Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Volume 17, Number 2, January 2012 ISSN 1531-7714

Defining Authentic Classroom Assessment

Bruce B. Frey, *University of Kansas*
Vicki L. Schmitt, *Logan-Rogersville School District, Rogersville, MO*
Justin P. Allen, *University of Kansas*

A commonly advocated best practice for classroom assessment is to make the assessments authentic. Authentic is often used as meaning the mirroring of real-world tasks or expectations. There is no consensus, however, in the actual definition of the term or the characteristics of an authentic classroom assessment. Sometimes, the realistic component is not even an element of a researcher's or practitioner's meaning. This study presents a conceptual analysis of authentic as it is used in educational research and training to describe an approach to classroom assessment. Nine distinct components or dimensions of authenticity are identified and only one of those is the realistic nature of the assessment.

A well accepted position among educational researchers and teacher educators is that the *best* classroom assessments are *authentic* (e.g. Archbald & Newman, 1988; Bergen, 1993; Gronlund, 2003; Meyer, 1992; Newman, Brandt & Wiggins, 1998; Wiggins, 1989a, 1989b). The term *best* typically means *valid,* and *authentic* is usually defined as having something to do with the *real world.* This position is difficult to translate into an assessment strategy, however, for two reasons. First, validity is not a characteristic of any assessment; it refers to the interpretation and use of assessment results. Secondly, there are a variety of definitions of authenticity presented in the research literature and in books and other materials used to train teachers. While most authors speak of authentic in the context of application outside the classroom, some do not and emphasize other aspects of assessments that determine their authenticity. Many advocates emphasize the role of the student in the process or the complexity of the task. Others present criteria that sound suspiciously like general recommendations for valid classroom assessment of any type or, sometimes a bit more specifically, valid *performance-based* assessments of any type. Such recommendations offer little in determining whether any specific teacher-made assessment is authentic and, therefore, produces the benefits presumably associated with authenticity.

Even a cursory examination of the existing literature reveals that there is not always agreement as to the important elements that make an assessment authentic. The "real world" element is often, though not always, emphasized, but there are a variety of other components cited, as well. For example, Bergen (1993) identifies three qualities of authentic assessment. Referring to assessment that is both performance and authentic, one criterion provided is that it is often group-based with each individual contribution required for success. The other two qualities refer to the complexity of the task- it measures many facets simultaneously and it is applied in a way that reflects the complex roles of the real world. The first criterion suggests that authentic assessment should involve a group project and a group evaluation, a characteristic that is rarely suggested as crucial by others and not likely to be part of any large scale standardized "authentic" assessment. Further, writing assessments, to pick just one example, are often cited as authentic under certain circumstances, but few writing assignments are group assignments. In fact, it is often the

individualized nature of writing that supports claims of authenticity.

Certainly, however, the *realistic* criterion is commonly presented as the underlying, critical defining factor for authenticity. For example, a discussion of the varied definitions of the terms performance assessment and authentic assessment is presented by Palm (2008). Beginning with a basic dictionary definition of authentic as meaning essentially that something is "real, true or what people say it is" (p. 6), he suggests that the term is used in various contexts as being *true* or *real*. Palm concludes that authenticity is defined as assessment that is real in terms of processes and products, assessment conditions or the presented context, and true to life beyond school, curriculum and classroom practice or learning and instruction.

This study conducted a conceptual analysis of *authentic* as it is used to describe a type of classroom or educational assessment. We gathered and reviewed journal articles, presentations, books and dissertations to identify concrete criteria for evaluating the authenticity of an assessment. We first developed a conceptual "map" of the term and at the conclusion of our review produced a table of nine dimensions of authenticity related to the context of the assessment, the student's role and scoring procedures.

Brief History of Authentic Assessment

The earliest reference to *authentic* tests is likely that made by Archbald and Newman in 1988, in a book critical of standardized testing, that sought to promote assessment centered on meaningful real-world problems or tasks. Assessment is authentic when it measures products or performances that "have meaning or value beyond success in school" (Newman, Brandt & Wiggins, 1998, p.19). According to Newman, assessments that ask questions and poses problems that have "real world" meaning to students meet one criterion for being authentic intellectual work, but there are two others related to disciplined inquiry that are unrelated to the realism of the assessment tasks. Wiggins was also an early proponent for the use of the term *authentic* to describe assessment with real-world application (1989). "'Authentic' refers to the situational or contextual realism of the proposed tasks" he has emphasized (Newman, Brandt & Wiggins, 1998, p.20). Terwilliger (1998) expressed concerns with Wiggins and others use of the term, viewing the label of *authentic* as a veiled criticism of traditional assessment approaches as somehow less authentic or inauthentic. Wiggins position is essentially that traditional assessment is not inauthentic, it is simply less direct and, probably, less meaningful to students. Wiggins argues that traditional assessment is not faithful to the domains of performances and contexts that are most important for higher order thinking and learning (1993). As he used the term, authenticity is akin to *fidelity*.

Since the early 90's, teacher educators, theorists and researchers have flocked to support authentic assessment as a more valid and productive approach towards student evaluation. There are at least a dozen books and hundreds of journal articles on authentic assessment as an approach. Some of these works, even the books that use *authentic assessment* in their titles, use the term without offering a direct definition, but most do, at least, offer a set of criteria that amount to a definition. A review of those publications reveals a wide range of descriptions for the term, some of which overlap with other classroom assessment terms, such as *performance-based assessment* and *formative assessment*- other modern approaches that have broad support as alternatives to "traditional" paper-and-pencil testing or standardized large scale assessment.

In education, of course, it is not uncommon for best practices or "hot" or innovative topics or methods to suffer from a confusion of understanding and a lack of consistent use of terms or definitions. The conceptual overlap between performance assessment, formative assessment and authentic assessment clouds the waters if one wishes to provide objective criteria (or, at least, criteria approaching objectivity) for judging the authenticity of an assessment. While exploring this dilemma, for example, Frey & Schmitt (2006) compared competing views as to comparability in meaning between performance assessment and authentic assessment. Figure 1 reflects that disagreement in the field.

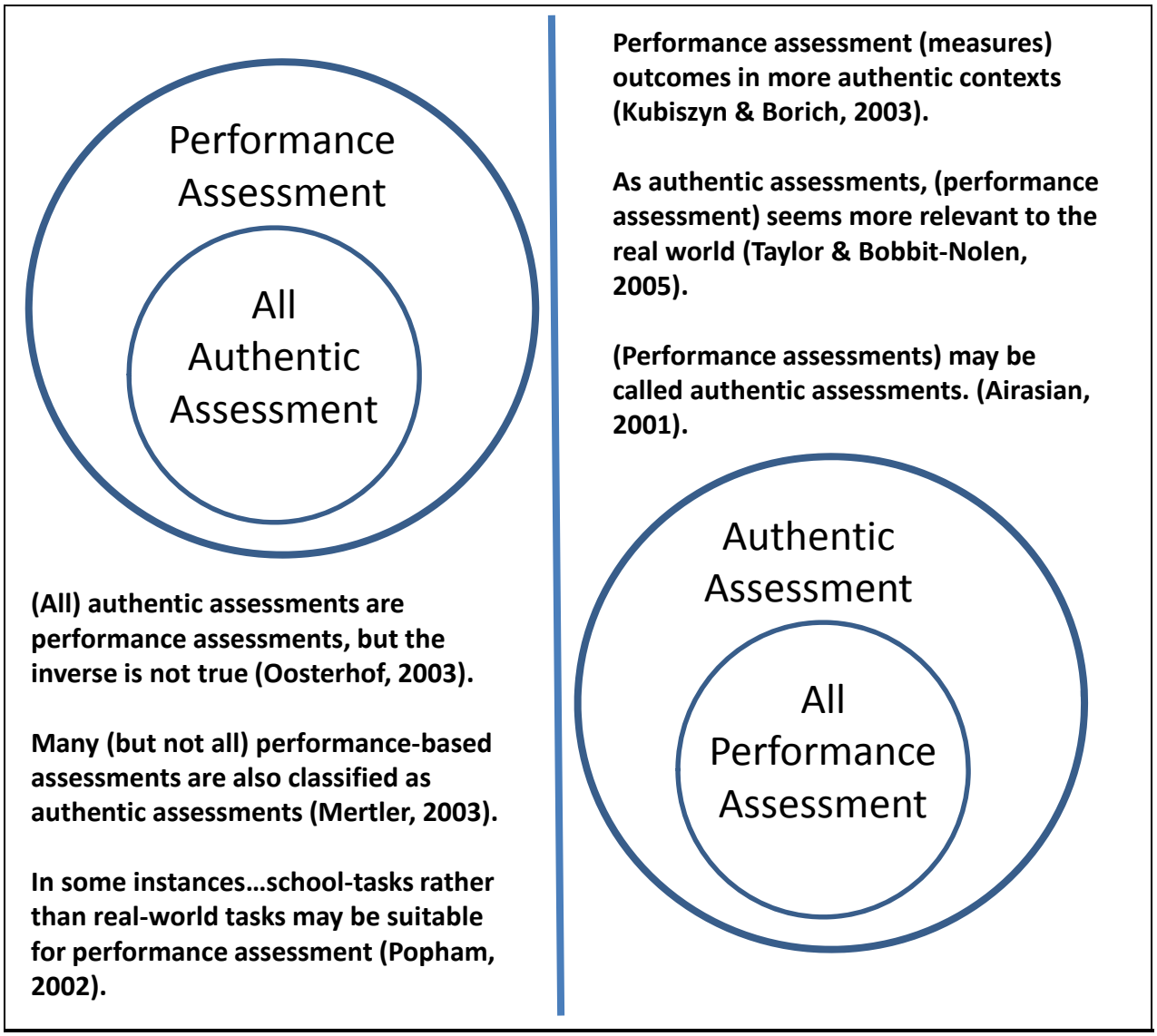

Figure 1: Two Opposing Possibilities for the Conceptual Relationship between Performance and Authentic Assessment (adapted from Frey & Schmitt, 2007)

Methods

For this study, we conducted a review of the literature to develop a comprehensive list of critical components that various authors and researchers believe determine the authentic nature of any classroom assessment. The methods used were similar to those utilized by others to identify guidelines for quality assessment in the absence of empirical studies (Frey, Petersen, Edwards, Pedrotti & Peyton, 2005; Frey & Schmidt, 2006; Haladyna & Downing, 1989; Haladyna, Downing & Rodriguez, 2002). That is, the consensus of experts, theorists, researchers, advocates and trainers was used in an attempt to identify the central components of authenticity. We examined the nature of authenticity in the context of what valid authentic assessments should look like for children in grades K-12 (e.g. Darling-Hammond, Ancess & Falk, 1995; Paris & Ayres, 1994; Wiggins, 1989a, 1989b, 1990, 1992,1993), adults engaged in professional development (i.e. college and graduate students; e.g. Gulikers, Bastiaens & Kirschner, 2004; Gulikers,

Bastiaens & Martens, 2005; Herrington & Herrington, 1998; Rennert-Ariev, 2005, Svinicki, 2004), pre-school children (e.g. Grisham-Brown, Hallam, & Brookshire, 2006; Puckett & Black, 2008) and students in a variety of specific subject areas (e.g. Bachman, 2000; Ekbatani & Pierson, 2000; Hirvela & Pierson, 2000; Montgomery, 2000; Valencia, Hiebert & Afflerbach, 1994).

We located 109 scholarly publications (mostly journal articles and books), by 100 different primary authors, that provided a definition of authentic assessment specific enough to identify characteristics of assessment that make them authentic. This was a subset of a much larger pool of scholarly publications that use the term without definition or only imply a definition. We found scholarly articles and books offering criteria for authenticity for a variety of different types of students and content areas and separated them by intended population for the purpose of analysis into pre-school (three publications), school-aged children K-12 (forty-seven publications), English as a second language and a foreign language (six publications), language arts (seven publications), foreign language (two publications), social studies (four publications), science (four publications), math (five publications), physical education (five publications), fine arts (two publications) and college and professional training (twenty-six publications).

The goal was to analyze the abstract concept of authenticity through the descriptions and discussions provided by experts, authors and researchers. In a few cases, statistical analyses suggested components or factors of authenticity, but most of the scholarship examined amounted to essays, thought pieces, informational articles and the definitional sections of studies involving authentic assessments as a variable or an intervention. The outcome of this process was a concept analysis, a model of the parts and pieces or types of authenticity that are important to those whose work we studied.

Identifying Characteristics of Authentic Assessment

There was some subjectivity and judgment required to identify separate components or elements in the definitions we found in the various publications. The first two authors and a colleague analyzed articles separately and discussed any difficulties in interpretation or classification with the rest of the research team. Careful attention was paid to not identify characteristics of *quality* or *effective* authentic assessment as required components of a definition. The original text phrases from the publications for each definitional component were sorted into broad categories and tentative labels or names for the categories were determined.

Two examples of the process of identifying different definitional phrases from the publications and the categories into which they were ultimately placed are illustrative. For instance, these definitional elements:

> "… results in a product or presentation that has meaning or value beyond success in school." (Wiggins, 2006, p. 51).
>
> "…emphasize(s) connections between assessment, learning and real-world issues." (Green, 1998, p. 11).
>
> "Performance is assessed in a context more like that encountered in real life…" (Dez, Moon & Meyer, 1992, p.38-39).
>
> were all classified as *realistic activity or context*.

As another example, these definitional phrases:

> "…emphasis she places on student self-assessment." (Wilson, 1993, p. 8).
>
> "The (assessment) would also serve well for a student's self-evaluation…" (McMann & McMann, 1992, p. 6).
>
> "The significant criterion for the authenticity of a writing assessment might be that the locus of control rests with the student…" (Dez, Moon & Meyer, 1992, p.39).

were all classified as *formative assessment*. Formative assessment occurs during instruction, typically does not affect student grades, and often involves students in self-evaluation of their learning progress.

Early in the review of literature, a conceptual, graphic map began to emerge as key definitional components of authentic assessment were drawn

from each study, paper, book or book chapter. This tentative and preliminary organizational scheme allowed for the literal phrasing (or very close paraphrasing) used in various works to be placed in columns below a small set of classifications or categories. Figure 2 presents this initial attempt to align the wording used by each author (or group of authors) based on similarity and to develop labels for the emerging elements. These first iteration categories were used to group the literature in subsequent analyses.

As additional definitions were catalogued and added to the database, the initial categories were revised slightly to allow for inclusion of all definitions. The goal was to use as few categories as necessary, while still maintaining conceptual clarity and precision. For example, the use of multiple scores and the use of portfolio assessment were initially treated as two distinct dimensions of authenticity; later, it became clear that it was most often the multiple indicators aspect of portfolio assessment that appeared to the authors as *authentic* and the categories were combined.

After a set of nine possible components of authenticity was finalized, we performed a frequency count of publications supporting the importance of each element. Percentages were calculated to show the relative "popularity" of each characteristic.

Results

Though it is the commonly assumed definition of authenticity, the requirement that the assessment be "realistic", was often not even mentioned as a necessary characteristic for an assessment to be authentic. Additionally, beyond realism, there were eight other characteristics frequently reported. Many of these seem unrelated to the realism criterion, but instead are of general importance for the reliability or validity of any assessment. The commonly reported dimensions of authenticity, grouped into three broad categories, are:

- the context of the assessment
  - realistic activity or context
  - the task is performance-based.
  - the task is cognitively complex.
- the role of the student
  - a defense of the answer or product is required.
  - the assessment is formative.
  - students collaborate with each other or with the teacher.
- the scoring
  - the scoring criteria are known or student developed.
  - multiple indicators or portfolios are used for scoring.
  - the performance expectation is *mastery*.

Table 1 summarizes the publications we reviewed that referred to authentic assessment in the context of K-12 education. Each article or book is listed and if one of the identified dimensions of authenticity is a required part of the definition as presented in that publication, then the cell associated with that dimension is shaded. Table 2 provides the same information for publications defining authenticity for assessments designed for adults at the college or job-training level. Table 3 summarizes publications that define authenticity specifically for certain academic subjects or for the pre-school level. For Tables 1 and 2, we provide percentages indicating the relative frequency with which each component of authenticity is included by the various authors. For example, from Table 1, fifteen percent of publications that define authentic assessment for school-aged children indicate that one characteristic of authentic assessment is that students must provide a defense of their work. There are only a few publications listed in each category on Table 3, so we have not provided percentages indicating relative frequency on that table.

Clearly, the degree to which an assessment task mirrors some reality outside of the classroom is a critical dimension of authenticity. It is emphasized by many of the authors in our study. However, it is

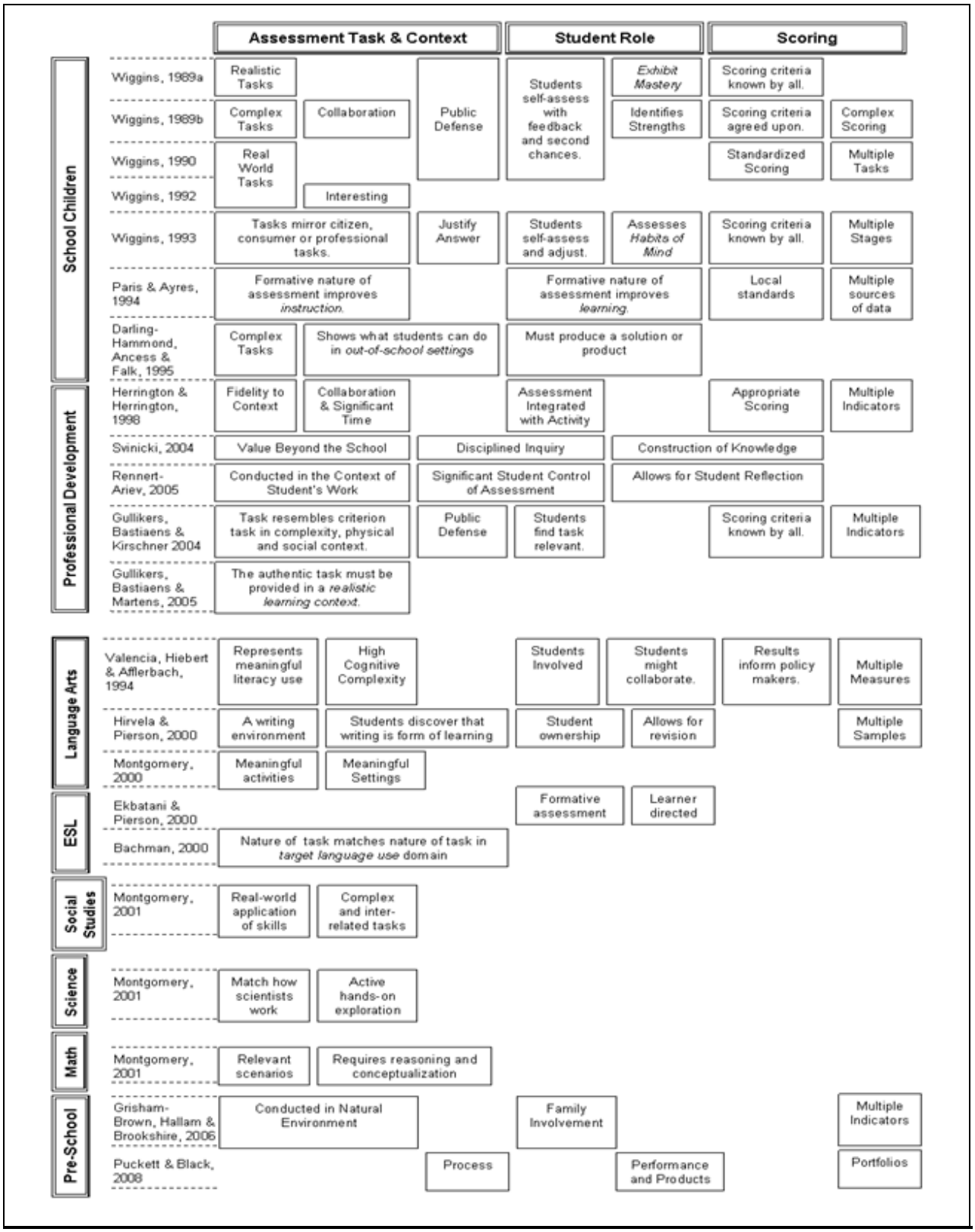

Figure 2. Initial pilot identification of elements of authentic assessment

Table 1. Definitions of authentic assessment for school-aged children (K-12)*

		Context			Student Role			Scoring		
		realistic activity or context	performance-based	cognitively complex	defense required	formative assessment	collaborative	known or student developed criteria	multiple indicators or portfolios	mastery expectation
	Percentage including Element	60%	23%	30%	15%	31%	20%	47%	53%	13%
1	Abernethie, 2006	■		■			■	■	■	■
2	Archbald, 1991	■	■				■	■	■	■
3	Bergen, 1993	■					■			
4	Biondi, 2001	■					■	■	■	
5	Borowski, Thompson & Zaccaria, 2001								■	
6	Brandt, 1996	■	■					■		
7	Bullens, 2002			■					■	
8	Burley & Price, 2003	■							■	
9	Cronin, 1993	■								
10	Cumming & Maxwell, 1999	■		■						
11	Darling-Hammond, 1994							■		■
12	Darling-Hammond, Ancess & Falk, 1995	■		■	■					
13	Dez, Moon & Meyer, 1992	■	■	■		■				
14	Dutt-Doner, & Maddox, 1998					■			■	
15	Engel, Pulley & Rybinski, 2003	■				■			■	
16	French, 2003	■		■					■	
17	Grabinger & Dunlap, 1995	■	■						■	
18	Green, 1998	■				■		■		
19	Herrington & Oliver, 2000	■				■	■			
20	Hunter, 2001	■							■	
21	Jolly & Kettler, 2000	■		■			■	■		
22	Kellaghan & Madaus, 1993		■						■	
23	Kieffer & Morrison, 1994								■	
24	Lawton, 2000				■	■		■		
25	Maden & Taylor, 2001	■			■				■	
26	Meisels, 1996		■			■			■	
27	Meisels, 2001								■	
28	Meyer, 1992	■	■			■		■		
29	Moorcroft, Desmarais, Hogan & Bekowitz, 2000		■					■	■	
30	Mueller, 2005			■				■	■	
31	Newmann, Brandt & Wiggins, 1998			■				■		
32	Paris & Ayres, 1994			■		■		■	■	
33	Ratcliff, 2001	■							■	
34	Reed, 1993	■								
35	Schnitzer, 1993							■	■	
36	Spinelli, 1998	■	■			■	■			
37	Stripling, 1993		■					■	■	
38	Suen, 1997			■						
39	Torrance, 1993		■					■		
40	Wiggins, 1989a	■			■	■		■		■
41	Wiggins, 1989b			■	■	■	■	■	■	■
42	Wiggins, 1990	■			■	■		■	■	
43	Wiggins, 1992			■						■
44	Wiggins, 1993	■		■	■	■		■	■	
45	Wiggins, 1996	■						■		
46	Wiggins, 2006	■				■	■			
47	Williams, 2004	■						■		

* Shaded areas indicated presence of the component.

Table 2. Definitions of authentic assessment for professional development (college and job training)*

		Context			Student Role			Scoring		
		activity or context	performance-based	cognitively complex	defense required	formative assessment	collaborative	developed criteria	indicators or portfolios	mastery expectation
	Percentage including Element	48%	8%	28%	20%	28%	16%	32%	52%	0%
1	Border, 1998			■					■	
2	Chance, 1997			■		■			■	
3	Darling-Hammond & Snyder, 2000	■						■	■	
4	Einbender & Wood, 1995	■								
5	Facione & Facione, 1996			■					■	
6	Fall, 1996	■							■	
7	Gulikers, Bastiaens & Kirschner, 2004	■		■	■			■	■	
8	Gulikers, Bastiaens & Martens, 2005	■								
9	Hanna, 2002			■		■	■			
10	Herrington & Herrington, 1998	■				■	■		■	
11	Howell, 1993		■					■		
12	Jackson, Draugalis, Slack & Zachry, 2002	■								
13	Jonietz, 1996				■			■		
14	LaLopa, 2004			■				■		
15	Lawver, Felstehausen, et al., 1994		■						■	
16	MacAndrew & Edwards, 2003				■		■	■		
17	Mallet, 2005, 2006				■	■			■	
18	Montgomery, 2002	■				■		■	■	
19	Oh, Kim, Garcia & Krilowicz, 2004				■				■	
20	O'Sullivan, 2005					■	■			
21	Rennert-Ariev, 2005	■		■		■		■		
22	Saunders, Saunders & Batson, 2001	■							■	
23	Schuwirth & van der Vleuten, 2003	■								
24	Svinicki, 2004	■								
25	Wenzel, Briggs & Puryear, 1998								■	

*Shaded areas indicated presence of the component.

only mentioned by 60% of school-aged assessment publications and 48% of adult assessment publications. Across all categories, close to half do not include realism in their definition of authentic assessment. Perhaps as commonly given as a criterion for authenticity is the use of multiple indicators for scoring or a multiple components system like the use of portfolios or scoring rubrics. This was required by authors of 54% of the K-12 articles and 52% of the professional development publications. For the school-aged publications, the relative frequencies of the remaining dimensions required as part of the definition of authentic assessment were *known or student developed criteria*, 47%, *formative assessment*, 31%, *cognitively complex*, 30%, *performance-base*, 23%, *collaborative*, 20%, *a required defense*, 15%, and *mastery expectation*, 13%. For the professional development publications, the relative frequencies of the remaining dimensions were *known or student developed criteria*, 32%, *formative assessment* and *cognitively complex*, 28% each, *a required defense*, 20%, *collaboration*, 16% and *performance-based*, 8%. No professional development articles presented a definition of authenticity including *mastery expectation*.

Authentic Assessment of School Children

Many researchers write of the nature of authentic assessment as it applies to K-12 education.

Table 3. Definitions of authentic assessment for various academic subjects and categories*

			Context			Student Role			Scoring		
			realistic activity or context	performance-based	cognitively complex	defense required	formative assessment	collaborative	known or student developed criteria	multiple indicators or portfolios	mastery expectation
1	Language Arts	Frazier, 1997	■				■		■	■	■
2		Hirvela & Pierson, 2000					■			■	
3		Karge, 1998						■	■	■	
4		Laurier, 2000								■	
5		Montgomery, 2001	■								
6		Stanford & Siders, 2001	■							■	
7		Valencia, Hiebert & Afflerbach, 1994	■		■		■	■		■	
1	ESL & Foreign Language	Bachman, 2000	■								
2		Diaz, 1999					■			■	
3		Ekbatani & Pierson, 2000					■				
4		Greenleaf, Gee & Ballinger, 1997		■			■			■	
5		Kohonen, 1997	■				■				
6		Lynch, 2003	■	■							
1	Social Studies	Avery, 1999		■							
2		McMann & McMann, 1992			■		■				
3		Montgomery, 2001	■		■						
4		Morris, 2001	■					■	■		
1	Science	Kamen, 1996		■						■	
2		Montgomery, 2001	■	■							
3		Nicaise, Gibney & Crane, 2000	■								
4		Vendlinski, Underdahl, Simpson & Stevens, 2002	■		■				■	■	
1	Math	Luitel, 2002		■						■	
2		McNaire, Thomson & Williams, 1998	■	■							
3		Montgomery, 2001	■		■						
4		Moschkovich, 1998	■								
5		Wilson, 1993	■				■				
1	Physical Education	Block, Lieberman & Connor-Kuntz, 1998	■	■					■		
2		Haywood, 1997		■					■		
3		Lund, 1997	■		■				■	■	
4		Mintah, 2003		■			■	■			
5		Smith, 1997		■						■	
1	Fine Arts	Bell & Bell, 2003		■			■				
2		Craig & McCormick, 2002		■							
1	Pre-School	Bagnato, 2005	■							■	
2		Grisham-Brown, Hallam & Brookshire, 2006	■							■	
3		Puckett & Black, 2008		■						■	

* Shaded areas indicated presence of the component.

For example, Wiggins (1989a), probably the most cited authenticity advocate, argues that teachers should "test those capacities and habits we think are essential and test them in context. Make them replicate within reason, the challenges at the heart of each discipline. Let them be- authentic." (p. 41) and presented four basic characteristics of authentic tests:

1. The task should be representative of performance in the field.
2. Attention should be paid to teaching and learning the criteria for assessment.
3. Self-assessment should play a great role.
4. When possible, students should present their work publicly and defend it.

In a separate article published the same year, Wiggins (1989b) again emphasizes the importance of real-world or representative tasks (e.g. conducting research, writing reports, assembling portfolios) and offers slightly extended criteria that included the ideal of collaboration among students and suggested that tasks and scoring should be complex. These four dimensions of authenticity are given:

1. *Structure and logistics.* The test becomes the task with learning occurring as part of the assessment. Authentic tests are public with evaluation based on judgment using agreed upon standards and prior experience and training. There are no unrealistic time constraints or secret questions**.** "Authentic tests require some collaboration with others." (p. 711).
2. *Intellectual design features.* The tasks are enabling and increase coherence of knowledge and level of problem-solving skills. Authentic tasks emphasize realistic but fair complexity; they stress depth more than breadth.
3. *Scoring.* Scoring must be complex and authentic tests cannot be scored on a curve, but instead are criterion-referenced, based on standards. As with formative assessment, self-assessment is central.
4. *Fairness and equity.* Authentic tests identify strengths. They don't use norm-referencing methods to widen the spread of scores.

In later work, Wiggins (1990, 1992, 1993) continued to emphasize that tasks should mirror real-world activities and elaborated that authentic tests should assess students' "habits of mind" (1993). Tasks are not authentic, necessarily, just because they are similar to real-world tasks, but they must mirror the complexity, collaboration, and high-level thinking that is necessary in the most intellectual of professional problem-solving and decision-making. The assessments act as instruction and skill-building opportunities, not merely as tools of evaluation.

Paris and Ayres (1994) describe authentic assessment in terms suggesting that authenticity requires that the assessments be formative. They join some who argue that authentic assessment, because it is formative, creates reflective students and teachers. Whether an assessment is authentic depends on local contexts, they contend that what is authentic in one school is not necessarily authentic in another, because authentic assessment is defined by locally valued outcomes of the curricula and must be aligned with instructional methods. The emphasis on the formative nature of the testing and the need for individualized (at the school or district level) customized context results in somewhat atypical criteria for authenticity. Authentic assessment:

1. supports classroom instruction,
2. collects evidence from multiple activities,
3. promotes learning and teaching among participants, and
4. reflects local values, standards and controls.

This list does not match Wiggins' lists except for the need for data from multiple sources. Interestingly, while a necessity, presumably, for reliable measurement, that particular requirement does not actually reflect the nature of "real world" activities, so is not driven by that key dimension of authenticity.

Another oft-cited set of broad characteristics for authentic assessment for school aged children is

provided by Darling-Hammond, Ancess & Falk (1995). In the same way that Paris and Ayres offer characteristics for authenticity that are more closely related to criteria for *formative* assessment, Darling-Hammond and colleagues present components of authenticity which while consistent with Wiggins, seem to broaden the criteria enough to allow consideration of quality *performance-based* assessments as authentic, even if they do not actually mirror real-world activities. They do explicitly agree with Wiggins (1989a) four points of authenticity, but go on to frame the authenticity requirements to include assessments that allow for demonstration of skills necessary for success outside of the classroom and the creation of products or solutions. This slight broadening of what the actual task can look like in order to be authentic should not be taken to mean that the authors do not emphasize the important of a strong link between the classroom task and the real-world. Demonstrating real-world competence remains the central dimension of authenticity. "Authentic assessments are also contextualized; that is rather than assembling disconnected pieces of information, the tasks are set in a meaningful context that provides connections between real-world experiences and school-based ideas. These assessments are connected to students' lives…" (p. 4).

Authentic Assessment of Pre-Service Professionals

Gulikers, Bastiaens and Kirschner (2004), Gulikers, Bastiaens and Martens (2005), Herrington and Herrington (1998), Rennert-Ariev (2005) and Svinicki (2004) are among those who have translated authentic assessment into the world of adults in professional training (e.g. teachers, nurses, etc.). In this context, the link between assessment and real-world professional activities is more crucial than ever and all emphasize that fidelity to situations that will be found in the field after graduation is an essential element of authentic assessment. An additional aspect of fidelity is the context of the assessment task. To be authentic, the assessment must be part of a realistic learning context (Gulikers, Bastiaens & Martens, 2005) and make sense in the context of student work (Rennert-Ariev, 2005). As with the standards for authenticity important for assessing school aged children, many authors include the requirement that scoring criteria are known by all and that multiple indicators are used for judgments (Gulikers, Bastiaens & Kirschner, 2004; Herrington & Herrington, 1998).

Authentic Assessment of Pre-School Children

Grisham-Brown, Hallam and Brookshire (2006) provide a well thought out vision of authentic assessment for pre-school aged children. Work, for children, is their play, and observing children in their natural playing and learning environment is the key for authenticity. Authentic assessment in the early childhood environment is described by the following characteristics: (1) conducted in natural environment; (2) uses multiple methods; (3) connection between assessment purpose and use; (4) involvement of families in assessment process. The involvement of families in the assessments parallels the role of the student in authentic assessment for school-aged children. Puckett and Black (2008) present a definition of authentic assessment for young children that includes the "four P's" of authentic assessment. The four words starting with *P* that are listed as characteristics of authenticity, however, seem to describe valid *performance-based* assessments in general, not what most other advocates would argue are the crucial dimensions of *authentic* assessment: *Process*, *Performance*, *Products* and *Portfolios*. Certainly, though, portfolios are a format that entails multiple evidence for scoring and that is consistent with typical guidelines for authenticity and they, like Grisham-Brown, Hallam and Brookshire, emphasize the importance of family involvement in the assessment.

Authentic Assessment in Subject Areas

Specific suggestions for what it means for assessments to be authentic in specific content areas are available. Hirvela and Pierson (2000), Montgomery (2000) and Valencia, Hiebert and Afflerbach (1994), among others, provide criteria for authentic assessment in the language arts. For assessment to be authentic, it should provide multiple examples, provide for student ownership and revision (when assessing writing skills), and include tasks that represent meaningful literacy use (Hirvela & Pierson, 2000). In her teachers'

handbook, Montgomery provides content-specific interpretations of authenticity standards for social studies, science and math, in addition to language arts. She is most concerned with the real-world nature of the application of the skills being assessed and the cognitive complexity of the problem. In assessing English-as-a-second-language (ESL) students, Ekbatani and Pierson (2000) argue that a test is authentic if it is learner-directed and acts as formative assessment. They agree with the criteria of Paris and Ayres (1994) that indicates that assessments that provide feedback to students and allow for student self-assessment are authentic. Bachman (2000) establishes only one key component of authentic assessment for ESL learners. The nature of the task must match the nature of the task in the *target language use* domain.

Discussion

The concept of authenticity is complex. Though some authors and researchers use the term simply as a synonym for *realistic*, it is clear that the idea of authenticity as used in the field encompasses much more than realism. This study focused on the various meanings of the concept expressed in the research and training literature. The preponderance of the publications reviewed concerned themselves with classroom assessment, but many authors, especially with the earlier works, were reacting to the "inauthentic" nature of most large-scale and standardized tests. As such, it is reasonable to assume that the conclusions reached here as to the definition of authenticity in a teacher-made testing context also apply to other student assessment contexts.

The *authentic* label is often placed on assessments that are performance-based or involve cognitively-complex tasks, without regard to whether the tasks are similar to those valued outside the classroom. Other definitions of authenticity are based on whether a defense is required, whether collaboration is involved, or the level of student involvement in determining scoring criteria. About a third of the time, a meaning is used that suggests that if the purpose of an assessment is formative, then it is authentic.

The imprecision reflected in the overlap between the use of terms such as *formative* and *authentic* is just one of many ways in which the concept of authenticity, as used in the scholarly literature, has drifted away from the basic "real world tasks with real world evaluation" definition. For example, as equally important as *realism* to the definition of authentic assessment in the publications we examined was that the evaluation includes multiple indicators of performance. We included the use of portfolios and the use of scoring rubrics in this category because they imply multiple scores. The use of portfolio assessment and the use of scoring rubrics for performance assessments are both considered best practices in classroom assessment, of course, but they do not necessarily reflect the real-world evaluation of real-world tasks. Similarly, it is probable that portfolios would likely add to the validity of most assessment strategies, and scoring rubrics with multiple indicators would likely increase the reliability of the scoring, but their use is not clearly more "realistic". Real-world job expectations might be evaluated by the performance on a single task, for example, though it is true that some real world judgments (receiving an advanced degree, selection for promotion) can be based on a portfolio of work.

Other popular components of authenticity, on the other hand, do seem associated with real-world tasks. Tasks or "assignments" outside of the classroom are often cognitively complex and the criteria for success are typically clear and known by all. It is also difficult to imagine a real-world task that is not performance-based to some degree. (As is often pointed out, few students end up with jobs where they get paid to fill out multiple-choice test bubble sheets.) Whether these dimensions are conceptually distinct from *realism* cannot be judged solely by the analysis of publications we performed and that question remains unanswered.

Of course, it is not wise to assume that aspects of authenticity not emphasized in published definitions are missing from the authors' actual conceptualizations or that most assessments labeled as authentic do not as correlated characteristics include those aspects of authenticity. For instance, might all assessments by necessity have to be

performance-based to be realistic? Frey & Schmitt (2007) have suggested as much. Can a requirement for a *defense* of one's work exist without students' having engaged in some sort of self-reflection and evaluation consistent with the nature of *formative assessment*?

As noted, many definitions of authenticity appear to be paraphrased descriptions of validity. Palm's (2008) observation of the term as usually meaning some version of "being true" is consistent with this finding. It is tempting in educational research and among practitioners (e.g. trainers, classroom teachers, education professors) to bunch together a small set of "good" characteristics of assessment ("tests should be valid and reliable" with "valid and reliable" often running together when spoken as if they are one thing) and to discuss these characteristics jointly until they run the risk of blending into one *good* thing, one best practice. So it becomes as if performance assessment is authentic assessment and authentic is valid and valid is reliable and reliable means multiple indicators and multiple indicators mean scoring rubrics and scoring rubrics mean performance assessment. There is consensus in the field of education that classroom assessments should be authentic, but there is no consensus of what that means.

What, then, is the correct or best definition of authenticity? Two strategies for defining a concept as useful and influential as authentic assessment are to trace the origins and first uses of the term in the relevant literature or to identify theoretically the crucial components of the definition in terms of what is needed for the assessment approach to have value. In other words, one can search for how the inventors of the term defined it or one can identify what authentic assessment must include in order to "work". Either approach might allow for a reasonable decision about the "right" definition.

Recommendations

The term *authentic* as applied to tests appears to have been used first by Archbald and Newman in 1988 in a book about the weaknesses of standardized testing and the difficulties in measuring "authentic" academic achievement in high school. As used here, assessments are only authentic if they have meaning or value beyond the score or grade that participation might produce. In other words, the assessment task itself should be meaningful. This suggests that assessments that require behaviors or cognitive operations that are not intrinsically meaningful, (e.g. responding to multiple-choice questions on an externally produced standardized tests) are not authentic. Conversely, the definition suggests that assessment tasks that are interesting, require complex thought, and require high levels of student participation are authentic. The other early advocate for authentic assessment, and the author most closely associated with the term, was Wiggins (1989a, 1989b). Among the critical dimensions emphasized in those early arguments was the need for a public defense of the student work and the value of a mastery approach to the task and the assessment. It is interesting to observe that these components are rarely mentioned in publications by others that came later, even in those works that cite Wiggins as the "father" of authentic assessment. While today's advocates for authenticity typically are opposed to most traditional norm-referenced large-scale standardized tests and prefer the mastery approach to assessment, the value of a defense, especially the public defense, of one's work is not often reflected in more recent writings on authentic assessment. The views of Wiggins, along with Archbald and Newman, support a definition of authentic assessment as assessment that poses an intellectually interesting and personally meaningful problem or task. It would be consistent with this definition to call these assessments realistic because, by definition, the questions, tasks or problems have value and interest beyond the classroom into the "real-world" of the students' values, abilities and motivations.

To determine a definition of authentic assessment based on what characteristics must be present for authentic assessment to "work", it is necessary to identify the unique goal of authentic assessment as opposed to the *other* kinds of assessment (e.g. traditional paper-and-pencil tests, standardized tests, and so on). Authentic assessment is supposed to engage the student; it works when the student has found it to be rewarding for its own

sake. Because tasks that require complex, interesting or creative cognitive activity and also seem meaningful or useful to students are the very tasks that are likely to be rewarding, those types of assessments should, reasonably, be considered authentic. It also suggests that student involvement in establishing scoring rules and self-evaluation of their own work adds to authenticity because it will increase student ownership and commitment to the task. This approach to definition results in criteria consistent with the early adopters of the term and consistent with the results of the first strategy of the two strategies for defining the term.

To settle on a parsimonious and academically useful definition of "authentic assessment", it is best to strip away those requirements that are not central to the unique worth of the approach. This means that reliability tactics such as the use of multiple indicators and portfolio systems are not definitional components. They add to the quality of assessment in general, no doubt, but are not exclusive to authentic assessment. Similarly, some suggested elements of authenticity which likely increase validity, such as the requirement of a public defense or that the assessment must be part of a formative system, are unnecessary as part of a definition. Those crucial elements which remain result in a description of a classroom assessment task that involves the student deeply, both in terms of cognitive complexity and intrinsic interest, and are meant to develop or evaluate skills and abilities that have value beyond the assessment itself. It is this type of assessment experience that is, realistically speaking, authentic.

References

Abernethie, L. (2006). Authentic assessment as a neo-liberal technology of government. *Teaching and Learning Forum*, 1-9.

Airasian, P.W. (2001). *Classroom assessment: Concepts and applications (4th ed.)*. Blacklick, OH: McGraw-Hill.

Archbald, D.A. (1991). Authentic assessment: Principles, practices, and issues. *School Psychology Quarterly, 6*(4), 279-293.

Archbald, D.A., & Newmann, F.M. (1988). *Beyond standardized testing: Assessing authentic academic achievement in the secondary school.* Washington, DC: Office of Educational Research and Improvement.

Avery, P. (1999). Authentic assessment and instruction. *Social Education 63*(6): 368-373.

Bachman, L.F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Achievement in the Secondary School, 17*(1), 1-42.

Bagnato, S. (2005). The authentic alternative for assessment in early intervention: An emerging evidence-based practice. *Journal of Early Intervention 28*(1): 17-22.

Bell, A., & Bell, M. (2003). *Developing authentic assessment methods from a multiple intelligences perspective* (Master of Arts Action Research Project, Saint Xavier University and SkyLight Professional Development Field-Based Master's Program, 2003).

Bergen, D. (1993-1994). Authentic performance assessments. *Childhood Education, 70*(2), 99-102.

Biondi, L.A. (2001). *Authentic assessment strategies in fourth grade.* EDU: Collaborative Action Research, Mangrove College.

Block, M., Lieberman, L., & Connor-Kuntz, F. (1998). Authentic assessment in adapted physical education. *Journal of Physical Education, Recreation & Dance 69*(3): 48-56.

Border, B. (1998). *The status of alternative assessments through the 1990s: Performance and authentic assessments in relation to vocational-technical education technical skills, workplace skills, and related academic skills.* Decatur, Georgia: Vocational Technical Education Consortium of States.

Borowski, M., Thompson, C., & Zaccaria, K. (2001). *Portfolios: authentic assessment* (Master of Arts Action Research Project, Saint Xavier University and SkyLight Professional Development Field Based Master's Program, 2001).

Bullens, D. (2002). *Authentic assessment: Change for the future* (Master of Arts Action Research Project, Saint Xavier University and SkyLight Professional Development Field-Based Master's Program, 2002).

Brandt, R. (1996). On Authentic Performance Assessment: A Conversation with Grant Wiggins. *Educational Leadership, 54*(4), 1-5.

Burley, H. P. & Price, M. (2003). What works with authentic assessment. *Educational Horizons*, 193-196.

Chance, B. (1997). Experiences with authentic assessment techniques in an introductory statistics course. *Journal of Statistics Education, 5*(3).

Craig, C.L., & McCormick, E.P. (2002). *Improving student learning through authentic assessment* (Master of Arts Action Research Project, Saint Xavier University and SkyLight Professional Development Field-Based Master's Program, 2002).

Cronin, J. (1993). Four misconceptions about authentic learning. *Educational Leadership, 50*(7), 78-80.

Cumming, J.J. & Maxwell, G.S. (1999). Contextualizing authentic assessment. *Assessment in Education, 6*(2), 177-194.

Darling-Hammond, L. (1994). Setting standards for students: The case for authentic assessment. *Educational Forum, 59*(1), 14-21.

Darling-Hammond, L. & Snyder, J. (2000). Authentic assessment of teaching in context. *Teaching and Teacher Education, 16*(5-6), 523-545.

Darling-Hammond, L., Ancess, J. A., & Falk, B. (1995). *Authentic assessment in action: Studies of schools and students at work.* New York: Teachers College Press.

Dez, M., Moon, J., & Meyer, C. (1992). What do we want students to know? ... and other important questions. *Educational Leadership, 49*(8), 38-42.

Dutt-Doner, K., & Maddox, R. (1998). Implementing authentic assessment. *Kappa Delta Pi Record, 34*(4), 135-137.

Ekbatani, G., & Pierson, H. (Eds). (2000). *Learner-directed Assessment in ESL.* Mahwah, NJ: Lawrence Erlbaum.

Einbender, L., & Wood, D. (1995). *An authentic journey: Teachers' emergent understandings about authentic assessment and practice.* New York : National Center for Restructuring Education, Schools, and Teaching (NCREST).

Engel, M., Pulley, R. & Rybinski, A. (2003). *Authentic assessment: It really works* (Masters action research project, Saint Xavier University and SkyLight, 2003).

Facione, N. & Facione, P. (1996). Assessment design issues for evaluating critical thinking in nursing. *Holist Nursing Practice, 10*(3), 41-53.

Fall, L.T. (1996). *Management by objectives: Authentic assessment in a public relations practicum.* Paper presented at the Annual Meeting of the Association for Education in Journalism and Mass Communication (79th, Anaheim, CA, August 9-13, 1996).

Frazier, C. (1997). The development of an authentic assessment instrument: The scored discussion. *The English Journal, 86*(1), 37-40.

French, D. (2003). A new vision of authentic assessment to overcome the flaws in high stakes testing. *Middle School Journal, 35*(1), 14-23.

Frey, B.B., Petersen, S.E., Edwards, L.M., Pedrotti, J.T. & Peyton, V. (2005). Item-writing rules: Collective wisdom. *Teaching and Teacher Education, 21,* 357-364.

Frey, B.B. & Schmitt, V.L. (2006, April). *Coming to terms with classroom assessment: Are performance-based assessments authentic or visa-versa?* Presented at the Annual Meeting of the American Educational Research Association, San Francisco.

Frey, B.B. & Schmitt, V.L. (2007). Coming to terms with classroom assessment. *Journal of Advanced Academics, 18, 3*, 402-423.

Grabinger, R. & Dunlap, J.C. (1995).Rich environments for active learning *ALT-J, 3*(2), 5-34.

Green, J. (1998). Authentic assessment: Constructing the way forward for all students. *Education Canada, 38*(3), 8-12.

Greenleaf, C., Gee, M., & Ballinger, R. (1997). *Authentic assessment: Getting started.* Washington, DC: Office of Vocational and Adult Education (ED), National Workplace Literacy Program.

Grisham-Brown, J., Hallam, R., & Brookshire, R. (2006). Using authentic assessment to evidence children's progress toward early learning standards. *Early Childhood Education Journal, 34*(1), 45-51.

Gronlund, G. (2003). *Focused early learning: A planning framework for teaching young children.* St. Paul, MN: RedleafPress.

Gulikers, J., Bastiaens, T., & Kirschner, P. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research & Development, 52*(3), 67-86.

Gulikers, J., Bastiaens, T., & Martens, R. (2005). The surplus value of an authentic learning environment. *Computers in Human Behavior, 21*, 509-521.

Haladyna, T. M. & Downing, S.M. (1989). A taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education, 2*(1), 37-50.

Haladyna, T.M., Downing, S.M., & Rodriguez, M.C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education, 15*(3), 309-334.

Hanna, N. (2002). Effective use of a range of authentic assessments in a web assisted pharmacology course. *Educational Technology & Society, 5*(3), 1-19.

Haywood, K. (1997). Authentic assessment and movement requires a developmental approach. Central District Scholar Address. *Missouri Journal of Health, Physical Education, Recreation and Dance, 7*(3-15), 3-15.

Herrington, J. & Herrington, A. (1998). Authentic assessment and multimedia: How university students respond to a model of authentic assessment. *Higher Education Research and Development, 17*(3), 305-322.

Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development, 48*(3), 23-48.

Hirvela, A., & Pierson, H. (2000). *Portfolios: Vehicles for authentic self-assessment.* In Ekbatani, G. and Pierson, H., editors, 2000: Learner-directed Assessment in ESL. Mahwah, NJ: Erlbaum, 105–26.

Howell, K. (1993). Bias in authentic assessment. *Diagnostique, 19*(1), 387-400.

Hunter, D. (2001). The large-scale, authentic assessment of listening and speaking as interactive communication: Issues in reliability. *Alberta Journal of Educational Research, 47*(2), 156-172.

Jackson, T., Draugalis, J., Slack, M. & Zachry, W. (2002). Validation of authentic performance assessment: A process suited for rasch modeling. *American Journal of Pharmaceutical Education, 66*, 233-243.

Jolly, J. & Kettler, T. (2000). Authentic assessment of leadership in problem-solving groups. *Gifted Child Today, 27*(1), 32-39.

Jonietz, P. (1996). *Qualitative/action research for authentic assessment*. Paper presented at the CEC/TED Conference (Washington, DC, November 8, 1996).

Kamen, M. (1996). A teacher's implementation of authentic assessment in an elementary science Classroom. *Journal of Research in Science Teaching, 33*(8), 859-877.

Karge, B. (1998). Knowing what to teach: using authentic assessment to improve classroom instruction. *Reading & Writing Quarterly, 14*(3), 319-331.

Kellaghan, T., & Madaus, G.F. (1993). The British experience with 'authentic' testing. *Phi Delta Kappan, 74*(6), 458-466.

Kieffer, R. & Morrison, L. (1994). Changing portfolio process: One journey toward authentic assessment. *Language Arts, 71*(6), 411-418.

Kohonen, V. (1997). *Authentic assessment as an integration of language learning, teaching, evaluation and the teacher's professional growth*. University of Tampere.

Kubiszyn, T., & Borich, G.D. (2003). *Educational testing and measurement: Classroom application and practice (7th ed.)*. New York: J.Wiley & Sons.

LaLopa, J. (2004). Developing a student-based evaluation tool for authentic assessment. *New Directions for Teaching and Learning, 100,* 31-36.

Laurier, M. (2000). Can computerized testing by authentic? *ReCALL, 12*(1), 93-104.

Lawton, D. (2000). Authentic assessment in South Brunswick schools. *Knowledge Quest, 29*(2), 30-38.

Lawver, D., Felstehausen, G., Cepica, M., et al. (1994). *Authentic assessment for occupational comepetency for career and cechnology education. Final report (year one)*. Austin: Texas Education Agency.

Luitel, B. C. (2002). *Authentic and portfolio assessment*. Curtin University of Technology.

Lund, J. (1997). Authentic assessment: Its development & applications. *The Journal of Physical Education, Recreation & Dance, 68*(7), 25-40.

Lynch, R. (2003). Authentic, performance-based assessment in ESL/EFL reading instruction. *Asian EFL Journal*.

MacAndrew, S. & Edwards, K. (2003). Essays are not the only way: A case report on the benefits of authentic assessment. *Psychology Learning and Teaching, 2*(2), 134-139.

Maden, J., & Taylor, M. (2001). *Developing and implementing authentic oral assessment instruments*. Paper presented at the Annual Meeting of Teachers of English to Speakers of Other Languages (35th, St. Louis, MO, February 27-March 3, 2001).

Mallet, D. G. (2005). An investigation of some authentic assessment for advanced undergraduate engineering and applied mathematics students. (http://www.studentweb.usq.edu.au).

Mallet, D. G. (2006). Authentic assessment for advanced undergraduate students. In Stacey, Andrew and Blyth, Bill and Shepard, John, Eds. *Australian Mathematical Society,* 1-14.

McMann, F.C. & McMann, C.J. (1992). Authentic evaluation in history. *Organization of American Historians Magazine of History*, Spring..

McNair, S., Thomson, M., & Williams, R. (1998). *Authentic assessment of young children's developing concepts in mathematics and science*. Australia and New Zealand Conference.

Meisels, S. (1996). Using work sampling in authentic assessment. *Educational Leadership*, 60-64.

Meisels, S. (2001). Parental reactions to authentic performance assessment. *Educational Assessment, 7*(1), 61-83.

Mertler, C. A. (2003). *Preservice versus inservice teachers' assessment literacy: does classroom experience make a difference?* Paper presented at the Annual Meeting of the Mid-Western Educational Research Association (Columbus, OH, October 15-18, 2003).

Meyer, C. (1992). What's the difference between authentic and performance assessment? *Educational Leadership, 49*(8), 39-40.

Mintah, J. (2003). Authentic assessment in physical education: prevalence of use and perceived impact on students' self-concept, motivation, and skill achievement. *Measurement in Physical Education and Exercise Science, 7*(3), 161-174.

Montgomery, K. (2001). *Authentic assessment: A guide for elementary teachers*. Austin, TX: Pro Ed.

Montgomery, K. (2002). Authentic tasks and rubrics: Going beyond traditional assessments in college teaching. *College Teaching*, 34-39.

Moorcroft, T., Desmarais, K., Hogan, K., & Bekowitz, A. (2000). Authentic assessment in the informal setting: How it can work for you. *The Journal of Environmental Education, 31*(3), 20-24.

Morris, R. (2001). Drama and authentic assessment in a social studies classroom. *The Social Studies, 92*(1), 41-44.

Moschkovich, J. N. (1998). Rethinking authentic assessments of student mathematical activity. *Focus on Learning Problems in Mathematics, 20*(4), 1-18.

Mueller, J. (2005). Authentic assessment in the classroom and the library media center. *Library Media Connection, 23*(7), 1418.

Newmann, F., Brandt, R.. & Wiggins, G. (1998). An exchange of views on semantics, psychometrics, and assessment reform: a close look at 'authentic' assessments. *Educational Researcher, 27*(6), 19-22.

Nicaise, M., Gibney, T., & Crane, M. (2000). Toward an understanding of authentic learning: student perceptions of an authentic classroom. *Journal of Science Education and Technology, 9*(1), 79-94.

Oh, D., Kim, J., Garcia, R., & Krilowicz, B. (2004). Valid and reliable authentic assessment of culminating student performance in the biomedical sciences. *The American Physiological Society, 29,* 83-93.

Oosterhof, A. (2001). *Classroom applications of educational measurement (3rd ed.).* Upper Saddle River, NJ: Prentice-Hall.

O'Sullivan, M. L. (2005). Authentic individual self-assessment, praxis and community. *Herdsa*, 351-357.

Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. *Practical Assessment, Research and Evaluation, 13*, 4. 1-11.

Paris, S.G. & Ayres, L.R. (1994). *Becoming reflective students and teachers with portfolios and authentic assessment.* ERIC Reproductive Services No. 378166.

Popham, W.J. (2002). Right task, wrong tool. *American School Board Journal, 189*(2), 18-22.

Puckett, M.B., & Black, J.K. (2008). *Authentic assessment of the young child: celebrating development and learning (2nd ed.).* Des Moines, IA: Prentice-Hall Inc.

Ratcliff, N. (2001). Using authentic assessment to document the emerging literacy skills of young children. *Childhood Education, 78*(2), 66-69.

Reed, L. (1993). Achieving the aims and purposes of schooling through authentic assessment. *Middle School Journal, 25*(2), 11-13.

Rennert-Ariev, P. (2005). A theoretical model for the authentic assessment ofteaching. *Practical Assessment, Research & Evaluation, 10*(2), 1-11.

Saunders, N., Saunders, G., & Batson, T. (2001). *Assessment and the adult learner: does authentic assessment influence learning?* Paper presented at the Annual Meeting of the Mid-Western Educational Research Association (Chicago, IL, October 24-27, 2001).

Schnitzer, S. (1993). Designing an authentic assessment. *Educational Leadership, 50*(7), 32-35.

Schuwirth, L.W.T., & Van der Vleuten, C.P.M. (2003). The use of clinical simulations in assessment. *Medical Education, 37*, 65-71.

Smith, T. (1997). Authentic assessment: using a portfolio card in physical education. *Journal of Physical Education, Recreation and Dance, 68*(4), 46-52.

Spinelli, C.G. (1998). *Teacher education reform: promoting interactive teaching strategies and authentic assessment for instructing an increasing diverse population of students.* West Long Branch, NJ: 120 Opinion Papers.

Stanford, P. & Siders, J. (2001). Authentic assessment for intervention. *Intervention in School and Clinic, 36*(3), 163-167.

Stripling, B. (1993). Practicing authentic assessment in the school library. *School Library Media Annual (SLMA), 11*, 40-57.

Suen, H.K. (1997). Concept map as scaffolding for authentic assessment. *Psychological Reports, 83*(3,1), 734.

Svinicki, M. (2004). Authentic assessment: Testing in reality. *New Directions for Teaching and Learning,* (100), 23-29.

Taylor, C.S., & Bobbitt-Nolen, S. (2005). *Classroom assessment:Supporting teaching and learning in real classrooms (2nd ed.).* NJ: Prentice-Hall, Inc.

Terwilliger, J. S. (1998). Rejoinder: response to wiggins and newmann. *Educational Researcher, 27*(6), 22-23.

Torrance, H. (1993). Combining measurement-driven instruction with authentic assessment: Some initial observations of national assessment in England and Wales. *Educational Evaluation and Policy Analysis, 15*(1), 81-90.

Valencia, S.W., Hiebert, E., & Afflerbach, R. (1994*). Realizing the possibilities of authentic*

assessment. In S. Valencia, & P. Afflerbach (Eds.), Authentic reading assessment: Practices and possibilities. Newark, DE: International Readding Association.

Vendlinski, T., Underdahl, J., Simpson, E. & Stevens, R. (2002). *Authentic assessment of student understanding in near-real time!* In: NECC 2002: National Educational Computing Conference Proceedings (23rd, San Antonio, Texas, June 17-19, 2002).

Wenzel, L., Briggs, K., & Puryear, B. (1998). Portfolio: Authentic assessment in the age of the curriculum revolution. *Journal of Nursing Education, 37*(5), 208-212.

Wiggins, G. (1989a). Teaching to the (authentic test). *Educational Leadership, 46*(7), 41-47.

Wiggins, G. (1989b). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan, 70*(9), 703-716.

Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research & Evaluation, 2*(2).

Wiggins, G. (1992). Creating tests worth taking. *Educational Leadership, 49*(8), 26-34.

Wiggins, G. (1993). Assessment to improve performance, not just monitor it: Assessment reform in the social sciences. *Social Science Record, 30*(2), 5-12.

Wiggins, G. (1996). Practicing what we preach in designing authentic assessments. *Educational Leadership*, 18-25.

Wiggins, G. (2006). Healthier testing made easy: Tests don't just measure absorption of facts. They teach what we value. *Edutopia*, 49-51.

Williams, J.B. (2004, December). Creating authentic assessments: *A method for the authoring of open book open web examinations.* Paper presented at *21st ASCILITE Conference: Beyond the Comfort Zone.* Perth, Western Australia.

Wilson, L. D. (1993). *What gets graded is what gets valued.* Paper presented at the Annual Meeting of the American Educational Research Association, Atlanta, GA. ERIC Document Reproduction Service No. 364577

Citation:

Frey, Bruce B., Schmitt, Vicki L., & Justin P. Allen (2012). Defining Authentic Classroom Assessment. *Practical Assessment, Research & Evaluation*, 17(2). Available online: http://pareonline.net/getvn.asp?v=17&n=2

Authors:

Bruce B. Frey
University of Kansas

Justin P. Allen
University of Kansas

Vicki L. Schmitt
Logan-Rogersville School District
Rogersville, MO

Corresponding Author:

Bruce Frey, Ph.D.
Psychology and Research in Education
University of Kansas
1122 West Campus Road, Room 643
Lawrence, KS 66045
BFrey [at] ku.edu