Examining increased flexibility in assessment formats

Brian Irwin and Stuart Hepplestone

Quality Enhancement and Student Success, Sheffield Hallam University, Sheffield, UK

There have been calls in the literature for changes to assessment practices in higher education, to increase flexibility and give learners more control over the assessment process (Boud and Falchikov 2006; Nicol and MacFarlane-Dick 2006; Taras 2002). This article explores the possibilities of allowing student choice in the format used to present their work, as a starting point for changing assessment, based on recent studies and current examples of flexible assessment practice in Higher Education. The benefits of this flexible assessment format approach are highlighted, along with a discussion of classic assessment considerations such as validity, reliability and marking concerns. The role of technology in facilitating assessment method choice is considered, in terms of new opportunities for providing student choice in the way they evidence their learning and present their work. Considerations for implementing flexible assessment choices into the curriculum are presented, along with a call that further research into such practice is needed to develop a comprehensive set of practical recommendations and best practice for implementation of flexible assessment choice into the curriculum. The article should be of interest to curriculum developers and academics considering implementing changes to the assessment process to increase student ownership and control.

Keywords: flexibility, assessment, choice, format, student-centred

Introduction

During the last decade there have been calls for more student ownership and flexibility in the assessment process (Boud and Falchikov 2006; Nicol and MacFarlane-Dick 2006; Taras 2002), from setting criteria against which to assess to using feedback to feed forward into
future assessments and careers. However, the reality of today's assessment practices is far from this proposed ideal (Crook, Gross, and Dymott 2006; Knight 2002). MacLellan (2001) reports that students do not exploit assessment to improve their learning nor view feedback on their work as helpful, believing criteria to be implicit and unclear. We are proposing that to bridge this gap between current assessment practice and the ideal world is to allow students to have a choice in the format used to present their work.

This paper will:

- Discuss the concept of allowing such choice using online assessment formats
- Discuss key issues around validity, reliability and skill development
- Propose considerations for implementing increasing flexibility in assessment and empowering students

To inform the discussion of choice in flexible assessment formats, we performed a partial literature search on eight leading journals concerned with innovations in learning, teaching and assessment, particularly in higher education, using the terms 'flexible', 'flexibility' or 'choice' and 'assessment'. From the matches, titles and abstracts used to identify relevant papers on the topic to be read in full. Google Scholar was also used with the same search terms to highlight the most relevant matches outside of these core journals, such as other journal articles, conference papers and proceedings.

**Background**

Nicol and MacFarlane-Dick's principles of good feedback (2006) propose a shift in the thinking of higher education assessment practices towards increased student control of assessment. The Re-Engineering Assessment Practices in Scottish Education (REAP) project
echoes the call for student empowerment within its principles of good formative assessment and feedback (Nicol 2007). Incorporating flexibility in the assessment process 'is one of the principles that underpins good assessment practice' (Clayton and Booth, 2000). MacLellan (2001, 311) agrees that alternative assessment has a place in formative assessment, but strongly opposes its use where assessment is based on making judgements: 'marking consistency in alternative assessment can make comparability of performance difficult to effect, thereby leaving alternative assessment a less than convincing form for use in higher education'.

Flexible assessment involves some element of choice on the part of the student (Australian National Training Authority 2002; Martin 2006; Wood and Smith 1999). Some of the ways this has been implemented include allowing individual students to decide the weightings applied to each assessment task (Jamieson 2005; Wood and Smith 1999), and the use of scored rubrics alongside self- and peer-assessment to offer choice in the method, subject, engagement with criteria and result (Ellis and Folley 2009b). However for this paper, student choice in assessment explores the idea of equivalent ways of demonstrating learning outcomes (McClenaghan 2006), where students can 'use different methods to show their understanding' (Hanafin et al. 2007, 444). Similar is Boud's (1995, 5) observation that 'what is important are learning outcomes, no matter how they [are] achieved'.

If the assessment criteria are clear about the desired learning outcomes, then students could use a variety of formats to meet those outcomes. For instance, with a traditional assessment format such as the essay, the underlying purpose is for students to demonstrate a coherent, well-structured critical argument. A student may able to accomplish this task using a web page, reflective blog or video presentation. For many students, alternative formats may
prove a less difficult and more suitable way to evidence their learning (Hall, 1982), particularly for students with disabilities (Konur, 2007), and goes some way to enhancing student equity (Martin, 2006).

Flexible assessment aligns with a competency-based approach to learning (e.g. workplace assessment, self- and peer-assessment, online assessment), and brings with it authenticity into the assessment process (McLoughlin and Luca 2001). The use of flexible assessment strategies can help prepare students for life after university, for example, in subjects with less defined career paths, students may choose a format that has transferability for their future careers. Sambell, McDowell, and Brown (1997) found that students are more engaged with learning in their assessments when they can see the future transferability of the assessment, suggesting that use of choice in assessment could also improve motivation and learning. In addition, Asafu-Adjaye (2001) identified that by providing flexible assessment regimes which calculate a student's final grade based their best scores from a range of compulsory and optional tasks can result in a decrease in failure rates, by forfeiting difficult tasks to avoid losing marks or having the opportunity to gain a high mark.

A complete theoretical and cultural change in assessment practices will be a difficult task, and how to accomplish this is often not discussed in the literature (Craddock and Mathias 2009). The introduction of some student choice into the assessment mode could be a starting point for this change, as research on student readiness for flexible assessment found that even first-year students were highly receptive to this idea, especially in comparison to greater responsibilities such as designing tasks and defining criteria (Francis 2008).
Research into flexibility in online assessment

Previous research into the role of online assessment has focused on the flexibility it provides for learning at a distance and the flexibility it provides for students to take automatically-marked self-tests at various times (Phillips and Lowe 2003). Flexibility has also involved giving students choice of on-campus or at-a-distant assessment, regardless of their enrolment status (Veenendaal 2001). Phillips and Lowe (2003) looked at mapping traditional forms of assessment into online equivalents, but they suggest tutors change their format to online to increase flexible access rather than allowing students to choose the format that best fits their needs. Similarly, Kehoe, Tennent, and Windeknecht (2004) gave students assessment method choice but it involved tutors creating multiple online assessments rather than giving students control for one assessment. This method resulted in an increased workload for tutors, and less engagement with criteria by students, as the tutors had the responsibility for aligning the assessment method with the criteria instead.

Assessment format choice between essays and online multiple-choice tests has been explored by Scouller (1998) and Cook (2001). However, both courses used a flexible mark weighting system whereby students could choose to engage with only summative assessments and ignore any formative assessment, meaning the assessment components designed to specifically promote learning would not be done by all students. Hall (1982) looked at student choice in assessment format and found student preference varies widely depending on their individual learning strategies and creativity in demonstrating their ability in the subject, and therefore simply providing a limited choice between an essay and an online test is unlikely to allow every student the opportunity to demonstrate their learning in their preferred way.
Discussion

Building flexibility into the process of choosing assessment methods can provide students with more control over their assessments, and thus their learning experience. However, there are certain areas that the literature has generally defined as central concerns for assessment practices such as validity, reliability, marking and the contribution of assessment towards learning, both intended and unintended. In this discussion section we will also explore the role of technology in enabling this flexibility and choice of formats.

Validity

Validity can be defined as ‘that assessment measures what it is designed to measure’ (Russell et al. 2006, 466), or in other words that the assessment is an accurate measurement of the learning. Russell et al. (2006) point out that validity has been widely discussed in assessment literature, especially in relation to positivist-oriented assessment: that student performances in an assessment should match an ideal response. However some research suggests that some assessments which are accepted as being valid, such as essays and exams, often measure the student’s ability to engage with a particular form of assessment and assessment constraints (such as time limits), rather than the learning in a more general way. For instance traditional assessment formats such as essays present barriers to some students, such as those with disabilities, due to having to deal with large amounts of text (Hanafin et al. 2007). Hanafin et al. (2007) found that the large amount of text can prevent some students from being able to demonstrate their learning effectively, even where those students were given additional time due to their learning contracts. Hanafin et al. (2007) also determined that unseen exams can also disadvantage students with disabilities due having to deal with text in a high pressure, time sensitive context. Sambell, McDowell, and Brown
(1997) makes a further point that unseen exams are perceived by students as assessing student memory, stress management skills and luck more than the learning outcomes.

These examples suggest that limiting the forms an assessment may take actually decreases its validity, as those with a disadvantage for that form of assessment would produce an assessment product that does not measure their true learning; and those with a natural disposition to or advantage towards that particular assessment form could over-perform compared to their actual learning. Therefore providing choice in the assessment format students could use would lead to the strongest validity for an assessment, as students would not be at a disadvantage due to personal characteristics not associated with the learning itself. However, MacLellan (2004) cautions that alternative assessment formats will not be valid in assessment contexts where there is less flexibility about the performance measurements needed, e.g. a teaching observation session for a trainee teacher. With this in mind, it makes sense to ensure that the flexibility being added to a specific assessment increases its validity rather than decreases it.

**Consequential Validity**

Related to validity is the concept of consequential validity, which is not concerned with measuring learning but instead the effect of the assessment on the learning process in general, including unintended effects (Boud 1995). Some research, such as Hall (1982) and Sambell, McDowell, and Brown (1997), has already looked at the potential consequential validity of using different assessment formats. In a study of the behaviour of one class, Hall (1982) found that when offered a choice of assessment modes, students generally chose a mixture of formats they were comfortable with and some elements of challenge in mode choice. However when there were tight deadlines to meet, the students defaulted to
assessment modes they were already skilled in. This suggests that the consequential validity of flexible assessment choice is high, due to the positive effect on learning to master new modes, however these potential benefits are dependent upon the overall design and time pressures of the assessment. In another study, Sambell, McDowell, and Brown (1997) found that students had established surface learning strategies for tackling traditional assessment formats such as exams and essays, but responded more deeply to innovative assessment formats and assessments where they could see the life-long applicability of their assessment. In that study the student engagement with alternative assessment formats led to additional motivation and learning, which points to a high consequential validity for non-traditional assessment formats. Struyven, Dochy, and Janssens (2005) also found the use of non-traditional assessment formats results in students applying a deep learning strategy. However Sambell, McDowell, and Brown (1997) and Struyven, Dochy, and Janssens (2005) did not explore the concept of choice with assessment formats, instead focusing on the alternative assessment formats themselves.

As both of these examples are limited in the number of students involved and more current research into the consequential validity of flexibility in format choice is needed to answer some further questions. Would students rely on a quick, surface approach to learning if they are already familiar with the format? Or does flexible assessment encourage a deeper approach as not struggling with the format would free up more time and allow them to engage with the learning and content in a deeper way? Will students take on the challenge of mastering new assessment formats, thus improving their learning outside of the core curriculum? Further research is needed to answer these questions and understand the
potential positive or negative effects on the general learning process when using flexible assessment formats.

Reliability

Reliability, or the repeatability of an assessment and its result, is often expressed in the literature as a concern about different assessments (Boud 1995). Yorke (2003) points out that this concern stems primarily from the summative assessment purposes, as this helps compare across student performances and creates a sense of equity and fairness to an assessment. It may be helpful to examine reliability in terms of two parts: the repeatability of an assessment for an individual and repeatability across a cohort. Concerns around the repeatability or consistency of marking between tutors are discussed in a later section on marking concerns.

Reliability for an individual

Reliability for an individual can be seen in the context of if the student performed the assessment again, would they produce a similar assessment performance. For example, an online test where a student is given different questions each time they take it is less reliable than a fixed piece of work such as an essay. Additional options and flexibility worsens the repeatability of the assessment, as students could choose different options for the assessment. However, it is worth asking if this is actually problematic. It seems like an issue only when there is an assumption of one correct way of demonstrating learning. Inclusion literature argues that such an assumption excludes students that cannot demonstrate their learning fully in that one way (Hanafin et al. 2007; Hockings et al. 2008). For instance if a student has dyslexia then they may have more difficulty expressing their learning in an essay than by through a video. Hanafin et al. (2007) suggests that assessments which are identical
for all students are inherently less fair and equitable, rather than more, meaning that assessments should measure learning outcomes and be adjusted to the needs of individual students (McClenaghan 2006). In this way, individual repeatability of the exact same assessment product is less important than repeatability of demonstrating the learning outcomes in the assessment.

**Reliability across a cohort**

Giving students flexibility in their assessment choices sometimes produces concerns that different students cannot be compared to each other. However, in criterion-referenced marking schemes, this is inconsequential as how students compare to the criteria is more important than how they compare to each other. Only in norm-referenced systems does reliability between students need to be considered, and higher education has moved towards criterion-referenced assessments (Boud 1995; Knight 2001). Recent research into staff marking practices suggests that tutors may not actually use the assessment criteria when marking, even when criteria are claimed to be the basis for marks (Bloxham, Ashworth, and Boyd 2009). If this is the case, flexible assessment methods could actually help move tutors away from norm-referencing tendencies when using criterion-referenced marking, as it becomes more difficult to compare student work:

> It was observed that the assessment process was fair and equitable as each assessment is explicitly measured against the performance criteria and learning outcomes; it is difficult to fall into the trap of measuring student against student in flexible assessment. (McCurdy 2000)

However, employers may want to compare between students, which is inhibited by flexibility (Lindsay 2007), so questions remain about if employers would see flexibility as lax standards or favouritism towards students, as has been noted is the attitude of some academics towards flexibility for students with disabilities (Hanafin et al. 2007; Konur 2007).
Determining if an assessment task is primarily for learning or evaluation and comparison may help determine how important the issue of employer concerns will be, and whether or not assessment method flexibility would be appropriate. There is a larger debate which could be had about the amount of importance to give employer concerns when designing assessment tasks as well. For instance, what is the role of higher education in developing the employability skills that business want? Are universities primarily vehicles for economic improvement? These questions are touched on by Knight (2002) in his exploration of the purposes of summative assessment in relation to employers.

**Marking concerns**

There is always a concern about the reliability of marking and its repeatability: would a mark be similar between different tutors marking the same piece of work? These concerns exist because not just because of a desire to compare students, but because students and tutors want a sense of fairness in the assessment process (Sambell, McDowell, and Brown 1997): therefore students should not be penalised or rewarded as a result of luck in who grades their work. This is a potential concern with the idea of student choice in assessment formats, as different markers are likely to have different experiences and attitudes towards a variety of formats, including unfamiliarity with various formats, personal preferences for particular ones and historical assumptions about the effectiveness or appropriateness of each format for academic work. If an academic has a strong belief that essays are the only way to display a critical approach to the literature, then this bias could affect the marking of student work in other formats. This highlights the importance of tutor discussion around giving choice to students in assessment, to unearth staff perceptions and biases about assessment formats (Ellis and Foley 2009a). Errington (2004) points out that the success of innovations in
teaching and learning in general are largely dependent upon teachers’ personal theories and past experiences.

One approach to addressing marking concerns could be requiring students to evidence how their work meets the assessment criteria. This would help in the situation where a tutor lacks understanding about or experience with a particular format. This has the added benefit of getting students to engage more deeply with the criteria as part of the assessment, a base requirement for good assessment practices (Nicol and MacFarlane-Dick 2006). It could also serve as the start of a dialogue around the assessment processes and any resulting feedback, another principle for encouraging self-regulation in learners (Nicol and MacFarlane-Dick 2006).

The potential for an increase in tutor marking time is also a consideration, though there was no evidence found in the literature about the effect of choice in assessment methods on tutor marking load. Any difference in marking time is likely to depend on the tutor’s personal preferences as well as any disabilities they may have. For instance, a dyslexic tutor may find marking multimedia assessments actually faster than marking traditional essays. Some tutors may read faster than others or have established routines for marking papers and thus find marking text-based materials easier.

Tutors can also consider the possibility of involving students in the marking and feedback process using peer assessment. Peer assessment has been shown to benefit students by encouraging them to engage with assessment criteria and evaluation skills (Gibbs and Simpson 2004). When using flexible assessment formats, students would also be exposed to new approaches and different perspectives on addressing the assessment
criteria, which could help them better understand how to demonstrate their learning and meet criteria.

However, Hounsell (2008) argues that peer marking is problematic when compared to peer feedback, which suggests that one solution would be for tutors to just be involved in marking, while the feedback came from peers. It is unclear if this would decrease marking time for tutors, but students would receive a larger amount of feedback to work with, consistent with Gibbs and Simpson’s (2004) call for students to receive sufficient feedback and practice making informed judgments about work (Boud and Falchikov 2007).

**Developing skills**

One important consideration is the role of assessment in promoting sustainability, preparing students for future needs and life-long learning (Boud and Falchikov 2007). In looking at the sustainability of flexible assessment formats, it is important to consider that some skills, such as writing, are essential for life-long success. If students are given control over format choices for all assessments, they may choose the same format each time and play just to existing strengths as Hall (1982) found, resulting in other important skills not developing. In the extreme this would lead to a student who avoids written formats entirely and leaves university unable to write coherently. However this scenario could be avoided if flexible assessment formats are seen as an important part of a varied assessment diet, rather than a single solution for all situations. One option is to promote format choice on one assessment while the others remain tutor-controlled to ensure key skills are learned (Tal 2005).

The need to develop particular skills also highlights the need for clear learning outcomes that identify the important components in assessment tasks. For instance, if the
development of writing skills is a critical learning outcome for an assessment task, that would help determine the amount of flexibility available for students. Through careful analysis of the underlying purposes behind the assessment tasks, opportunities for flexibility can be identified.

Format choice has the potential to help students meet life-long employment needs by engaging in more authentic assessment activities for their context (McCurdy 2000). However, students may make assumptions about employment based on stereotypes or out-dated conceptions, or just lack an awareness of what skills are needed for employment (Hall 1982). Therefore, guidance will be needed to help bridge the gap and ensure students have realistic and up-to-date knowledge about long-term skills needs as considered and recommended by the Leitch Review of Skills (Leitch 2006). Students may in general need additional guidance and scaffolding to be prepared to select from a range of assessment formats, particularly if their past experiences consist of less flexible assessments. Boud (1995) notes that students' past experiences of assessment have an effect on how they perceive their role in learning, meaning students may conceive of their role not as an active participant, but as a passive recipient. Constructing an identity as an active learner is seen as the first step in students' developing informed judgement (Boud and Falchikov 2007), and is crucial to adopting an approach like this that requires students to make, justify, and evaluate their choices – the third step in Boud and Falchikov's scheme for promoting sustainable assessment.

The role of technology

As mentioned previously there has been some research into the role of online assessment with regards to flexibility in terms of student convenience in taking the assessment. However, this research has not focused on the ability of online technologies to allow student choice in
assessment methods. Online assessment offers many benefits for choice including new formats such as blogs and wikis, improved ease of use for tools, and accessibility advantages. However, the use of such formats raises some considerations in terms of choosing and supporting those tools.

**Online formats**

In looking at online formats, an obvious first consideration is web pages. Due to their ability to link between various pages, they have been used as an assessment format which demonstrates the interrelationship between concepts (Landow 1997). More recently the concept of Web 2.0 has seen an explosion of new tools built on top of the web, many of which focus on collaboration as a key goal (O'Reilly 2005). Wikis and blogs are two such tools that have seen increasing use as higher education assessment formats (Alexander 2006). Alexander suggests the number of online tools will increase in the future, providing both an opportunity and a challenge in terms of additional choices to select an appropriate format from. The existence of new formats and tools allows more choice for students, meaning they have more options for choosing the mode that best showcases their learning.

Most online applications, such as web 2.0 tools, are also designed to be easier to use, with simple, rich interfaces (O'Reilly 2005). These improvements in interfaces mean there will be less technical complications and limitations for students who want to use more complicated assessment formats such as audio and video. For example, using YouTube to upload a video is considerably easier than recording a video using a camcorder and then mailing cassette tapes to everyone who needs to see it. The existence of these easy to use tools which allow simple distribution of the final assessment product facilitates the use of flexible assessment formats in the curriculum. Many web 2.0 tools such as blogs and wikis
also allow monitoring of the process of creation as well as the viewing of the finished product, allowing assessors access to new information about the assessment that could be used in the marking process (Millwood and Terrell 2005).

Accessibility options are also improving with technology, allowing for individual users to determine colours, text size, and use assistive technology to improve their ability to engage with an assessment format. This means students and staff should have access to many different formats, regardless of any disability they may have, which results in increased choice for them. Developing technologies such as text-to-speech and speech-to-text offer further scope to help all learners engage with different assessment formats.

These changes show online assessment method flexibility is a viable option for contemporary assessment practice, as there are increased options for students to choose from and decreased barriers to using those options. However, the only research identified about the use of online components to achieve flexible assessment formats was McCurdy (2000), a case study on the use of flexible assessment methods, including online components. The study is limited, however, as its findings reflect only the tutor's perspective and do not cover modern web tools due to when it was published. More research is needed into the effectiveness of online tools and new technologies to facilitate student choice in assessment format, including student perceptions and experiences of using different online tools to accomplish this.

**Helping students with technology**

While the addition of new technologies into the range of assessment processes increases opportunities for using different formats, it also raises questions of the amount of support needed for students and tutors using different technologies. For instance, if a student decides
to create a video will the university have the support available in educational guidance departments as it does for essay writing skills? Does the university have a role in approving which technologies that are fit for student use, as some Web 2.0 companies will go out of business? These questions are particularly important in light of increasing numbers of available Internet technologies. Thankfully many of the most successful web-based technologies are those that are most intuitive and easy to use. Many of them contain good self-help support, which reduces the need for university-owned support materials and staff. However, the university is still likely to have a role in helping students select technologies that are both fit for purpose and reliable. Students will know of some of the most popular tools but may not be aware of other options available to them. They may also need guidance in how to duplicate work offline, in case a web technology company goes bankrupt, taking valuable assessment information with them.

**Implementing Flexible Assessment Formats**

Staff and students, both of whom, have existing perceptions of and experiences with assessment, may be resistant to complete change, but open to the smaller changes in power, responsibility and maturity that some choice allows. Wood and Smith (1999, 233) warn that flexible assessment ‘must be introduced with planning, organisation and careful consideration of the objectives of the learning’, and suggest that greater flexibility can be implemented as students progress through their studies. Providing students choice in their format they use to produce their assessment is a good starting point for introducing wider flexibility into the assessment process. Based on a consolidation of points made by practitioners who have experimented with flexible assessment methods (Hall 1982; Sambell, McDowell, and Brown
1997; McCurdy 2000; Ellis and Folley 2009a) and points raised in this paper's discussion, the following are recommendations for implementing flexible assessment formats:

- Determine the degree of flexibility in format choice to build into assessments across the course
- Refine assessment criteria to ensure they are directed at demonstrating the desired learning outcomes instead of incidental learning outcomes that are specifically linked to the assessment format used previously
- Have an open and honest enquiry into any preconceptions of the worth of different assessment formats
- Agree marking strategies ahead of time
- Consider the role of technology in supporting student use of different assessment formats
- Ensure support for staff and students with any assessment formats that will be allowed
- Discuss with students why you are introducing assessment format choice and what the implications are for them

**Conclusion**

The use of flexibility in assessment formats supports core agendas in higher education such as accessibility and promoting autonomous learners, and has been called for by the literature in these fields. Flexible assessment modes can be seen as a first step towards a more student-led pedagogy, while increasing student engagement in the assessment process and setting criteria. Changes in online technologies have made it easier and more beneficial to use online assessment techniques as a mechanism to achieve this flexibility.

However, there are potential obstacles in terms of stakeholder attitudes, new marking practices and ensuring the development of well-rounded students. Further research, particularly based on case studies and real practice examples, is needed to determine the
relevance and importance of these potential barriers and to develop possible solutions for them. It is also worth examining the effect on students that format choice produces, in terms of approaches to learning taken, trends with choosing formats and impacts on student performance, particularly for students with disabilities. Interviews of staff and students would yield further insights into the advantages and disadvantages of this approach, as well as practical considerations when implementing it. A more comprehensive set of good practice guidelines for implementing flexible assessment choice format can be created from this research to ensure effective change in the assessment process.

Brian Irwin is a Principal Lecturer in Academic Innovation: e-learning at Sheffield Hallam University. His other research interests include realising the transformative potential of e-learning and student-centred pedagogies.

Stuart Hepplestone is a Senior Lecturer in Curriculum Innovation at Sheffield Hallam University. He specialises in the use of online assessment and feedback to enhance the student experience.

References


Bloxham, S., M. Ashworth, and P. Boyd. 2009. Investigating the mysteries of marking: the implications of staff marking practices for academic development. Paper presented at the 17th Improving Student Learning Symposium, September 7-9, in Imperial College London, UK.


Adelaide, Australia: Australasian Society for Computers in Learning in Tertiary Education.


