

3

The new learning paradigm brings with it the need for a change in assessment practices as well. In this chapter, one of the assessment practices that is most consistent with this paradigm, authentic assessment, is discussed.

Authentic Assessment: Testing in Reality

Marilla D. Svinicki

Recently a significant change has occurred in the types of instructional strategies that are being used in education in general and higher education in particular. Although the lecture-and-discussion methods are still the dominant mode of instruction, some faculty have begun to heed the advice of Barr and Tagg (1995) to move to a learning paradigm in which the learner becomes an active constructor rather than a passive recipient of knowledge. If those who are making the change are doing so because they have actually adopted the philosophy behind the learning paradigm, they must realize that with it comes the need to adopt a new assessment paradigm as well. If instruction is becoming more learner centered, then assessment should head in that direction as well (Anderson and Speck, 1998).

What Is Authentic Assessment?

One of the most commonly occurring changes in assessment that accompany changes in instruction is the move toward more authentic assessment methods. What does it mean to say that an assessment is authentic? It means that the assessment is based on student activities that replicate real-world performances as closely as possible. Assessment is no longer restricted to paper-and-pencil or even computer-drill-and-practice-type tests. Wiggins (1998) provides six characteristics of an assessment that would qualify it as authentic.

1. The assessment is realistic; it reflects the way the information or skills would be used in the “real world.”

2. The assessment requires judgment and innovation; it is based on solving unstructured problems that could easily have more than one right answer and, as such, requires the learner to make informed choices.
3. The assessment asks the student to “do” the subject, that is, to go through the procedures that are typical to the discipline under study.
4. The assessment is done in situations as similar to the contexts in which the related skills are performed as possible.
5. The assessment requires the student to demonstrate a wide range of skills that are related to the complex problem, including some that involve judgment.
6. The assessment allows for feedback, practice, and second chances to solve the problem being addressed.

Wiggins also gives a comparison between authentic assessments and typical tests. For example, a typical test “must be unknown in advance to ensure validity” whereas an authentic test “is known as much as possible in advance” because it is based on predictable skills and situations. If the assessment is truly authentic, the kinds of skills it invokes would have been practiced many times before in a wide range of situations so that they become predictable as indicators of learning. He also points out that authentic tests are iterative, “containing recurring essential tasks,” whereas typical tests are “one shot.” Typical tests have only correct answers whereas authentic tests are aimed more at the quality of the response and its justification. Typical tests sample from the possible universe of testable content, but authentic tests involve integrated challenges that require the learner to assemble components into the final product. Typical tests infer student understanding based on the correlation between what is tested and what is desired; authentic tests go directly to the desired outcome. Typical tests are more summative in nature, but authentic tests provide diagnostic information and feedback to the student so that they can see where and how to make corrections (Wiggins, 1998).

Wiggins does not mean that there are clear lines or a dichotomous categorization of tests as either authentic or not authentic. Instead, he demonstrates that there is a continuum of authenticity along which assessments might fall. For example, he offers the following three scenarios as a demonstration of such a continuum: An inauthentic measure would be to “write a paper on laws.” A more authentic measure might be to “write a persuasive essay on why a law should be changed.” The difference is the context in which the task is situated and the nature of the task, which is much more realistic because a citizen could conceivably be petitioning for a law to be changed. The most authentic task, however, would be to “write a proposal to present to appropriate legislators to change a current law.” Here there is a realistic context and a behavior that is identical to what might occur in life. The difference between the second two is the audience for the assignment. One is still the professor, whereas the other moves the audience to outside the classroom.

This distinction between intended audiences echoes the criteria that had been laid out by earlier experts and that could be used as guidelines for creating an authentic assessment. Newmann, Secada, and Wehlage (1995) listed the following as components of an authentic assessment:

Construction of knowledge

1. Student organization of information (higher-order skills)
2. Student consideration of alternatives

Disciplined inquiry

3. Core disciplinary content knowledge
4. Core disciplinary processes
5. Written communications to elaborate understanding

Value beyond the school

6. Connecting problems to the world beyond the classroom
7. Involving an audience beyond the school

What Does Authentic Assessment Look Like?

After reading the foregoing descriptions, readers might be saying to themselves that this sounds a lot like what is done in clinical teaching, internships, performing arts, and many other more advanced teaching venues. They would be correct in that observation. There are many fields in which the quality of a student's learning could *only* be assessed in the final performance. Giving a recital or caring for a patient is a classic example of authentic assessment. Can those models be used as the basis for authentic assessment in other fields?

Let us consider some examples of authentic assessments in more traditional settings to clarify what they entail. I have compiled some examples from a range of disciplines, each of which example might be modified for use in other similar settings.

- Snavely and Wright (2003) described their use of a *research portfolio* in an honors thesis course for library students. The portfolio was designed to mimic the information research process used by professionals in the field, including topic definition, search strategy development, database selection, evaluation of resources, an iterative process that involved revision at each step, and teacher and student notes. The portfolio traced the entire process that an informational professional would conduct in the situation of researching a topic for a client. In this case, the students were actually their own clients because the portfolio became the basis for their theses.

- Wellington, Thomas, Powell, and Clarke (2002) described using multidisciplinary teams of students from engineering, marketing, accounting, and industrial design who come together to work on *real-world problems* provided by industrial partners of the institution. The problems involve all the skills that are the target of these programs, including understanding of the problem, design strategies, communications with the client industry,

and participation in multidisciplinary teams. The students' work is evaluated by both the instructors and the industry partners.

- Cleveland (2004) teaches a class in management information systems in the University of Texas at Austin McCombs School of Business, a class that enrolls more than eleven hundred students each semester. The purpose of the course is to learn beginning business concepts such as statistics and information management, but applying them to the realities of business can be a challenge. To solve this challenge, Cleveland sets her class up as a *simulated environment*—in this case, a town with groups of students organized as businesses who must develop a business or product idea, go through the process of researching its feasibility in the “town” given its demographics, apply for any pertinent patents or trademarks, and finally, present their business plan at a major business fair at the end of the semester. The groups have to essentially market their business idea to a large group of invited judges representing the town officials and residents. In reality, of course, the judges are invited experts from throughout the university and private sector companies that are sponsors of the business school. Because students in the class come from a range of backgrounds and with a range of skills, they must learn to tap the talents of every member of the group and make the hard decisions that a real business would have to make. The accompanying business plan notebook serves as a record of all that they have done to lead up to the final booth display at the fair.

- Students at the American Film Institute work in teams to create *real products*, in this case, a series of short films, that are then screened by all the faculty and students as well as visiting dignitaries associated with the film industry. Students are drawn from all the components of the institute and represent screen writing, directing, cinematography, set design, and sound. Each student makes a unique contribution to the production, and all must work together to produce the final product.

The foregoing examples are fairly elaborate, but there are much simpler ways of using authentic assessment in classes. For example, most service-learning programs would be considered as candidates for authentic assessment, depending on the type of outcomes the program seeks and the kinds of formative and summative output the students are required to produce. They certainly take place in real environments with real problems, and most use something like a portfolio assessment strategy.

Another example of commonly used evaluation formats ripe for authentic assessment is the research paper. If the paper itself is written for a real purpose other than for the instructor as audience, it approaches authentic assessment. For example, in a sociology class, students might construct a family history based on interviews of their relatives and produce a portfolio that includes old and new photos, interview transcripts, a family tree, and so on. This could even be expanded into a family Web site that includes video clips of family members. The consumers of this product, in addition to the

instructor, would be the family itself. The research paper in another format has become the portfolio in which the students display not only the results of their efforts but also the process by which they were achieved and the self-assessment of the success of that process.

In many undergraduate science or science education courses, instructors are beginning to assign projects that involve the production of materials, Web-based or in-class activities, for elementary school use or for dissemination to the public. The college students do the research work to identify an appropriate content, simplify and clarify it for the target age group (the characteristics of which they have researched as well), create the storyboard or the support materials, and package the whole for delivery to local elementary schools or other public venues such as museums. In some cases, the college students even teach their sample lesson to the elementary students or create Web sites that are accessible to the public.

Pluses and Minuses of Authentic Assessment

Much may be said in support of authentic assessment from a technical measurement perspective. Because authentic assessments track the real world so closely, they are likely to have a great deal of face validity both for students and for any outside evaluator. They can also be motivating for students if the students are allowed to choose an assessment type that fits their own interests and skills. Certainly the students will be able to see the value of what they are learning if the product they produce then becomes a reality in public policy or the real needs of people. From a technical psychological perspective, authentic assessments are likely to produce a great deal of transfer from the classroom to the real world after graduation. And having a concrete target at which to aim allows the students to assess their own progress more readily during learning.

There are, of course, almost as many drawbacks to authentic assessment. As with all innovations associated with the new learning paradigm, authentic assessment requires a lot of time and effort on the part of both the students and the instructor. There is probably no real solution to that increased time commitment; it is simply a trade-off for the quality of learning, motivation, and transfer that result from this type of activity.

From a measurement perspective, there are reliability-in-grading issues. Can different works from different students be graded consistently and comparatively? Fortunately, because this type of assessment is becoming more popular and because it echoes the kinds of assessment done historically in clinical or performance venues, much is known about how to improve the quality of evaluation of the products (Palomba and Banta, 1999; Walvoord and Anderson, 1998). A more difficult technical problem is that because no two assignments are identical within a class, it is difficult to make cross-student comparisons. On the other hand, if one is to be consistent with the learning paradigm, you would not be making cross-student comparisons

anyway; you would only be comparing a student's efforts with his or her own previous efforts.

Authentic assessment also raises some interesting issues all its own. There are safety issues: When students work in real-world settings, are we putting them or those they work with at increased risk? There are management issues: How much supervision is required, and who should provide it? Who owns the product once it has been finished, the student, the school, or the applied setting? How should it be archived and by whom? The most common form of documentation of such activities are portfolios, which can run into large manuscripts and accompanying materials. There are ethical issues: Is everyone being treated fairly and receiving equal benefit from such assignments, both the students and the individuals they work with? Can students observe the level of professionalism that ensures the privacy and ethical treatment of others? There are student issues: Students who have been highly successful all their student days are now being evaluated on a whole different set of skills and attitudes. Are they ready? Are we ready?

Is It Worth It?

The benefits of using authentic assessments are many, and it is certainly more consistent philosophically with the learning paradigm toward which we are moving in higher education. The pragmatics of designing and implementing these assessments may only be temporary impediments. More and more elementary and secondary schools are moving toward some degree of authentic assessment, however small, despite the press of high-stakes performance testing that is pushing in the other direction. Our students may be coming to expect more because of the real experiences they have already had. Certainly, more is being demanded of us from the public that employs our graduates.

Authentic assessment has the promise to meet this last challenge more than anything that has thus far happened in higher education. What we can measure often becomes what is valued; therefore, we should target our evaluation at what we want valued. It seems far more consistent with our values to evaluate our students in authentic ways so that the real-world skills and attitudes they need become the things we all focus on during their education.

References

- Anderson, R. S., and Speck, B. W. (eds.). *Changing the Way We Grade Student Performance: Classroom Assessment and the New Learning Paradigm*. New Directions for Teaching and Learning, no. 74. San Francisco: Jossey-Bass, 1998.
- Barr, R., and Tagg, J. "From Teaching to Learning: A New Paradigm for Undergraduate Education." *Change*, 1995, 27(6), 12-25.
- Cleveland, L. "Cleveland's Good for Business." University of Texas at Austin Web site: <http://www.utexas.edu/features/archive/2004/cleveland.html>. Accessed November 14, 2004.

- Newmann, F., Secada, W., and Wehlage, G. *A Guide to Authentic Instruction and Assessment: Vision, Standards, and Scoring*. Madison, Wis.: Wisconsin Center for Education Research, 1995.
- Palomba, C., and Banta, T. *Assessment Essentials: Planning, Implementing, and Improving Assessment in Higher Education*. San Francisco: Jossey-Bass, 1999.
- Snively, L. L., and Wright, C. A. "Research Portfolio Use in Undergraduate Honors Education: Assessment Tool and Model for Future Work." *Journal of Academic Librarianship*, 2003, 29(5), 298–303.
- Walvoord, B., and Anderson, V. *Effective Grading: A Tool for Learning and Assessment*. San Francisco: Jossey-Bass, 1998.
- Wellington, P., Thomas, I., Powell, I., and Clarke, B. "Authentic Assessment Applied to Engineering and Business Undergraduate Consulting Teams." *International Journal of Engineering Education*, 2002, 18(2), 168–179.
- Wiggins, G. *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*. San Francisco: Jossey-Bass, 1998.

MARILLA D. SVINICKI is associate professor of educational psychology and former director of the Center for Teaching Effectiveness at the University of Texas at Austin.