

Running Head: The State of Authentic Assessment

The State of Authentic Assessment

by

Shelley Keyser

(Master's Student—Department of Instructional Psychology and Technology)

and

Scott L. Howell, PhD

(Adjunct Associate Professor and Director of Evening Classes)

Brigham Young University

December 12, 2008

Abstract

The purpose of this review article is to briefly summarize the state of authentic assessment as it evolves theoretically and then integrates itself into educational practice. Some of the questions it seeks to answer from the literature include: how did it begin; what were the circumstances surrounding its inception; how does it relate to performance and alternative assessment; what are obstacles to wider adoption; what impact is technology having in its emergence; and what should happen next? The article identifies authors and cites studies that have informed the discussion on authentic assessment. It is evident from the review that authentic assessment is an important educational concept that at the time of this writing is in a fledgling state of definition and integration.

The State of Authentic Assessment

Introduction

Authentic assessment is still an ambiguous concept to educators—some refer to it as a specific assessment that reflects a real-world context while others describe it as an assessment aligned to real-world activities or some combination thereof. While a survey of the most current literature shows some convergence in meaning, it is still evident that this assessment strategy still seeks definition.

In 1988 the term *authenticity* was used to describe a type of achievement, and later, any assessments related to that achievement (Archbald and Newman). In the late 1980s and early 1990s the U.S. Department of Education determined to analyze assessment of student performance as part of its educational reform (Kane et al. 1995), and researchers began to ask more questions about authenticity—what it meant and how it should be applied.

Since these educational reforms, authentic assessment has frequently been identified with another form of assessment called *performance assessment*. Today, educators still seek confluence on the meaning of authentic assessment, on how it is related to other forms of assessment, and on how it can be implemented into classrooms.

The purpose of this article is to provide the latest perspective on the meaning of authentic assessment and its integration among educators through a literature survey. An understanding of what authentic assessment is, where it came from, how it has been implemented so far, and where it is going in the future will not only inform the discussion but also encourage its practice.

Defining Authentic Assessment

The term *authentic* is something real or genuine. When education prepares students for real or genuine tasks they may encounter in a real-world setting it is considered *authentic*—or at least

more authentic. The term's ambiguity originates from the actuality that tasks inside the educational setting have varying degrees of replication outside the educational setting. This dilemma gives rise to many questions associated with authentic assessment, including the ultimate determination of whether the form of assessment is authentic, mostly authentic, or not.

Early uses of the term *authentic* in education referred to any use of instruction and assessment that mirrored real-world tasks rather than a contrived learning experience of the classroom-based worksheet or test. In that context, authentic achievement was an achievement that bore some resemblance to that required to be successful in the real world and not the staid classroom environment. Furthermore, authentic assessment became assessment used to assess real-world experience and achievement (Archbald and Newmann 1988).

When considered individually, each assessment represents some degree of the real world no matter how small. With this in mind, it is not enough to define an assessment as just categorically authentic or not. The assessment must be analyzed to determine the presence or degree of authenticity, and the elements of that authenticity, to learning outcomes aligned with real-world tasks.

Some researchers have classified authentic assessments and their properties into more descriptive sub-groups. Cumming et al. (1999) classified authentic assessments as *performance*, *context*, *complexity*, or *competence*, and then associated the first three with relevant learning theories. The first authentic assessment dynamic mentioned by Cumming is *performance*, who emphasized that assessment tasks separated from their real world context will decrease the integrity of assessment results. The second authentic assessment dynamic is *context*, and it posits that students have minimal transfer ability from the classroom to real-world contexts; it was developed in response to situated learning theories. Its advocates explain that performance in one

context is not indicative of performance in another, thus, it is important to teach and assess in real-world contexts. The third authentic assessment dynamic is *complexity*. It suggests that students are better prepared to develop and later use problem-solving skills through learning and assessment opportunities inherent in complex scenarios that mirror an authentic setting. Fourth, Cummings introduced the dynamic *competence* to include those activities that retain differences when transfer occurs from the laboratory or classroom setting to the real world. For example, performing carpentry tasks in a well-equipped classroom would require different skills, and solicit different results, than doing carpentry in a real environment with less than state-of-the-art tools.

Rule's (2006) literature review of examples of authentic assessment in higher education concluded that there are four characteristics of authentic activities:

- 1) involve real-world problems that mimic the work of professionals;
- 2) include open-ended inquiry, thinking skills, and metacognition;
- 3) engage students in discourse and social learning; and
- 4) empower students through choice to direct their own learning.

These characteristics not only help recognize an authentic assessment but also help provide theoretical constructs to describe significant elements or properties of authentic assessment.

In their research, Gulikers et al. (2004) recognized that assessments inevitably have varying levels of authenticity. They observed differences between a students' perspective of authenticity to that of an educator, raised questions of reliability and validity, and also conducted studies to discover the more important elements of authenticity. From their findings, they presented a continuum framework based on student needs and learning outcomes. The framework they developed seeks to help designers determine which level of authenticity most suits the learning

and assessment situation. The five dimensions of Gulikers framework are: (a) the assessment task, (b) the physical context, (c) the social context, (d) the assessment result or form, and (e) the assessment criteria.

While authentic assessments do not exist outside of on-the-job or in-the-real-world setting, some assessments are more authentic than others. Typically the more authentic assessments better mirror reality than those that approximate it. Much of the current literature and extant research still seeks to define authentic assessments and create a more widely accepted theoretical construct from which clearer communication and better research can take place.

Historical Background and Context

While a desire to see authenticity in assessment is not new, the term *authentic* first appears in reference to educational tasks and achievements and not necessarily assessments (Archbald and Newmann 1988). Wiggins (1989) suggests that authentic assessment should be associated with authentic achievement.

In the 1980s (and thereafter) the nation looked for an alternative to standardized assessments, and many states committed themselves to increased performance assessments as part of the educational reform movement. California began with open mathematics assessments (e.g., open-ended questions and short investigations) and Vermont followed by implementing portfolios. However, concerns with the costs of developing and administering performance assessments and the reliability and validity of results frustrated widespread adoption as a nationwide standard (Kane et al. 1995).

As with authentic assessment, researchers still seek to define performance assessment. Currently, the two terms—*authentic* and *performance assessment*—are associated together, and sometimes even with a third term, *alternative assessments*. Authenticity in an assessment

typically references those elements that most closely align with real-world tasks. *Performance assessment* emphasizes competency manifest through real-world actions; *alternative assessment* emphasizes an assessment other than traditional multiple-choice assessments . . . (Miller et al 2008).

Marzano et al. (1993) response like so many to educational reforms in the 1980s was emphasizing performance assessment and trying to find ways to more fully integrate it into the educational experience. The research team identified three reasons why educational assessment required reform through the use of performance assessments:

- changing nature of educational goals;
- relationship between assessment and teaching and learning; and
- limitations of the current methods of recording performance and reporting credit.

He then outlined five dimensions of learning as part of an instructional and assessment framework to help meet reform demands. Marzano (et al.) also included a discussion of rubrics as an important part of the performance assessment experience.

Other researchers (e.g., Haertal 1999) stressed the importance of focusing on performance assessment as a classroom tool and not necessarily as an answer to educational reform. He worried that the widespread adoption of authentic assessments could also create another problem—another type of test for teachers to teach to! While Wiggins (1990) generally agreed with these concerns he also felt that a move toward authentic tasks in classroom assessment would improve teaching and learning effectiveness. So it is, everyone agrees that authentic assessment is good—as long as they are talking about the same thing—though concerns about feasibility (Wiggins 1990) and what it is that actually defines and constitutes authenticity (Gulikers et al. 2004) linger.

Authentic Assessment in a Larger Context

As a result of recent educational movements and new technologies, teachers and school districts are looking for wider applications and better methods for integrating authentic assessment into their instructional plan. Boyd-Batstone (2004) developed a systematic approach for using standards-based, authentic assessment in the reading context. Among other things he emphasizes the importance of working from specific content standards whenever preparing authentic assessments. He also encourages the careful recording of student activity by the teacher in the authentic assessment setting. Hopefully, other disciplines will establish a systemic approach to developing, and then applying, authentic assessments to their field.

Once a teacher has determined which content standards they want to assess they must define what student mastery is, and what it looks like when it occurs. After the teacher has developed and administered an authentic assessment the results are submitted to interpretation that must remain as reliable and valid as the results themselves (Boyd-Batstone 2004).

The grading rubric used as part of an authentic assessment is an important—even a critical—part of the assessment experience. Andrade (2000) explained several ways rubrics could be used for not only authentic assessment but also authentic learning. Rubrics help communicate teachers' expectations to parents and students and help students direct their learning experience and preparation for any authentic assessments. Good construction of a rubric includes specifying which constructs to assess and what performances are required to demonstrate mastery.

A survey of recent advances in authentic assessment would be incomplete without a word on technology and how it has enabled authentic assessment. While simulations that replicate a real-world atmosphere are usually costly the costs will only decrease as technology advances and cost efficiencies are realized. One example of a simulation project that includes authentic assessment

is the Quest Atlantis project (Barab et al., 2005). Although more elaborate and more far-reaching, Quest Atlantis is similar to what some teachers are doing using video games, simulations, and technology. In an effort to give secondary students a more authentic social and scientific experience, Quest Atlantis forfeits environmental authenticity for authenticity in factual knowledge, conceptual knowledge, and decision-making. In this virtual game students interact with people in an imaginary world who seek to scientifically discover Earth. The students are given real-world assignments, like carrying out a science project within the community. Using games, simulations, stories, and other techniques, educators are able to introduce more authenticity into their assessments and more closely align those assessments with intended learning outcomes.

Conclusion

In the wake of educational reform, many states have mandated statewide performance assessment. In response to the educational movements, K-12 educators have identified some solutions to problems of efficiency and reliability that include making rubrics and aligning assessments to learning outcomes. These states have also published some of their successes in using performance assessments all in an effort to disseminate more authentic assessment practices.

Melding theoretical models that help define and describe authentic assessments with scoring rubrics that then align learning outcomes to those assessments promise not only more meaningful assessments but also better learning by assessment. Advances in technology will also allow educators greater access to authentic experiences and thereby enable more authentic assessments for their students.

The field needs not only more definition but also more applied research focusing on successful interventions and practices. Also a more thorough understanding and analysis of the interrelationship of performance and authentic assessments will help research bases converge so that theoretical constructs and applications will advance.

Authenticity is the element of every successful assessment that resembles a real-world skill or activity and aligns itself with a learning outcome. Since educational reforms in the late 1980s, educators and researchers alike have shown increasing interest in defining and applying authentic assessment. These emerging theoretical and applied models promise to inform future research and best practices.

References

- Andrade, H.G. (2000). Using Rubrics to Promote Thinking and Learning. *Educational Leadership*. 57(5) obtained from <http://www.ascd.org/readingroom/edlead/0002/andrade.html> on August 4, 2008.
- Archbald, D. A. & Newmann, F. M. (1988) *Assessing authentic academic achievement in the secondary school* (Reston, Virginia, National Association of Secondary School Principals).
- Barab, S., Thomas, M., Dodge, T., Carteaux, R., & Tuzun, H. (2005). Making Learning Fun: Quest Atlantis, a Game Without Guns. *Educational Technology Research and Development*. 53(1) 86-107.
- Boyd-Batstone, P. (2004). Focused anecdotal records assessment: A tool for standards-based, authentic assessment. *International Reading Association*, 58(3), 230–240.
- Burley, H., & Price, M. (2003). What Works with Authentic Assessment. *Educational Horizons*, 81(4), 193–196.
- Cumming, J.J., & Maxwell, G.S. (1999) *Contextualizing Authentic Assessment*. *Assessment in Education: Principles, Policies, and Practices*. 6(2) 177–194.
- Gulikers, J.M., Bastiaens, T.J., & Kirshner, P.A. (2004). A Five-Dimensional Framework for Authentic Assessment. *Educational Technology Research and Development*, 52(3), 67–87.
- Buzzetto-More, N.A., & Alade, A.J. (2006). Best Practices in E-Assessment. *Journal of Information Technology Education*. 5, 251–269.
- Haertal, E.H. (1999) Performance Assessment and Educational Reform. *Phi Delta Kappan*. 80(9) 662–669.

- Herrington, J., & Kervin, L. (2007). Authentic Learning Supported by Technology: 10 suggestions and cases of integration in classrooms. *Educational Media International*, 44(3), 219–247.
- Herrington, J., Reeves, T.C., Oliver, R., & Woo, Y. (2004). Designing Authentic Activities in Web-based Courses. *Journal of Computing in Higher Education*, 16(1), 3–29.
- Kane, M.B., & Mitchell, R. Ed. (1996) *Implementing Performance Assessment: Promises, Problems, and Challenges*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Khattari, N., & Sweet, D. (1996). Assessment Reform: Promises and Challenges. *Implementing Performance Assessment: Promises, Problems, and Challenges*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Marzano, R.J., Pickering, D., & MicTighe, J. (1993). *Assessing Student Outcomes: Performance Assessment Using the Dimensions of Learning Model*. Virginia, U.S.: Association for Supervision and Curriculum Development.
- Miller, M.D., Linn, R.L., & Gronlund, N.E. (2008). *Measurement and Assessment in Teaching*. New Jersey, U.S.: Merrill.
- Reeves, T.C., Herrington, J., & Oliver, R. (2002). Authentic activities and online learning. *HERDSA*, 562–268.
- Rule, A. C. (2006). The components of authentic learning. *Journal of Authentic Learning*, 3(1), 1–10.
- Williamson, D.M., Bauer, M., Steinberg, L.S., Mislevy, R.J., Behrens, J.T., & DeMark, S.F. (2004). Design Rationale for a Complex Performance Assessment. *International Journal of Testing*. 4(4), 303–332.

Wiggins, G. (1989). A true test: toward more authentic and equitable assessment, *Phi Delta Kappan*, 70 703–713.

Wiggins, G. (1990). The case for authentic assessment. ERIC Digest. (ERIC Document
Reproduction Service No. ED 328 611).