

The Assessment Wheel: facilitating a collegial response to the inherent tensions and trade-offs in assessment design for Higher Education

Lisa Day
London Metropolitan Business School
London Metropolitan University

Keywords: *assessment, communities of practice, assessment principles*

Introduction

This paper offers a fresh perspective on the bewildering business of assessment design. The Assessment Wheel is a tool to help practitioners discuss and debate assessment. It makes subtle and conflicting trade-offs and tensions in assessment design more explicit and discussable. The Wheel is firmly grounded in the literature on assessment in Higher Education. It is hoped that it might be adopted more widely as a teaching tool for new practitioners in Higher Education and to facilitate local communities of practice in developing and reviewing assessments.

Frameworks for Designing an Effective Assessment

Brown (2001) offers a model for designing assessment that aligns the assessment method with aims, intended learning outcomes and methods of learning. Brown's model provides a useful starting point but might be viewed as an example of rational curriculum planning (Knight, 2001) in taking a relatively sequential approach. Whilst linear models appeal to our natural desire for order, in reality, assessment design is part of the messy business of course design, which tends to be holistic, rather than sequential. Based on a rational model, the idea that there might be 'one best' or 'right' form of assessment seems plausible. In reality, assessment choice is often complex and contentious. Ecclestone (2003:96) refers to the broader, 'problematic context of political and social tensions, different goals for education in terms of motivation and autonomy, and competing ideas about knowledge as the basis for learning'.

The Messy Business of Assessment Design

Ecclestone and Swann (1999:379) argue that lecturers are, 'under great pressure to provide high-quality marking and feedback despite an increase in the numbers of students to assess'. Bloxham and Boyd (2007) identify pressures, such as resource

constraints, modularisation, non-traditional students and issues with increased plagiarism. Ecclestone (2003:85) refers to, 'underlying tensions and dilemmas in assessment'. There are often underlying trade-offs to be made in assessment and the need to recognise that you simply cannot satisfy all of the people all of the time. There are no easy answers and assessment choice will depend very much on the particular context. However, there is a need to be able to discuss assessment methods so that some sort of agreement can be reached. Jenkins (1998) circular Ouija Board construct is a model that presents curriculum design powerfully and visually as a series of forces that pull the designer in different directions. However, the search for a similar construct that focused on assessment was less fruitful.

The Assessment Wheel

A review of academic literature on assessment in Higher Education highlighted a range of factors that might be described as principles of good assessment (for example, Brown et al., 1997; Toohey 1999; Bloxham and Boyd, 2007; Ecclestone, 2003; Gibbs, 1995). Common themes are validity, reliability and fairness. Toohey (1999) also mentions support for learning and cost. Bloxham and Boyd (2007) cite comparability, consistency, transparency, practicability, equity and attribution. Gibbs (1995) identifies the issue of acceptability. These different themes have been summarised into eight main categories: fairness; reliability; validity; practicality; support for learning; fit; acceptability and purpose. Each of these was derived from the literature but, in some cases, this involved grouping a number of points to establish a meaningful but manageable set of ideas.

The eight categories, or principles, are presented in Figure I as a circular framework, which is designed to emphasise the holistic nature of assessment choice, in contrast to more traditional linear models. Cummings and Wilson (2003:413) argue that a wheel indicates that all elements or perspectives are equally valid, and that people should work around the wheel and gather ideas to feed into decision making. The wheel might be used at any point in the module or assessment design process, and in any way that practitioners find useful, both in groups and as individuals. O'Brien and Dyson (2007:22-23) argue that models are, 'a deliberate simplification that nevertheless aims to capture the essence of what...is going on' and can be used as a diagnostics tool, to fuel debate, or to help practitioners develop and rehearse their ideas.

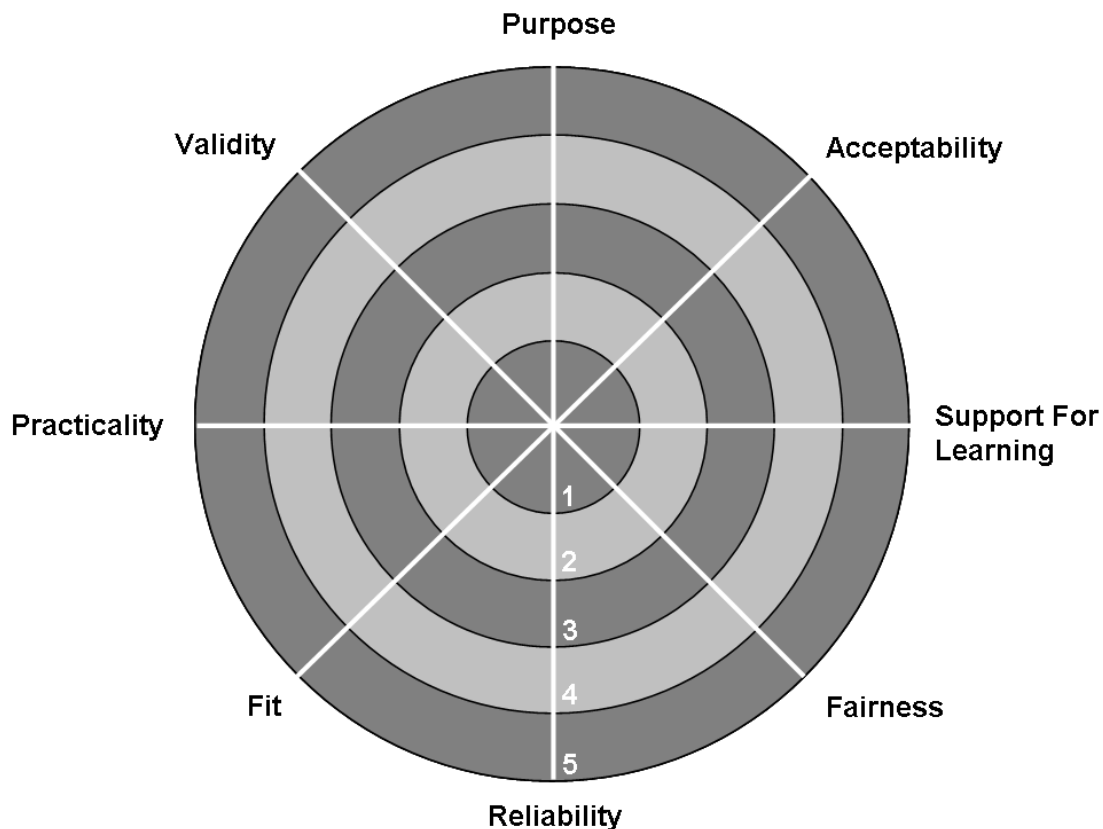


Figure 1 The Assessment Wheel

Using the Assessment Wheel

The wheel might be applied, like a 'spidergram', to consider which elements are most important for a particular situation, or which elements are strong or weak for a current or proposed assessment. The Wheel highlights the dilemmas in assessment choice but is synthesized in such a way that it can help to make tensions more explicit and therefore discussable. It might also discourage circular debate between colleagues who are implicitly giving different weight to one or more aspect. It is important to recognise that not every element can be rated high and informed compromises are needed. The Wheel helps people to articulate the underlying rationale for their assessment and reach an informed agreement.

Bloxham and Boyd (2007:225) argue for, 'professional learning ...within communities of practice' but note that the, 'competitive element of academic practice and the level of institutional and external accountability that exist' are often a barrier to this. So, a particular objective of the framework is to facilitate local, collegial, debate around assessment within a particular 'community of practice' (Lave and Wenger, 1991:98). The model has been tested through a pilot study which considered a MBA Strategic Management module and on a Learning Technologies module. The details of the pilots are outside the scope of this particular paper.

The Eight Principles in more Depth

The Principles have been summarised and some key questions developed to provide some clarity around the inevitably fuzzy boundaries of the different categories. The list of questions is by no means exhaustive and experienced practitioners will doubtless wish to add to or adapt them.

Validity

A valid assessment is fit for purpose and enables students to develop and demonstrate the module learning outcomes. Ecclestone (2003:37) mentions 'authenticity', which is closely related and refers to whether the assessment has credibility with employers, and reflects the 'real-life' situation it is designed to prepare students for. There may be great debate as to what constitutes an 'authentic' assessment and the balance between developing workplace and academic knowledge and ability.

Example Questions for Validity

- Does the assessment enable students to demonstrate their competence with regard to the learning outcomes of the module?
- Is the assessment 'authentic'? Does it reflect the sort of activities that a practitioner might be expected to perform in the workplace?

Practicality

Toohey (1999:181) identifies cost as a key consideration and Bloxham and Boyd (2007) highlight 'practicability'. Practicality concerns the feasibility of the assessment, such as marking load and general ease of implementation.

- Are there any issues concerning cost, time or access to resources that make this assessment impractical for students, lecturers or any other party?
- Could any aspect of the assessment be made more efficient or effective through the use of e-technology?

Support for Learning

Support for Learning (Toohey, 1999:180) concerns clarity of the pedagogical purpose of the assessment and is broader than just meeting the learning outcomes. For example, a purpose of the assessment might be to encourage life-long learning (Boud and Falchikov, 2007), or facilitate early formative feedback, or develop students' self-assessment skills through peer review (Bloxham and Boyd, 2007:15).

- Does the assessment facilitate appropriate formative support and feedback?
- Does the assessment facilitate student engagement and lasting learning?

Fairness

Bloxham and Boyd (2007:41) mention 'equity' and ensuring any reasonable adjustments are made to ensure the assessment is appropriate for students' special educational needs or disabilities. Fairness might also include encouraging academic honesty or, if group work is involved, considering issues of participation by all members. It might also include 'transparency' and ensuring the assessment criteria are clear to students.

- Does the assessment enable reasonable adjustments to be made for students with special educational needs or disabilities?
- Does the assessment encourage academic honesty and manage group work fairly?

Reliability

Ecclestone (2003:111) defines reliability as, 'assessment that is designed to ensure that the same range of results gained by learners could be reproduced in a different cohort of learners who are deemed to have similar abilities' and states that it, 'aims to enable assessors to standardise their judgements against those of other assessors'. Reliability refers to the ability to mark consistently and might include the provision of marking schemes. Multiple choice tests are an example of an assessment which is considered to have high reliability.

- Is it possible to mark the assessment reasonably accurately and consistently? Is there a process for moderating and reviewing marker consistency?
- Is it a 'high stakes' assessment that has a considerable impact on the final award classification? (in which case reliability is likely to be of increased importance)

Fit

Fit refers to how the assessment fits with wider course objectives. This includes the timing of assessments and the variety of assessment methods across a programme. Bloxham and Boyd (2007:166) argue that 'a wide range of assessment methods can promote equity and...an inclusive assessment culture' but too much variety can prevent students from becoming familiar with a form of assessment and from developing their proficiency. This category also captures whether the assessment fits with formal internal and external quality guidelines and requirements.

- Does the assessment enable students to experience an appropriate range of different assessment methods across the programme?
- Does the assessment fit the appropriate internal and external formal guidelines, such as Level Descriptors, QAA Guidelines?

Acceptability

Acceptability concerns the cultural norms and tacit and explicit expectations of key stakeholders, such as students, institution, lecturers, subject-discipline and

professional bodies or future employers. Bloxham and Boyd (2007:164) note that 'assessment methods have long traditions within different disciplines'. Gibbs (2006:20) notes that 'students are also often conservative...they are instinctively wary of approaches with which they are not familiar or that might be more demanding'. Levels of conservatism may well be higher for 'high stakes' assessment where the assessment has a significant impact on the students' future or overall course grade (Lines, 2005:149). Considering acceptability does not mean that novel assessments should be avoided or discouraged. The Assessment Wheel might be used to support innovative assessment by demonstrating that, whilst it may be lower on acceptability, this is mitigated by improved validity or support for learning.

- Will the assessment be acceptable to all stakeholders? What is their appetite for conservatism or innovation?
- How is this type of activity assessed by other institutions?

Purpose

Gibbs (1995:8) argues that 'if a number of tutors work in parallel to implement a course it is crucial that they understand and are committed to the same rationale'. Gipps (1994:3) states that, 'the first question to be asked...when considering the form of assessment to be used is 'what is the assessment for?'. He argues against simply selecting the 'best' from a menu of assessment options, and states, 'we need to resist the tendency to think in simplistic terms about one particular form of assessment being better than another: consideration of form without consideration of purpose is wasted effort' (Gipps, 1994:15). Finally, Earl (2003:13) states that 'it is not possible to use one assessment process for the many purposes that we want it to fulfil. Different purposes require vastly different approaches, and mixing the purposes is likely to ensure that none of them will be well served'.

- Is there a common understanding of the purpose and objectives of the module?
- Is there a common understanding of the purpose of the assessment(s)? Is there agreement as to the learning outcomes that should be demonstrated?

Conclusion

The Assessment Wheel is a tool to help practitioners debate and design efficient and effective assessments. In contrast to more traditional linear models, the Wheel encourages practitioners to articulate and grapple with the complex tensions and trade-offs that underlie assessment choice. It provides a vocabulary and framework to support practitioners in developing and presenting a well-considered assessment solution for their particular context. Further research is encouraged into the use of the Assessment Wheel as a tool for teaching lecturers in Higher Education about assessment design and for developing and reviewing assessments in a range of different contexts.

References

- Bloxham, S., and Boyd, P. (2007) *'Developing Effective Assessment in Higher Education: A Practical Guide'*, Open University Press, McGraw-Hill Education.
- Boud, D., and Falchikov, N. (2007) 'Introduction: assessment for the longer term'. In *Rethinking Assessment in Higher Education: Learning for the Longer Term*, eds. David Boud and Nancy Falchikov, 3-13. Routledge.
- Brown, G. (2001) 'Assessment: A Guide for Lecturers Series No 3. Learning & Teaching Support Network' (LTSN).
- http://www.heacademy.ac.uk/resources/detail/assessment/assessment_series (accessed June 24, 2012)
- Brown, G., Bull, J. and Pendlebury, M. (1997) *Assessing Student Learning in Higher Education*, Routledge.
- Cummings, S. and Wilson, D. (2003) 'Strategy as Orientation and Animation'. In *Images of Strategy*, eds. Stephen Cummings and David Wilson, 411-437. Wiley-Blackwell.
- Earl, L.M. (2003) *'Assessment as Learning: Using classroom assessment to maximize student learning'*. Corwin Press Inc, Sage Publications.
- Ecclestone, K. (2003) *'Understanding Assessment and qualifications in post-compulsory education'*. National Institute of Adult Continuing Education, NIACE.
- Ecclestone, K. and Swann, J. (1999) 'Litigation and Learning: tensions in improving university lectures assessment practice', *Assessment in Education: Principles, Policy & Practice* 6, no. 3: 377-388.
- Gibbs, G. (1995) *'Assessing Student Centred Courses'*. The Oxford Centre for Staff Development. Oxford Brookes University.
- Gibbs, G. (2006) 'Why Assessment is Changing'. In *Innovative Assessment in Higher Education*, eds. Cordelia Bryan and Karen Clegg, 11-22. Routledge.
- Gipps, C.V. (1994) *'Beyond Testing: Towards a theory of educational assessment'*. The Falmer Press
- Jenkins, A. (1998) *'Curriculum Design in Geography'*. Geography Discipline Network (GDN) Cheltenham & Gloucester College of Higher Education.
- Knight, P. (2001) 'Complexity and Curriculum: a process approach to curriculum-making', *Teaching in Higher Education* 6, no. 3: 369-381.
- Lave, J. & Wenger E. (1991) *'Situated Learning: Legitimate peripheral participation'*, Cambridge University Press
- Lines, D. (2005) 'Developing a variety of assessment methods, including self and peer-assessment – post workshop report'. In *Enhancing Practice: Reflections on Assessment* Volume I, ed. David Lines, 223-227. Gloucester: QAA.

O'Brien, F.A. and Dyson, R.G. (2007) '*Supporting Strategy: Frameworks, Methods and Models*', John Wiley & Sons.

Toohy, S. (1999) '*Designing Courses for Higher Education*', Open University Press.

Biographical note:

Lisa Day teaches Strategic Management in the Business School and is course leader for the MA Management Studies and Diploma in Management Studies. She has a Cranfield MBA and LondonMet MA in Learning and Teaching in Higher Education. She is currently studying for a PhD in Strategic Management with Lancaster University. **Email:** l.day@londonmet.ac.uk