A review of contemporary trends in Higher Education assessment

Sally Brown
Leeds Metropolitan University
sally@sally-brown.net

| Presented: 26/09/2014 | Accepted: 20/02/2015 | Published: 22/06/2015

Assessment is of high importance and contributes significantly to student learning, especially when it is designed to be fit-for-purpose, authentic and focused on integrating assessment with learning (commonly termed ‘assessment for learning’). Drawing on international good practice, this article proposes twelve current assessment trends and ten principles for assuring good practice, while recognising that this is an ever changing field which will continue to challenge academics involved in assessment.

Key words: assessment, fit-for-purpose assessment, technologies for assessment, international diversity in assessment, inclusive assessment

Resumen
La evaluación tiene una gran importancia, y contribuye significativamente al aprendizaje de los estudiantes, sobre todo cuando está diseñada para ser adecuada a su propósito, cuando es auténtica y está centrada en la integración de la evaluación con el aprendizaje (lo que comúnmente se denomina evaluación para el aprendizaje). Sobre la base de buenas prácticas internacionales, este artículo propone doce tendencias actuales de evaluación y diez principios para asegurar buenas prácticas, al tiempo que se reconoce que este es un campo en constante evolución, que continuará siendo un reto para los académicos que se impliquen en la evaluación.

Palabras clave: evaluación, evaluación adecuada a su propósito, tecnologías para la evaluación, diversidad internacional de evaluación, evaluación inclusiva

Resum
L’avaluació en té un gran importància, i contribueix significativament a l’aprenentatge dels estudiants, especialment quan està dissenyada per a ser adequada al seu propòsit, quan és autèntica i està centrada en la integració de l’avaluació amb l’aprenentatge (el que comunament és denomina avaluació per l’aprenentatge). Sobre la base de de bones pràctiques internacionals, aquest article proposa dotze tendències actuals d’avaluació i deu principis per assolir bones pràctiques, al temps que reconeix que aquest és un camp en constant evolució, que continuarà sent un repte per als acadèmics que s’impliquen en l’avaluació.

Parauls clau: avaluació, avaluació adequada al seu propòsit, tecnologies per a l’avaluació, diversitat internacional d’avaluació, avaluació inclusiva
Assessment for Learning: http://www.northumbria.ac.uk/sd/central/ar/academy/cetl_afl/

Assessment has such a crucial role to play not only in evaluating what students have learned and what they can do, but also in motivating them to learn and ensuring that they direct their energies appropriately (Boud 2007).

“Research and experience tell us very forcefully about the importance of assessment in higher education. It shapes the experience of students and influences their behaviour more than the teaching they receive”. (Bloxham and Boyd 2007, 3)

Poor assessment, conversely leads to poor academic conduct, with an excessive focus on working to the test rather than deeply engaging with the subject, and a higher tendency to plagiarise, cut corners and behave strategically (Kneale, 1997). This article aims to review strategically what are the key trends in contemporary higher education, and to propose some principles for assuring good assessment practice.

Twelve current assessment trends

1. Assessment for learning

   Perhaps the strongest perceivable trend is to aim to ensure that assessment is fully integrated with student learning. Higher education practitioners recognise the importance of assessment as a driver for learning and a means by which it occurs rather than just a method of judging student performance. Researchers in the UK and elsewhere (Bloxham and Boyd 2007, Boud and Associates 2010, Gibbs and Simpson 2005, Ibarra-Sáiz and Rodríguez-Gómez 2010, Sambell, McDowell and Montgomery 2012, Brown and Race 2012) all propose that assessment needs to change, to become more central to the learning process, rather than a subsequent add-on, and to fully engage students if it is to be valuable in its own right. Assessment for Learning particularly as defined by the team who led the Assessment for Learning (A4L) Centre for Excellence in Teaching and Learning at the University of Northumbria implies that A4L:

   -“Emphasises authenticity and complexity in the content and methods of assessment rather than reproduction of knowledge and reductive measurement.
   -Uses high-stakes summative assessment rigorously but sparingly rather than as the main driver for learning.
   -Offers students extensive opportunities to engage in the kinds of tasks that develop and demonstrate their learning, thus building their confidence and capabilities before they are summatively assessed.
   -Is rich in feedback derived from formal mechanisms e.g. tutor comments on assignments, student self-review logs.
   -Is rich in informal feedback e.g. peer review of draft writing, collaborative project work, which provides students with a continuous flow of feedback on ‘how they are doing’.
   -Develops students’ abilities to direct their own learning, evaluate their own progress and attainments and support the learning of others.” (Brown 2015, 107)

2. Strategic assessment design

   Good assessment is not a matter of good luck: assessment that engages students and fosters learning is designed strategically to fit the occasion, the students, the context, the subject and the level. Assessment needs to be constructively aligned with the curriculum (Biggs and Tang 2011), so that student learning of the specified outcomes in the published documentation associated with a programme is fully represented in the assignments that students undertake. Those who adopt a fit-for-purpose assessment approach seek to offer assignments that are clear about the purpose of a particular task, use appropriate methods and approaches, are assessed by the right people at the right time to support student learning (Brown 2015). Students need confidence that they will be assessed fairly and justly and they need to be really clear about what is required of them and what standards of work are expected. Assessors need to be confident that their workload is manageable and they are well-placed to make judgements about the quality of work. Institutions need to be able to justify the standards embodied in the qualifications they award, and employers need reassurance that graduates are able to fulfil employment requirements. This implies a highly professional and nuanced approach to assessment design that is purposeful and focuses more on enhancement and less on monitoring (Quesada-Serra et al. 2014). It is also important that curriculum designers avoid assessment design being undertaken at a module-only level: a key current trend in assessment is to consider assessment across a whole programme to ensure a coherent and integrated assessment experience for students (McDowell 2012).

3. Assessment literacy

   Students globally may have widely divergent experiences and expectations of what university-level assessment requires. Those with parents or other family members who have been in higher education may have a better understanding of what terms like criteria and weighting mean than students with less social and cultural capital. By building assessment literacy development activities into, for example, the crucial first six weeks of the first semester of the first year of university study, academics can enhance student achievement and retention. (Yorke 1999). Sambell (2013)
makes a strong case for enabling students to have a sophisti-
cated and articulated understanding of what goes on inside
the black hole that assessment is sometimes perceived to be. Price et al. (2012) further argue for building into the cur-
riculum a range of activities that require students to engage
fully with the practicalities and nuances of assessment prac-
tice as a means of engendering engagement and maximising
students’ chances of success. This is particularly important
in contexts where students from different nations study
away from home, because learned behaviours that worked
in one context may not be useful in others, for example, the
to which memorisation and reproduction in assign-
ments is valued.

4. More (or less?) time on assessment activities
There are competing and potentially mutually exclusive
tensions concerning the amount of time and resource to
spend on assessment. There is, for example, evidence of ex-
pectations by students, particularly in countries where they
pay substantial higher education fees that they should receive
detailed formative and developmental feedback that is sup-
portive, personalised and helps them integrate their learning
(Nicol and Mcfarlane Dick 2006). A number without words
of advice is rarely considered sufficient in these cases.
At the same time there are often directives by university
managers that academics need to do more with less, that is,
spend less time on assessment. For some this means making
best use of available technologies to reduce the drudgery of
repetitive marking and to similarly cut down on the time ac-
ademicians and administrators spend on managing the assess-
ment process, but other less scrupulous managers regard
assessment as a time-waster for lecturers who could be
doing other academic work, particularly research.
The tendency in coming years will be for assessors to
make more strategic decisions about how they use their time
on assessment-related activities: avoiding drudgery as far as
possible but making sure that all time accorded will be spent
prudently to give maximum pay-off in terms of learning.

5. Technologies to enhance learning
There have been significant advances in the potential for
technologies to support assessment in the last three decades
(Beetham 2013, JISC 2007, JISC 2010). Computer-based
assessment (CBA) nowadays goes well beyond simple mul-
tiple-choice questions with a single correct answer and in-
stead requires significant cooperative design and testing,
using expertise in question design, subject knowledge and
the requisite technologies to ensure effective assessment.
CBA can encompass a wide range of simulation activities,
non-text tasks and case study analysis, as well as more tra-
ditional multi-choice questions. When predicting the future
of assessment, few commentators would demur from the ex-
pectation that technologies will be increasingly used to sup-
port all aspects of assessment.
At a wider level, implementing a means of electronic
management of assessment (EMA) is crucial nowadays.
EMA is ‘the way that technology can be used to support the
management of the whole assessment and feedback lifecy-
cle, including electronic submission of assignments, mark-
ing, feedback and the return of marks and feedback’ (Ferrell
2014, 4). Such systems, including those from vendors in-
cluding Blackboard, Moodle, Turnitin, Livetext and so on,
can increase efficiency but need a systematic and institu-
tion-wide approach if they are to have high impact. While
administrators and students are often highly in favour of
EMAs, academics are sometimes reluctant to cede control
of some aspects of the quality control of assessment and
there is often a marked reluctance regarding online marking
(Farrell 2014, 30). Without doubt, the days are numbered
for the practice of hand marking paper scripts with a pen,
since this approach is so problematic for students who can’t
read poor handwriting nor decode cryptic comments and is
so time-consuming and repetitive for academics.

6. Assuring the quality of assessment
There is a growing interest at institutional, national and
international level in assuring the quality of assessment,
with detailed guidance on assessment processes and good
practice provided. Such advice is frequently provided by
Professional, Subject and Regulatory Bodies (like Law Soci-
eties and Engineering Councils), governmental bodies (in-
cluding the Quality Assurance Agency and the Higher
Education Academy in the UK, Australian Learning and
Teaching Council, and the Higher Education Authority in
Ireland,) as well as international groups European initiatives
(McAleese 2013). Assessment is rarely seen nowadays as
a private matter to be left to the discretion of the individual
assessor, nor indeed to individual institution’s (sometimes
arcane) regulatory frameworks.
There are often advanced expectations that staff new to
assessment will be trained and inducted into how to mark
fairly, validly and consistently, and that experienced mark-
ers will engage in moderation activities to ensure inter- and
intra-marker reliability, so several markers working on a
shared set of scripts, and individual markers working their
way through many scripts will ensure that each student is
judged justly against agreed criteria (for example, in the UK
QAA, 2013, Indicator 10, p. 14 and HEA, 2012). Many na-
tions in Europe, Australasia and elsewhere now argue for
training and CPD for university teachers to be undertaken
to improve both teaching and assessment, (e.g. Ibarra-Sáiz
and Rodríguez-Gómez 2010). Recommendation 4 of the Re-
port to the European Commission on improving the quality
of teaching and learning in Europe’s higher education insti-
tutions (McAleese 2013, p. 15) argues that:

“All staff teaching in Higher Education institutions in
2020 should have received certified pedagogical train-
ing. Continuous professional development as teachers
should be a requirement of teachers in the higher edu-
cation sector”.

This trend toward professionalisation of all aspects of
university teachers’ work, especially assessment, is likely to
be an ongoing one, not just in Europe but across the globe.

7. More challenges by students of assessment judgments
Formerly student complaints about assessment tended
to be about due process not being followed but nowadays in-
creasingly students (and their parents in fee-paying coun-
tries) are more likely to challenge the basis on which
judgments are made. This has led in some cases to more
risk-averse behaviour on the part of assessors, who may pre-
fet to use assessment types which rely less heavily on judg-
ment of performance and instead, for example, use multiple
choice tests which are seemingly more objective. However,
even though such methods are on the surface less subject to
bias (with any subjectivity hidden within the design of the
questions, for example) they may be less effective at judging higher level learning outcomes than more fit-for-purpose assignments. Assessors and those designing assessment are likely in coming years to undertake risk assessments associated with changing assessment practices, although they would also be wise to consider the risks of not changing old practices in a time when student expectations are high.

8. Assessment Agency

Currently the vast majority of assessment is undertaken by academics and their proxies including senior students and graduate teaching assistants. Studies in Australia, (Sadler 2010, Boud 1995), the UK, (Fulchikhov 2004, Race 2001), Spain, (Rodríguez-Gómez, Ibarra-Sáiz and Jiménez 2013) and elsewhere argue for the benefits of involving students in their own and each other’s assessment as an important means of helping them gain a deeper understanding of what assessors are seeking in assessed work. The benefits of self and peer assessment are demonstrably so valuable that it is worth overcoming the recognised reluctance of some students to put their grades in the hands of fellow students. To work well, self and peer assessment rely not on snap judgments based on feelings, but on careful calibration of evidence of achievement against explicit criteria. Self and peer assessment should never be undertaken without clear briefing, sensitive training, risk-free rehearsal and dialogic opportunities.

The UK National Union of Students assessment and feedback benchmarking tool2 proposes that to be effective, assessment criteria should be:

“clear, easily accessible and linked to learning outcomes. Students fully understand and are supported to use them. They are designed in partnership with students to ensure accessibility”. (NUS, 2014).

Using such assessment criteria as part of an ongoing iterative and incrementally developed process can enable students to assess themselves and each other fairly and validly. Where this is the case, such assessment can be used summatively, as well as the more common formative use.

Employers, placement supervisors, practice mentors, clients and others are all increasingly being involved in assessment of students, and again briefing and moderation are essential features of broadening the range of people undertaking assessment, together with as before a very clear focus on marking according to evidence presented against the assessment criteria.

9. Authentic assessment

This implies that the tasks students are required to undertake should be closely linked with published learning outcomes, and should be realistic in format. For example, if assessors want to know whether students understand the interpersonal issues involved in effectively working as a member of a team, it is much more authentic to give them a group task to undertake and reviewing their individual and collective performance of it than it is to ask them simply to write an exam answer about teamwork. Peer assessment of colleagues working in their own groups is a valuable way to evaluate behaviours that are not normally visible to the tutor. Employers particularly value graduates who can demonstrate skills and capabilities developed in realistic assessment scenarios.

A case study of good practice in using authentic assessment approaches by Lopez-Pastor (in Brown 2015, 130-2), describes using incremental and ongoing feedback assessment with early years student teachers in Spain. In addition to evaluating students’ competences, the programme requires students to demonstrate their reflection on practice and to demonstrate development over their period of study. It is popular with the teachers who supervise the students’ practice because it is perceived as rigorous and supportive, and also because they are able immediately to resolve any problems that students are experiencing before the end of the teaching practice.

In coming years, there is likely to be continuing pressure from students, employers and other stakeholders for assessment to be meaningful and to relate to the kinds of real-world contexts students will experience on graduation.

10. Growing concerns about plagiarism and cheating

These and other forms of poor academic conduct have led to significant efforts both to design out opportunities for students to cheat (e.g. Newstead 1996, Carroll 2002).

In effect academics can come at this from four directions:

· using technologies like Turnitin or Safeassign to catch students who submit part or whole assignments that are not their own, using frequently cut and pasted from the internet, through text matching systems against extant resources;
· designing assignments and systems that reduce opportunities for plagiarism, like in-class activities, scrupulously invigilated exams and incrementally reviewed tasks;
· having well-publicised and effective systems to catch and punish those who cheat;
· developing a culture of honour where students themselves regard such behaviour as unacceptable. (after McDowell and Brown 2001).

Of these four, the ones most likely to succeed are designing out opportunities to plagiarise and using technologies, though it is much more difficult to prevent or catch out students who commission bespoke assignments through so-called ‘essay mills’ where unscrupulous postgraduates and even academics write assignments to order for hard cash.

11. International diversity in assessment practice

Advances in global higher education practices and staff/student mobility, with a wider recognition that there are multiple approaches to assessment in different nations globally, with the potential for mutual learning, and no monopoly by a single nation or group of nations on correct ways of assessing, are discussed in Brown and Joughin (2007). Issues of divergence are likely to include:

· The amount and quality of feedback students can expect from their tutors prior to submission;
· The amount and quality of feedback students are likely to receive after the assignments have been

marked, and the extent to which students are expected to respond to it;
- The extent to which marks, once awarded, are negotiable;
- How grading systems work: for example a ‘B’ in the UK system is a reasonably good mark but in the US system it would be considered a poor contributor to the Grade Point Average;
- Approaches to referencing vary: in some nations these are rather open whereas elsewhere, for example in the UK, there are very rigorous expectations about acknowledgment of sources and citing references correctly;
- The types of assessment methods in use which are highly variable: in some nations multiple choice questions prevail and elsewhere there is a focus on open text answers. Time constrained unseen exams predominate in some countries with others using more oral assessment (as in Scandinavia and Northern Europe) and others use more continuous assessment in the form of essays and reports;
- How much external scrutiny there is of assessment: in some nations including Ireland, New Zealand and the UK, national quality bodies take a close interest in how quality and standards are assured. Elsewhere, systems of external examiners and moderators are not used.
- The extent that technology is used for assessment: in nations including the US and Singapore computer-based assessment is much more widely used than is the case in a number of southern European nations.

Implications for practice of the global diversity of assessment in higher education are a requirement for induction for staff and students about the different assessment practices they are likely to encounter when working and studying outside their home nations.

12. A commitment to inclusive assessment

Assessment judgments should be based on evidence of the achievement of stated criteria, notwithstanding the race, ethnicity, gender, cultural or social background or physical health or status of the student being assessed and for this reason many nations have adopted strong anti-discriminatory legislation to ensure the inclusivity of assessment in all levels of education, for example the Special Educational Needs and Disabilities Act (SENDA) in the UK. SENDA requires, for example, that reasonable adjustments be made for students with disabilities, such as the provision of software to enable visually-impaired students to hear rather than see examination papers, or for students with chronic fatigue syndrome to undertake elements of a prescribed task incrementally rather than at one sitting.

Sensitive higher education institutions don’t timetable required assessed tasks on Sabbaths, late in the day during fasting times over Ramadan or on Catholic feast days. They don’t include among criteria for presentations making eye contact with assessors for students from cultures (for example among Maori people in New Zealand or some Asian nations) where doing so is deemed inappropriate or improper. Organising and managing inclusive assessment practices is time consuming and complex, but less so than trying to sort out matters at the time of an assessment or dealing with justifiable complaints afterwards.

Whereas formerly students were largely expected to sink or swim, with little account being taken of special circum-

stances or needs, nowadays students and their parents are likely to be intolerant of assessment practices which are seen as being unjust and unreasonable.

Conclusions

Reviewing trends is not a foolproof process, and it is impossible to anticipate all that is likely to happen in coming years. However, having reviewed the state-of-the-art assessment context, here are proposed some suggestions that are likely to remain current in the coming years.

Ten principles for assuring good assessment practice

1. Every decision taken, whether it is on what is to be assessed, when it is to be assessed, by whom, and how, needs to be taken consciously, with a clear imperative in mind to enhance student learning. Established patterns may have to be challenged, and this may be disruptive, but this is unavoidable if assessment is to satisfy all the multiple requirements laid upon it.

2. A strong focus on prompt and developmental formative feedback is likely to have high impact on effective student learning, so it needs to be foregrounded. Both curriculum delivery and current assessment practices may need to change in order to make space and time for the centrality of formative feedback which can not just provide comments and corrections on students’ work to date, but can influence students’ future work: this is commonly described as feed-forward.

3. Innovations in assessment should be undertaken prudently and with due regard for the extensive literature in the field. Scholarly approaches to change are likely to be better received by students, academic colleagues and quality assurance staff than impromptu implementation of bright ideas.

4. Student perspectives on assessment design should be sought early in the curriculum creation process, since as potential end-users, they are likely to have helpful comments on how to ensure assignments work well.

5. While assessment will always exist in a context of power imbalance between the marker and the marked, it is always helpful to foster the dialogic elements of assessment to maximise its impact on learning.

6. Assessment needs to be fair and seen to be fair for students to take it seriously. For this reason, plagiarism needs to be tackled rather than glossed over, and students need to be able to perceive that they are treated with equity and justly (Flint and Johnson 2011).

7. Hygiene factors associated with assessment must be carefully assured. Assignment submission dates and exam dates must be made available early and adhered to. Criteria need to be explicit and accessible and assignment briefings need to be made available in a durable format so that, for example, students unavoidably absent from briefings should be able to access the video or text version on the web. Security of tests and exams should be safeguarded. Exam rooms should provide facilities sufficient to ensure students can undertake tests in comfort, and exams should be impartially invigilated.

8. Where possible assessment tasks should be engaging and even potentially enjoyable: students should feel satisfied in completing them and academics should be able to recognise the learning that has taken place under their care.
9. Designing manageable, effective and incentivising assignments should not be a task for a single individual: avoidance of poor assessment design requires a minimum a small team to discuss, confer and check draft assignments and to monitor their effectiveness in practice. This is particularly true of e-Assessment. The most effective universities have systems in place to ensure the reliability and validity of assessment, and to ensure they align with national as well as Professional, Regulatory and Subject body requirements.

10. Assessment must always be undertaken respectfully and feedback should focus on the submitted work rather than the individual’s personality or previous track record. Good assessment helps students enhance their self-efficacy while poor assessment can undermine self-confidence and result in higher rates of student failure and under-achievement.

If we achieve all of these we can be confident that assessment is fit-for-purpose and fully integrated with learning: that is not ubiquitously the case in universities nowadays but it is a worthy aspiration towards which to aim.

References


Boud, David (2007). Reframing assessment as if learning were important. Rethinking assessment in higher education: Learning for the longer term, 14-25.


Carroll, Jude (2002). A handbook for deterring plagiarism in higher education, Oxford: Oxford Centre for Staff and Learning Development


http://repository.jisc.ac.uk/5599/1/EMA_REPORT.pdf. Accessed 01.01.2015.


Price, Margaret; Rust, Chris; O’Donovan, Berry; Handley, Karen and Bryant, Rebecca (2012). Assessment Literacy:
the foundation for Improving student learning, Oxford: Oxford Centre for Staff and learning Development.


Race, Phil (2001). A Briefing on Self, Peer and Group Assessment, in LTSN Generic Centre Assessment Series No 9 York, UK: HEA/LTSN.


Sambell, Kay and Hubbard, Anntain (2004). The Role of Formative 'Low-stakes' Assessment in Supporting Non-traditional Students' Retention and Progression in Higher Education: Student Perspectives, Widening Participation and Lifelong Learning, 6 (2), pp. 25-36.


---

| How to cite this article |