A FRAMEWORK FOR DETERMINING THE AUTHENTICITY OF ASSESSMENT TASKS: APPLIED TO AN EXAMPLE IN LAW

Dr Kelley Burton
School of Law
QUT, Australia
kj.burton@qut.edu.au

Abstract

Authentic assessment tasks enhance engagement, retention and the aspirations of students. This paper explores the discipline-generic features of authentic assessment, which reflect what students need to achieve in the real world. Some assessment tasks are more authentic than others and this paper designs a proposed framework supported by the literature that aids unit co-ordinators to determine the level of authenticity of an assessment task. The framework is applied to three summative assessment tasks, that is, tutorial participation, advocacy exercise and problem-based exam, in a law unit. The level of authenticity of the assessment tasks is compared and opportunities to improve authenticity are identified.

Introduction

Student feedback is a key driver of the need to embed authentic assessment. The Australasian Survey of Student Engagement (AUSSE) Report confirms that "students feel their educational experience has been most valuable when they are challenged to learn in a supportive environment and have encountered work-relevant learning experiences" (Australian Council for Educational Research, 2009, p. 23). Authentic assessment is valued by students because it represents what they need to achieve in the workplace (James, McInnis & Devlin, 2002).

Authentic assessment has the potential to have a positive influence on students, for example, it "can raise aspirations and increase intrinsic student motivation through explicit demonstration of career alignment and relevance of curriculum activities" (QUT Office of Teaching Quality, 2009, p. 8). According to West and Heath (2009), the "net goal of authentic pedagogies is to engage learners in the integration of theory and practice earlier, more frequently and more substantially" indicating that it would be beneficial to the learning of many students if theory and practice were incorporated in all levels of a course or degree (p. 465).

Burton (2009) posited that adopting a fixed approach to the concept of 'authentic assessment' and hinging it around what is done in the workplace today is flawed as this may not be relevant in the future, leading to the suggestion that authentic assessment should be fluid because of the shifting nature of our world. Fluidity will help prepare students for lifelong learning. With a view to unpacking the notion of authentic assessment, this article attempts to tease out the discipline-generic features of 'authentic assessment.' It then integrates the features into a framework that aids unit co-ordinators to appreciate the level of authenticity or the "appropriate degree of realism" of their assessment tasks (Gronlund, 2003, p. 124). The framework also highlights to unit co-ordinators how to improve the authenticity of their assessment tasks. This article applies the

authentic assessment framework to three assessment tasks, that is, tutorial participation, advocacy exercise and problem-based exam, in a law degree unit.

Unpacking 'authentic assessment'

Authentic assessment is "closely aligned with activities that take place in real work settings, as distinct from the often artificial constructs of university courses" (Boud & Falchikov, 2007, p. 23). It has a sense of realism and addresses real world problems (Frey & Schmitt, 2007; Newman, Brandt & Wiggins, 1998). However, equating authentic assessment to the real world is only a starting point because the requirements of the real world are discipline-specific and, as mentioned above, alter over time, for example, as practises improve and as the regulation of professionals change. Assessment can be seen as authentic if it resembles issues or activities that are realistic in the workplace or real life situations in which the students may find themselves (Gulikers, Bastiaens & Kirschner, 2006). Thus, authentic assessment is a relative notion contingent on what happens in practice, which varies across disciplines. To gain a deeper understanding of authentic assessment that can be applied in the future and across disciplines and in the context of all units, this article seeks to briefly explore six approaches that unpack the features of 'authentic assessment' and then centres on the prominent features of authentic assessment.

Herrington and Herrington (1998; 2006), who are leading scholars on authentic assessment, developed guidelines for determining authentic assessment. They categorise their guidelines into four groups, that is, context, student factors, task factors and indicators. The first group, context, emphasises fidelity of the task conditions and connectedness to the real world. Student factors include problem-solving skills, higher order thinking, the production of knowledge rather than the reproduction of knowledge, significant student time, collaboration, effective performers, polished products and depth of knowledge. Task factors include wide range of responses, complexity, ill-structured, judgments and multiple steps and integrated assessment. The indicators include learning, validity, reliability and criteria.

Similarly, Herrington, Oliver and Reeves (2002) developed 10 elements of an authentic task in an online environment, which are: has real world relevance, can be applied across subject areas and beyond specific outcomes, offers the opportunity to reflect and collaborate, is investigated over a sustained period of time, creates a polished product valuable in its own right, can be viewed from multiple perspectives using a variety of resources, permits diverse outcomes, involves a complex task that require students to determine sub-tasks, and is seamlessly integrated with assessment. While the ten elements were designed for an online environment, they are equally applicable to face-to-face and blended learning environments. Their work is labelled as Approach 2, and in many respects, it mirrors the features identified by Herrington and Herrington (1998; 2006) with continuing emphasis on relevance beyond the classroom to the real world, diversity of outcomes, complex tasks and integration with assessment.

Approach 3 consists of a five dimensional framework designed by Gulikers, Bastiaens and Kirschner (2006). These are: the nature of the assessment task, the physical context of the assessment (conditions of the assessment), the social context of the assessment (collaboration), the output of the assessment (polished products) and assessment criteria. These dimensions have already been embraced by Approaches 1 and 2, and do not elucidate any additional features of authentic assessment.

Approach 4 is based on the work of Frey and Schmidt (2007) who recognise the following features of authentic assessment: nature of the stimuli, complexity, conditions, resources, consequences and whether tasks are determined by an assessor or student. These features have already been canvassed by the earlier approaches.

Another approach to unpacking the features of authentic assessment has been adopted by Keyser and Howell (2008) and this is Approach 5. Whilst using some different terminology, their approach essentially isolates the features highlighted in the earlier approaches. In particular, their

features of authentic assessment include mimicking the work of professionals in real world problems, open-ended inquiry, thinking skills, meta-cognition, collaboration and complex tasks that require students to make choices.

The sixth and final approach considered in this article to understand the features of authentic assessment is espoused by Burkill, Dunne, Filer and Zandstra (2009). Approach 6 places emphasis on product as well as process, the development of real world and higher order cognitive skills (analysis, synthesis and evaluation), the integration of a range of skills into a whole project and the construction of novel ideas and responses. These features largely coincide with the features identified in the earlier approaches.

The six approaches provide a rich insight into the features of authentic assessment, and the prominent features are discussed below.

Fidelity of the task to the real world

Herrington and Herrington (1998, 2006) identified *context* as a feature of authentic assessment which focuses on the connectedness to the real world. Herrington, Oliver and Reeves (2002) label this as completing a task in real world conditions, while Gulikers, Bastiaens and Kirschner (2006) refer to this as the physical context of the assessment and Frey and Schmidt capture the essence of this by centring on consequences. The importance of authentic assessment having consequences in the real world has also been propounded by Boud and Falchikov (2007). This feature is over-arching and reflects the starting point of authentic assessment. Other features provide a greater insight into authentic assessment.

Polished product valuable in its own right

One of Herrington, Oliver and Reeves' (2002) elements of an authentic task in an online environment is that authentic assessment can be applied across subject areas and beyond specific outcomes. This element correlates to the output of the assessment and whether a polished product valuable in its own right is developed (Gulikers, Bastiaens & Kirschner, 2006; Herrington & Herrington, 1998, 2006;). This is consistent with Burkill, Dunne, Filer and Zandstra's (2009) assertion that authentic assessment requires the construction of novel ideas rather than the reproduction of knowledge.

'Consequences' is one of the features of authentic assessment identified by Frey and Schmidt (2007), and it appears to be a broad and vague criterion for authentic assessment. Arguably it sits within the task factors from Herrington and Herrington's (1998; 2006) approach and focuses on a polished product. Keyser and Howell's (2008) approach to authentic assessment is the only approach canvassed in this article that omits the need for a polished product.

High order thinking, reflection and self-assessment

Similarly to collaboration, higher order thinking is connected to greater student engagement (Australian Council for Educational Research, 2009). Reflection is an example of higher order thinking and falls within the realm of Herrington and Herrington's (1998; 2006) student factors. As student engagement and retention are significant, authentic assessment tasks should expect students to utilise higher order thinking skills.

Notably, Herrington, Oliver and Reeves (2002) identified reflection as being important to authentic assessment. This is analogous to the higher order thinking feature of authentic assessment in Herrington and Herrington's (1998, 2006) previously discussed. Assimilating reflection, particularly in the first year experience, is integral to developing "emotional intelligence necessary to survive" practice (McNamara, Field & Brown, 2009, p. 2). The reflection feature of authentic assessment has the ability to conflict with another feature of authentic assessment, that is, the seamless integration of the task and authentic assessment. To overcome this potential

conflict, reflection should be realistically embedded in the assessment task, rather than being added as a separate or contrived activity.

Approach 6 situates self-assessment within authentic assessment (Burkill et al., 2009), and arguably this corresponds with reflection. Self-assessment, and even peer assessment for that matter, are important to tertiary education, particularly to the first year experience because it "cultivate[s] responsibility for engagement in the learning process" (Krause, 2005, para 8). The other five approaches did not raise self-assessment as a feature of authentic assessment.

Collaboration

Herrington and Herrington (1998, 2006) and Herrington, Oliver and Reeves (2002) identified collaboration as a feature of authentic assessment. Similarly, Gulikers, Bastiaens and Kirschner (2006) identified social context of assessment as an important dimension of authentic assessment. Collaboration has the impact of enhancing student engagement in learning and unit co-ordinators should capitalise on an opportunity to embed teamwork in assessment in a creative manner, both within and outside the classroom environment (Krause, 2005).

Make judgments and choices, determine sub-tasks, complete a complex and ill-structured task, open-ended inquiry and create novel ideas and responses

Another student factor from Approach 1, previously listed, requires significant student time. The incorporation of this feature is perplexing because the term 'significant' is vague and relative. Does it mean days, weeks or months? Where is the dividing line between significant and insignificant from a time perspective? The length of time spent on an assessment task surely does not distinguish an authentic assessment task from an artificial assessment task. The amount of student time spent on an authentic assessment task should depend on how long it takes a professional to do the same task in practice. Taking this view, significant student time more appropriately reflects the need for an authentic assessment task to require students to make judgments and choices, determine sub-tasks, complete a complex and ill-structured task under realistic conditions, and create novel ideas and responses. This view is consistent with the work of Frey and Schmidt (2007) which lists complexity as a feature of authentic assessment, while Keyser and Howell (2008) noted the need for an open-ended inquiry. Further, significant student time is already canvassed within the context and task factors identified by Herrington and Herrington (1998; 2006), and Burkill, Dunne, Filer and Zandstra's (2009) feature of constructing novel ideas and responses.

Task seamlessly integrated with assessment

Unit co-ordinators should be aware that simply adding a reflective task onto an assessment task does not necessarily invoke authentic assessment. The reason for this assertion is that authentic assessment requires the assessment to be integrated in the task rather than being separate. Herrington and Herrington (1998; 2006) and Herrington, Oliver and Reeves (2002) canvass this feature. Integrating assessment with an activity is consistent with the good practice that unit co-ordinators should focus on "'assessing for learning' rather than simply 'assessing of learning'" because the former approach enhances student engagement (Krause, 2005, para 8).

Nature of the assessment task or stimuli

Gulikers, Bastiaens and Kirschner (2006) listed the nature of the assessment task as the leading dimension. Similarly, Frey and Schmidt (2007) posited that nature of the stimuli is a feature of authentic assessment. Without funnelling down what the nature of the authentic assessment is, this feature does not shed any light on authentic assessment or how it can be distinguished from artificial assessment.

Indicators of learning, validity, reliability and criteria

Herrington and Herrington (1998; 2006) identified indicators such as achieving multiple indicators of learning, validity, reliability, and using appropriate criteria. Similarly, Gulikers, Bastiaens and Kirschner (2006) posited that assessment criteria are an integral dimension of authentic assessment. However, these features are not compelling because they are not unique features of authentic assessment, but rather contentious issues for all types of assessment (Price, O'Donovan, Rust, & Carroll, 2008). As they do not pinpoint authentic assessment, these features are unconvincing. Consequently, Herrington and Herrington's (1998; 2006) student factors and task factors provide a greater insight into the notion of 'authentic assessment'.

To recap the literature above, some approaches attribute a different label to a feature of authentic assessment, but the essence of the feature remains the same, and as such they have been mapped against each other. A particular feature of authentic assessment is more compelling as the number of approaches emphasising it increases. Thus, according to the synthesis of the six approaches above, the most compelling features of authentic assessment, which have been supported by at least three of the approaches, are:

- Fidelity of task to the real world including conditions and a variety of resources (Approaches 1, 2, 3, 4, 5 and 6);
- Polished product valuable in its own right (Approaches 1, 2, 3, 4 and 6);
- Higher order thinking, reflection, meta-cognition and self-assessment (Approaches 1, 2, 5 and 6);
- Collaboration (Approaches 1, 2, 3 and 5);
- Requires students to make judgments, choices and determine tasks (Approaches 1, 2, 4 and 5);
- Complex task (Approaches 1, 2, 4 and 5);
- Wide, diverse and novel responses (Approaches 1, 2 and 6);
- Ill-structured, open-ended inquiry and construction of novel ideas (Approaches 1, 5 and 6); and
- Task is seamlessly integrated with assessment (Approaches 1, 2, and 6).

Authentic assessment framework

Some assessment tasks better reflect the workplace, or real world, and are thus more authentic than those that reflect the "staid classroom environment" (Keyser & Howell, 2008, p. 4). To aid unit co-ordinators to determine the degree of authenticity of an assessment task, this section develops an authentic assessment framework, which is based on the literature discussed in the first sections of this paper.

Recently, Burton (2009) argued that establishing the degree of authenticity of an assessment task should rely more heavily on task factors, in the sense of what students do, rather than the physical setting of the task (context). Consequently, irrespective of whether the assessment task is completed in the workplace, a mock workplace or a classroom environment is not totally conclusive of whether it is authentic, and unit co-ordinators should focus more on whether students are mimicking professionals.

When students engage in higher order thinking, meta-cognition, reflection and self-assessment, it is important that these features are not simply tacked onto the end of an assessment task, but are seamlessly integrated with the assessment. As a result, the questions presented in Table 1 combine these two features.

Three other features have been merged together in Table 1 below, that is, complex task; wide, diverse and novel responses; and ill-structured, open-ended inquiry and construction of novel ideas. The reason for considering these features together is that deciding whether a task is complex is subjective. It hinges upon the task being ill-structured and requiring open-ended inquiry, which both in turn, generate the construction of novel ideas as well as wide, diverse and novel responses.

These features have been simplified in a question to be asked by unit co-ordinators by focussing on the types of outcomes of authentic assessment.

Another feature of authentic assessment that is connected to outcomes is that the polished product needs to be valuable in its own right. The other questions to be asked by unit co-ordinators pertain to collaboration and whether the assessment task requires students make judgments, choices and determine sub-tasks.

Table 1: *Converting the compelling features of authentic assessment into questions*

Compelling features of authentic assessment	Questions to be asked by unit co-ordinators
Fidelity of task to the real world including conditions and a variety of resources	1. Is the student required to mimic a professional in the real world? 2. Is the student required to complete the assessment task using resources similar to that in the workplace? 3. Is the student required to complete the assessment task under realistic conditions?
Polished product valuable in its own right	4. Does the assessment task produce a valuable, polished product?
Higher order thinking, reflection, meta-cognition and self-assessment	5. Is higher order thinking or meta-cognition seamlessly integrated with the assessment task? 6. Is reflection seamlessly integrated with the assessment task? 7. Is self-assessment seamlessly integrated with the assessment task?
Task is seamlessly integrated with assessment	
Collaboration	8. Does the student collaborate with other stakeholders (for example, professionals/students) when completing the assessment task?
Requires students to make judgments, choices, and determine tasks	9. Does the student need to exercise judgment or choice in determining sub-tasks of the assessment task?
Complex task	10. Does the assessment task produce a novel or diverse responses?
Wide, diverse and novel responses	
Ill-structured, open-ended inquiry and construction of novel ideas	

The 10 questions in Table 1 may be answered with a Yes (Y) or No (N) response. Where a question is answered No, the authentic assessment framework illuminates to unit co-ordinators a strategy on how to enhance the authenticity of an assessment task.

Once a unit co-ordinator completes the framework, the overall result for an assessment task can be mapped onto a continuum with "artificial and decontextualised" at one end and "authentic and situated" at the other (Gulikers et al., 2006, p. 337). Table 1 develops 10 questions to be asked by unit co-ordinators, which makes it easier to calculate the degree of authenticity as a fraction. Further, the mapping process facilitates comparisons to be made between the authenticity of assessment tasks in a single unit or across units. Such comparisons will help to determine whether

authentic assessment is embedded in all levels of a degree rather than being left until the end, and whether authentic learning experiences later in the degree have been appropriately scaffolded. The authentic assessment framework will now be applied to three assessment tasks.

Applying the authentic assessment framework

As a case study example, the authentic assessment framework is here applied to *Fundamentals of Criminal Law*, which is a core semester subject in a law degree at the Queensland University of Technology. Some students complete this unit in their second year of a full-time law degree, whilst other students complete this in their first year of the law degree as a graduate. Fundamentals of Criminal Law offers a diverse range of summative assessment, that is, tutorial participation (TP, which is worth 20%), advocacy exercise (AE, which is worth 20%) and a problem-based exam (PBE, which is worth 60%), and as such provides a rich example to which the authentic assessment framework can be applied. Table 2 presents the application of the authentic assessment framework to this unit of study.

Each student engages in TP (tutorial participation) in three tutorials. Approximately 15-20 minutes is dedicated to TP, and a student demonstrates their ability to identify legal issues, knowledge of legal authorities, application of legal authorities, and make a conclusion. In addition to their oral contributions, a student demonstrates these problem-solving skills on a one-page written document that is submitted to the tutor. TP has a degree of authenticity of 3/10 (three of the 10 questions are answered yes).

A student is allotted 10 minutes in which to represent the prosecution or defence in an AE (advocacy exercise), that is, a mock criminal law scenario. The AE is completed in a moot court room in front of the tutor, who acts as a judge. The AE has a degree of authenticity of 6/10 (six of the 10 questions are answered yes).

The PBE (problem-based exam) contains four questions based on factual scenarios of about 1-3 pages each and require a student to provide written advice for the prosecution or defence. There is 30 minutes perusal time and two hours of working time. Thus, the PBE has a degree of authenticity of 3/10 (three of the 10 questions are answered yes).

Table 2: *Application of the authentic assessment framework to Fundamentals of Criminal Law*

Questions to be asked by unit co-ordinators	TP	AE	PBE
1. Is the student required to mimic a professional in the real world?	N	Y	N
2. Is the student required to complete the assessment task using resources similar to that in the workplace?	N	Y	N
3. Is the student required to complete the assessment task under realistic conditions?	N	Y	N
4. Does the assessment task produce a valuable, polished product?	N	N	N
5. Is higher order thinking or meta-cognition seamlessly integrated with the assessment task?	Y	Y	Y
6. Is reflection seamlessly integrated with the assessment task?	N	N	N
7. Is self-assessment seamlessly integrated with the assessment task?	N	N	N
8. Does the student collaborate with other stakeholders (for example, professionals/students) when completing the assessment task?	N	N	N

9. Does the student need to exercise judgment or choice in determining sub-tasks of the assessment task?	Y	Y	Y
10. Does the assessment task produce a novel or diverse responses?	Y	Y	Y
Total degree of authenticity = 12/30	3/10	6/10	3/10

Applying the authentic assessment framework to tutorial participation (TP), advocacy exercise (AE) and problem-based exam (PBE) resulted in 3/10, 6/10 and 3/10 respectively. This means that the advocacy exercise is the most authentic assessment task in this unit. The total degree of authenticity of the assessment tasks in *Fundamentals of Criminal Law* is 12/30, which seems reasonable considering that authentic assessment should be integrated throughout the degree and that the assessment tasks in this unit will help to scaffold those in later units. The authenticity of the assessment tasks in this unit could be improved, particularly by focusing on producing a valuable, polished product; seamlessly integrating reflection and self-assessment; and requiring collaboration. Requiring students to work in teams on a real criminal law case is an example of how to improve the authenticity of the assessment in this unit.

Conclusion

The authentic assessment framework described in this paper equips unit co-ordinators with a tool to help them determine the degree of authenticity of their summative assessment tasks, and make comparisons between them. Applying the authentic assessment framework to *Fundamentals of Criminal Law* indicated that the advocacy exercise was more authentic than tutorial participation and problem-based exam. Determining the total degree of authenticity of the assessment tasks in this course is the first step in making worthwhile comparisons with other units, so that authentic assessment is embedded across all levels of a course or degree to enhance student engagement and retention, and to ensure that later authentic assessment is appropriately scaffolded.

References

Australian Council for Educational Research. (2008). *Attracting, engaging and retaining: New conversations about learning: Australasian student engagement report Australasian survey of student engagement* . Retrieved March 10, 2010, from http://www.acer.edu.au/documents/AUSSE_ASER-Report.pdf

Boud, D., & Falchikov, N. (2007). *Rethinking assessment in higher education: Learning for the longer term.* New York: Routledge.

Burkill, S., Dunne, L., Filer, T. & Zanstra, R. (2009). Authentic voices: Collaborating with students in refining assessment practices. In *Assessment in Different Dimensions, Proceedings of the ATN Assessment Conference 2009, 19-20 November 2009* (pp. 84-93).

Burton, K. (2009). Does the summative assessment of real world learning using criterion-referenced assessment need to be discipline specific? In *Assessment in Different Dimensions, Proceedings of the ATN Assessment Conference 2009, 19-20 November 2009.* (pp. 94-103).

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

Gronlund, N. (2003). *Assessment of student achievement* (7th ed.). Needham Heights, NJ: Allyn & Bacon.

Gulikers, J., Bastiaens, T. & Kirschner, P. (2006). Authentic assessment, student teacher perceptions: The practical value of the five-dimensional framework. *Journal of Vocational Education and Training, 58,* 337-357.

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. & Herrington, A. (2006). Authentic conditions for authentic assessment: Aligning task and assessment, in *Critical Visions, Proceedings of the 29th Higher Education Research and Development Society of Australasia (HERDSA) Annual Conference, Western Australia, 10-12 July 2006.* (pp. 146-151).

Herrington, J., Oliver, R., & Reeves, T.C. (2002) Patterns of engagement in authentic online learning environments, in *Winds of Changing in the Sea of Learning, Proceedings of the 19th Annual Conference of the Australian Society for Computers in Tertiary Education (ASCILITE), Auckland, New Zealand, 8-11 December 2002.* (pp. 279-286).

James, R., McInnis, C. & Devlin, M. (2002). *Assessing learning in Australian universities.* Retrieved March 1, 2010, from http://www.cshe.unimelb.edu.au/assessinglearning

Keyser, S. & Howell, S. (2008). *The state of authentic assessment.* Education Resources Information Center (ERIC) database. (ERIC Document No ED503679) (10 March 2010).

Krause, K. (2005). *Good practice guide: Enhancing student engagement in the year: 10 strategies to success.* Retrieved March 1, 2010, from http://www.griffith.edu.au/gihe/pdf/gihe_tipsheet_web_ese.pdf

McNamara, J., Field, R., & Brown, C. (2009). Learning to reflect in the first year of legal education: The Key to Surviving Legal Education and Legal Practice, *Conference Proceedings of the First Year in Higher Education Conference, 29 June-1 July 2009.* (pp. 1-10).

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

Price, M., O'Donovan, B., Rust, C. & Carroll, J. (2008). Assessment standards: A manifesto for change. *Brookes eJournal of Learning and Teaching*, *2*(3). Retrieved March 10, 2010, from http://bejlt.brookes.ac.uk/article/assessment_standards_a_manifesto_for_change

QUT Office of Teaching Quality. (2009). *Protocols: First year experience*. Retrieved March 10, 2010, from http://www.otq.qut.edu.au/about/policyandpro/guidelines/documents/protocols_fye.pdf

West, D., & Heath, D. (2009). A learning community approach: Embedding authentic assessment in social work curriculum, in *The Student Experience, Proceedings of the 32nd Higher Education Research and Development Society of Australasia (HERDSA) Annual Conference, Darwin, 6-9 July 2009.* (pp. 463-472).