



OXFORD UNIVERSITY CENTRE FOR EDUCATIONAL ASSESSMENT

Report of Roundtable (23rd June 2008) as compiled by Velda Elliott

Policy Drivers

- Increasing public understanding of systems and terms is vital – as is convincing govt that it is.
- Awareness of academics that assessments/qualifications can never be as robust as govt would like them to be: to date has been kept somewhat quiet.
- ‘Standards’ are highly problematic. Little understanding among the public/ govt of what this means. Technical standards of assessment agencies are also important. Reconsidering what a standard is: relative or absolute?
- Focus in UK on school standards: what about Higher Education? Are assessment paraphernalia and mechanisms as robust in HE as in schools?
- Most ‘political drivers’ are in fact driven by votes. It doesn’t matter about the research when politicians are working on a short term view. How can researchers influence the real policy debates before they become ‘done deals’?
- Changing design of courses and curricula without fully thinking through implications of technical assessment considerations, e.g. change to modular assessment.

Theories and Models of Educational Assessment

- Personalising Learning brings up many assessment issues:
 - Challenge of preventing plagiarism or cheating in situations where assessments become available in different people at different times and are not withdrawn after administration.
 - Technical competency-based assessments problematic in this context.
 - Looking outside traditional systems for models of personalised assessment.
 - Standardised or non-standardised tests, relative to course of learning.
 - Possibility of using assessment as a motivating rather than demotivating force.

15 Norham Gardens, Oxford, OX2 6PY, United Kingdom

Phone: +44 (0) 1865 274002

www.education.ox.ac.uk/assessment/

Email: admin.oucea@education.ox.ac.uk

- Changes in assessment systems are used to drive curricula change: not assessment for learning but learning for assessment in most cases.
- Assessment for learning has been reified in England as a series of teaching strategies. Can research in assessment move on from formative assessment? How does/can summative assessment feed back into learning?
- Role of the teacher in assessment: lack of trust in TA, despite studies showing reliability can be achieved.
- Assessment of group-based work
 - Group based work is increasingly important, particularly in HE assessment
 - Group working skills important for employers
 - Currently often assessed *ad hoc*/ variably
 - Ways of monitoring assessment of group work
- Broader models of assessment: a more 'personalised' report; training employers to interpret such reports.

Stakeholder Perception

- Issues of communication
 - Notions of standards in performance works well for teachers in assessing levels of performance but cannot then be communicated to parents because it is in teacher language. Getting the language right to talk to different stakeholders is an essential issue.
 - What do people want? Have moved into giving parents the mark, which is what they want to know, and some extra information which they ignore. Is it possible to move to giving just the information?
 - Once the agenda is in the public domain researchers will always be responding to public discourse rather than setting terms appropriately.
- Different purposes loaded on to assessment arrangements
 - When try to separate the functions of assessment get accused of moving away from public accountability.
 - Abuse of purposes of data in the policy machine.
 - Govt bad at articulating fitness for purpose, and at aligning instruments with those purposes
- The accountability agenda
 - Why do we have such a competitive perception of/ aspirations for our education system?

- Parents say schools aren't good enough: perceived by govt as call for accountability measures.
 - Question of how big an effect the school can actually have versus other influences on attainment
- Public understanding of assessment is a thorny issue; in essence however, even if people know that assessment is meaningless (invalid etc) because everyone is in the same boat it will still be seen as fair.
 - Teachers in UK have been deskilled in terms of assessment.
 - How can we make assessments good enough that teaching to the test becomes a desirable outcome?
 - Average standards have become aspirational targets: can anything be done to challenge this?
 - How does accessibility interact with difficulty?
 - Standard setting is different from standard maintaining: initially it can be quite arbitrary, but maintaining it becomes another challenge.

Technical Issues

- Accountability and league tables
 - Matthews effect: learning trajectories
 - Phantom effect: random error masquerading as school effect
 - Other issues of bias skew league tables: if we have to have them, and they are a good thing in helping to identify schools requiring help, then there is a need to ensure that they are indeed measuring what they should be, and to find the best model for the purpose.
 - Teacher effects do not seem to be transferable across classes.
- Problem of lack of underlying psychometric scale, for example beneath NC levels.
- Inspection of schools: data on inter-rater reliability for OFSTED.

The E-Agenda

- Assessment at item level
- Transposition of items from paper to screen: comparability? Same measures of reliability?
- Future of item banks which can be used around the world: issues of reliability and validity in different cultural contexts.
- Innovative item design
- Reporting of assessment
- Potential to report back much more quickly (instantly?): has implications for use of 'summative' data by teachers and pupils for formative purposes.
- User engagement with test is different: can't just flick back a page; how do you deal with stimulus material? Research to be done in terms of user interface with the test, and how an examinations process can work.
- Is there a role for multiple choice, previously rejected by many in the UK?
- Computer aided assessment
- Some subjects problematic
- Short answer questions problematic
- Not a lot of research being done in this area in the UK
- Commercialised systems no longer being critically examined for areas of improvement because they are operating in a competitive marketplace.
- Worries about using proxies which can be artificially manipulated by candidates once system is in place
- Are existing human marker processes all that desirable?
- Introducing it from the bottom up so that teachers themselves become the advocates
- Original notions of seamless integration of assessment into instruction can be accomplished using technology.
- E-assessment has great potential in home environment: self-assessment, supporting personal learning.