If a teacher wants students to be able to logically evaluate an argument, the instructor might assign a debate. An effective authentic assignment directly tests desired skills while preparing students to advance beyond the classroom.

As with traditional testing, the scoring method should not place any group or individual at a disadvantage. A scoring algorithm can be applied to a diverse group (with proper controls)

to determine its equity. If certain groups or individuals are disadvantaged by the criteria, the assessment has not helped to correct the biases of many traditional tests. The evaluation shoul not evaluate parents or instructors. The scoring method must focus on performance of academic skills, not personalities or other distracters. "Tests are more authentic to the extent they call for depth of understanding, disciplinary grounding, and connection to extracurricular work." (Harrington, 1993, p. 66)

An authentic project contrasts clearly with traditional assessment. Authentic assessmen requires an independent performance similar to real-world problems, which allows students to demonstrate deep understanding of essential academic skills.

<u>Advantages to Authentic Assessment</u>

Authentic assessment provides students an opportunity to fully demonstrate their abilities in a range of areas. (Kerka, 1990, p.1) The system can be better integrated into education. Authentic assessment allows preparation that might otherwise be called "teaching the test," and it permits collaboration that might constitute cheating in a traditional environment. (Archbald, 1991, p.11)

The standards for authentic assessment are clearly defined and presented to students in advance. Students know what is expected of them and can better prepare for their performance exam. Students can, in a sense, contract their grade – choosing to perform to a particular level in exchange for a certain level of credit. "In authentic assessment, students have greater clarity about their obligations and are asked to master more engaging tasks, and teachers can see assessment results as meaningful and useful for improving instruction." (Wiggins, 1990, p.2)

Authentic assessment examines higher-level thinking skills, unlike traditional testing tha often evaluates factual recall. "Research is quite clear that standardized tests have serious limitations for assessing performance on open-ended problem-solving tasks – tasks involving creativity, sustained inquiry, and analysis of diffuse problems." (Archbald, 1991, p.7) These are the areas attended by authentic tests. Students cannot perform well on authentic tests withou demonstrating an ability to take their knowledge base and creatively apply it to a new problem. Authentic assessments require higher abilities than traditional methods of evaluation.

Institutions of learning aim to instill students with the talents needed after graduation. Traditional tests (and the curriculum they define) do not stress those talents. The cognitive demands of authentic assignments more closely match the cognitive demands of the workplace. (Archbald, 1991, p.7) Authentic assignments help students not only develop job skills, bu rehearse real-world assignments.

Research suggests authentic assessments develop stronger students. Traditional tests often limit students. Instructors who implemented authentic assessments in the classroom were surprised by their students' performances. (Allen, 1998, p.5) Since authentic assignments usually incorporate student choice, students select activities that engage their interest. As a result, students often invest more in their projects and accomplish more than they would have otherwise.

Effective authentic assessment brings many advantages to the classroom. Students perform better, receive richer experiences, and demonstrate more significant talents. Despite its benefits, authentic testing is not as widespread as its traditional counterpart, and authentic assessment must overcome hurdles before it is more widely implemented.

Disadvantages of Authentic Assessment

The largest drawback of authentic assessment is its high cost. Wiggins, one of the biggest promoters of authentic assessment, estimates scoring authentic tests cost 200 times the price of scoring a traditional exam. For every penny spent on traditional tests, an authentic alternative would cost $2. (Wiggins, 1990, p.4) Most of this cost is in time and development o authentic programs. The change to authentic assessment represents a significant investment for educators.

Some critics claim authentic assessment fails to remedy the problems of standardized testing. Authentic assessments can be as contrived as their traditional counterparts. (Terwilliger, 1996, p.6) "There is no absolute distinction between performance tests and other classes of tests." (Terwilliger, 1996, p.13)

Promotion of authentic assessment is primarily based on the failures of standardized tests. The academic world, however, does not unanimously denounce traditional testing. "While [traditional exams] are imperfect, these tests effectively provide information about student standing relative to peers on the domains tested." (Archbald, 1991, p.6) Some critics believe authentic tests cannot require the same factual knowledge as traditional assessment. (Terwilliger, 1996, p.11)

Authentic assessment may claim to address some problems in the current classroom test, but not without presenting its own dilemmas.

Conclusions

Any assessment method will evolve into the curriculum. Educators should understand this principle, and they then begin to design tests accordingly. Authentic assessment represents another option for instructors to add to their repertoire. It may not work in every situation, but may often prove useful.

Much of authentic assessment's high cost will disappear with time. The cost o developing scoring rubrics and training teachers will decrease as the method becomes more popular. Even so, authentic scoring will never be as simple or inexpensive as traditional scoring.

One research study compared the performance of fifth-grade students on traditional and authentic assessments. Of the 106 students in the study, 74 scored better on the authentic exam. (Cross et al, 1998, p.9) The researchers suggested this supported authentic means as superior, but the fact alone is value-neutral. What is significant, however, is that only 4 students scored the same on both tests. While the study does not prove alternative assessment is better or worse than traditional testing, it does indicate the two methods test different elements.

Authentic assessment is a different creature with its own strengths and weaknesses. Understanding these differences can lead to successful application of the available methods. Research will not reveal a simple solution to assessment problems, but will introduce new options.

References

Allen, Nancy J. (1998). Not in Our State: Authentic Assessment for Native American Students Using Digital Photography.San Diego, Ca.: National Association for Research in Science Teaching (ERIC Document Reproduction Service No. ED 418 851)

Archbald, Doug A. (1991). Authentic Assessment: What it Means and How it Can Help Schools. Madison, WI: National Center for Effective Schools Research and Development (ERIC Document Reproduction Service No. ED 366 624)

Brualdi, Amy (1998). Implementing Performance Assessment in the Classroom. Washington, D.C. (ERIC Document Reproduction Service No. ED 423 312)

The Challenge 2000 Multimedia Project. <http://pblmm.k12.ca.us/PBLGuide/Mmrubric.htm>

Cross, B. Jill; Greer, Tracy; Pearce, Mary (1998). Improving Student Reading Comprehension Skills Through the Use of Authentic Assessment. St. Xavier University. (ERIC Document Reproduction Service No. ED 420 841)

Elliot, Stephen N. (1995). Creating Meaningful Performance Assessments. Reston, VA (ERIC Document Reproduction Service No. ED 381 985)

Harrington, Diane (1993). Authentic Teaching and Assessment: Policy and Practice. Examples from the Field. Columbia Univ., NY (ERIC Document Reproduction Service No. ED 326 482)

Kerka, Sandra (1995). Techniques for Authentic Assessment. Washington, D.C. (ERIC Document Reproduction Service No. ED 330 020

McLeese, Heather (1997). The Need for More Authentic Assessment Measures in Schools. Mich.: Western Michigan University.

Perrone, Vito (1991) On Standardized Testing. Urbana, Ill. (ERIC Document Reproducti Service No. ED 338 445)

Terwilliger, James S. (1996). Semantics, Psychometrics, and Assessment Reform:A Close Look at 'Authentic' Tests. New York, NY: National Council on Measurement in Education (ERIC Document Reproduction Service No. ED 397 123)

Wiggins, Grant (1989). Teaching to the (Authentic) Test. Educational Leadership. 46(7), 41-47

Wiggins, Grant (1990). The Case for Authentic Assessment. Washington, D.C.: American Institute for Research. (ERIC Document Reproduction Service No. ED 390 639)