ICT in Assessment:
A three-legged race

Robert Harding
Director of Interactive Technologies in Assessment and Learning (ITAL Unit)
University of Cambridge Local Examinations Syndicate (UCLES)
1 Hills Road, Cambridge CB1 2EU
Email: r.d.harding@ucles-red.cam.ac.uk

Patrick Craven
Manager, Computer Based Assessment
Oxford, Cambridge and RSA Examinations (OCR)
Westwood Way, Coventry CV4 8JQ
Email: craven.p@ocr.org.uk

1. Introduction

The potential for the use of new technologies in education was recognized from the earliest days of interactive computing, and courseware (computer based learning materials) usually includes some self-assessment for formative purposes. Insufficient attention has been paid to the link between formal assessment and use of courseware. On the one hand, teachers and learners focus on examination syllabuses. On the other, examinations bodies can use only the simplest forms of ICT based assessment (objective or multiple choice tests) until learners are familiar with newer styles. UCLES is addressing this problem on several fronts. These include launching new ICT styles of tests and commissioning new types of examination questions built around on-line interactive simulations. Many projects have the participation of the ITAL Unit.

2. What we have been doing

Here are a few examples of UCLES activities in this field:
- CommuniCAT (EASA 2000 Award[1])
- The CLAIT and WP markers from OCR provide automatic marking in a range of desktop application skills
- MEI A-Level mathematics modules use computers in the live examination
- Work on electronic portfolios and distance verification/authentication for NVQs and other portfolio-based qualifications
- Online O-Level mathematics revision[2], Singapore, from SNP Multimedia with ITAL support.

3. The Three-legged Race and other Issues

The link between summative, certificated examinations and classroom practice is like the tie between a pair of contestants in a three-legged race. Neither can move forward without the other.

Why is this? We must first ask why examinations exist at all. Examinations strengthen the **credibility** of education, and the reasons certainly include:
- A goal to work towards – recognition for candidates' achievements
- Confidence by the public in the standards represented by awards
- Part of the way that the education profession discharges its duty of accountability
- A reliable and fair method of measuring the performance of the education system

There is no question but that ICT can make an immense contribution to learning. It is however self-evident that in order to be fair, a summative assessment system must be based on activities that are familiar to the learner. So the examination boards are shackled to the learning system. They *could* break free and run ahead using non-certificated formative assessment products, but this does not avoid the need for certificated examinations, as without these candidates would have no way of knowing that their achievements had been recognized. The public might then lose confidence that standards were being properly measured and maintained. In other words, educational credibility would be damaged. Some critics of examination boards blame them for 'holding back' the use of ICT in education. However, examination boards may feel 'held back' in their desire to use ICT in the examination process by the lack of ICT provision and use in schools.

---

[1] http://www.easa-award.net/index2.html

[2] http://www.col.com.sg/

There are other issues. For example, if there are two different styles of examination, can we be confident that they are comparable? Might one form give an advantage to those who are already more privileged (e.g. might have superior IT skills and so find an ICT based test easier than other forms taken by the less privileged)?

4. Roles for ICT in assessment

There are two categories of roles:
1. There are many reports from the Higher Education sector that ICT can be used to improve efficiency and effectiveness in learning, teaching and testing
2. There are similar claims of opportunities in association with new styles of learning.

4.1 Improving efficiency - some caveats

We include ICT based multiple choice (objective) testing under the first category, since it seems reasonable to assume that objective testing measures the same things whether taken using paper and pencil or ICT[3]. Note however that it does have consequences for education because of the ease and speed with which such tests can be administered. The drawback that is commonly cited is that if this were to become the dominant form of testing, UK education would be impoverished as schools adopt "teach to the test" tactics to maintain their ratings and 'richer' forms of assessment requiring more contextualisation and human intervention are marginalised.

We have talked about ICT improving 'efficiency'. We want to warn that this does not necessarily mean reducing costs. We think that the burden on the education system of maintaining office-type ICT facilities will be considerable. But just as a modern efficient office without desktop ICT is hard to imagine, so we think education will want and need ICT. It is here to stay, and any argument that it should not appear in students' lives until they leave education and go into work is surely now a lost cause.

What form will ICT provision take? It is hard to be sure, but we doubt that "one per student" at any level of education is feasible in the near future. Even if it were, there need to be special arrangements come examination time, and this is quite a daunting task. In a typical secondary school of say 1000 pupils, there might be around 200 entries for a typical GCSE. Think of the technical challenge of setting up 200 computers, probably moving many of them from their usual positions, checking their discs, software and networking connections, and supervising them. There will need to be special arrangements to ensure candidates cannot see each others' screens, and the invigilators will need a high level of ICT literacy. We do not think this could be done without substantial additional resource.

The level of teacher expertise with ICT is a major factor. In his BETT 2001 keynote speech, Lord Puttnam welcomed the NOF teacher training initiative but doubted that it would in itself be enough to deliver the major paradigm shift that is needed if ICT is to make a major impact on Education. Professional development needs to be sustained over a much longer period. (UCLES recognized this difficulty over three years ago and is one of the founding sponsors of the TEEM Project[4], which as all who visited BETT 2001 will know is now making a significant contribution to classroom practice.) We must also recognize that our cadre of examiners is drawn from the teaching profession. If we want our examination system to be more than just objective testing (for the reasons already put forward), then examiners will need training in the use of ICT in this context. They will also need to have suitable equipment, another considerable cost to add on to the assessment process. We think that the cost and difficulty of examiner retraining is probably not being given adequate attention in this debate.

4.2 Opportunities with new styles of learning

We mentioned the difficulties but let us not underestimate the opportunities. Some of our greatest challenges will conversely offer the greatest opportunities and if overcome will realise the greatest benefits to the educational community. We need to debate what is being assessed and how it relates to current learning goals, or else we lose continuity and possibly comparability of standards. Here are some opportunitites.

Progress Tracking and performance reports are most commonly used with portfolio-based assessment such as NVQs and Key Skills. They allow the candidate, assessors and external auditors to identify a candidate's achievement and progress.

---

[3] But note that this could be a controversial claim and it certainly needs to be tested by research.
[4] http://www.teem.org.uk

Authentic vocational skills tests are an increasingly common form of test for competences which can be measured in a truly objective manner. Such tests can take this type of assessment to a higher plain providing a more valid test of aptitude. At the far end of the spectrum simulations are increasingly being used to assess practical skills where the opportunities to demonstrate competence are too risky, expensive or do not occur frequently enough to support training programmes. Classic examples of this include flight simulators, battle training and training support for field engineers.

If we have a single hope for the use of ICT it is to "Give teachers back the time to teach" and this can be achieved in a variety of ways. Developing and maintaining the credibility of an assessment is at the heart of the challenge where accreditation is concerned. As we have said earlier, these fundamental issues do not change just because ICT is used. All else flows from this. We must remember that where ICT is concerned we are dealing ultimately with software. The public may accept bugs and flaws in applications software but not so in an accreditation system. It is both our biggest challenge and the greatest opportunity. An effective ICT solution can yield many benefits to teachers/lecturers in key areas such as marking scripts, management of tests and reporting results. ICT gives us the opportunity to make significant improvements to the perceived bureaucracy of current systems but introducing ICT solutions is fraught with more personal challenges (ie fears of ignorance, lack of tangibility, steep learning curve, etc). These challenges associated with fear of change are both external and internal and Awarding Bodies cannot initiate progress alone.

There are also opportunities to explore new methods of assessment using high level simulations based on 'gaming technology', assessment of teamworking skills where an individual's contribution might also be tracked and ultimately the ability to interrogate process rather than product. Such radical shifts in assessment methodology may require new 'metrics' and will certainly require new 'question types'. We are then drawn back to the previous challenge/opportunity we identified, to establish their credibility.

4.3 Technological challenges

If CBA systems are taken to their extreme candidate authentication will always be the 'holy grail' for those wishing to gain acceptance of such solutions. We currently choose to take the 'Threat Model' approach and insist that the CBA is conducted under administered conditions. If we are to maintain the rigour expected of us by regulatory bodies and employers true online assessment is still hampered by this obstacle. User IDs, PINs and Passwords are no deterrent, as they are no guarantee that users are who they claim to be. Biometrics - such as voice, iris, fingerprint, face and scent recognition are no guarantee that the user is on their own. The same is true of cognitive fingerprinting which although confirming that the user is inputting data is no guarantee that the user is not being coached. Video conferencing technology probably comes closest because our perceptions of the world around us are largely based on visual clues. In theory it offers a solution, in practice it is not sufficiently widespread and is plagued by technical and feasibility issues. There are still however many opportunities for further research into this area especially as the world of commerce begins to address the issues of digital authenticity.

5. Striking the balance

The diagram illustrates the competing forces on the design of an examination.
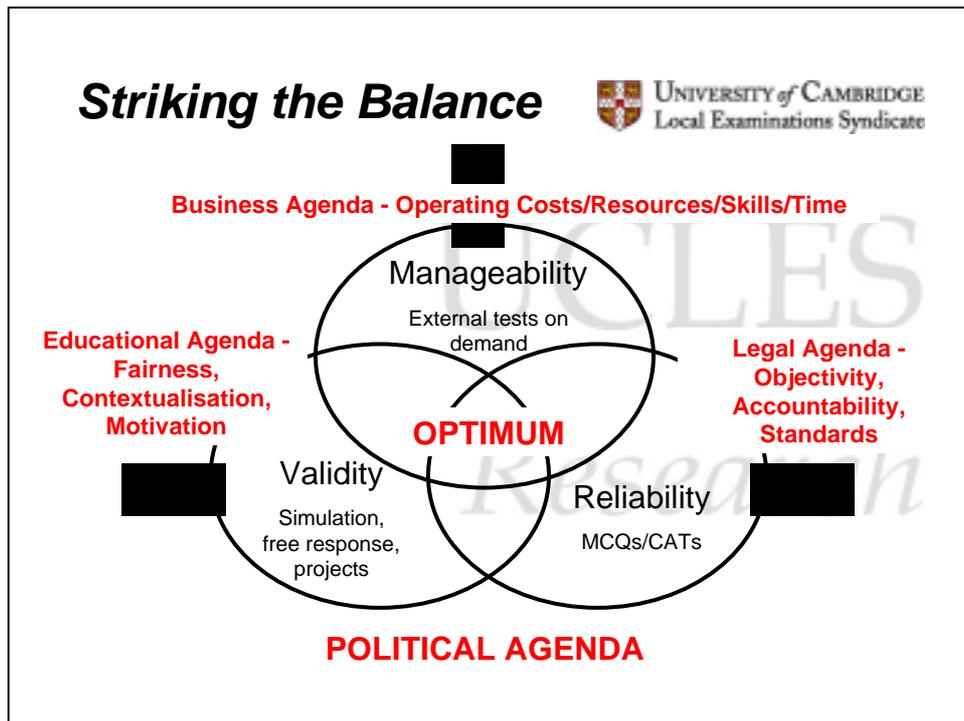
Reliability and auditablility
A reduction in dependence on human subjective judgement is possible (as with for example OCR's WP Marker). As already noted however, with other forms of objective testing there may be a price to pay in the possible effect on the style and quality of education. Perhaps in time advances in artificial intelligence will make it possible to mark free-form text to acceptable standards but this does not presently seem to be feasible.

Validity
What qualities is an examination supposed to test? Some, such as spelling, are relatively easy to define and measure, but more advanced qualities such as 'problem solving' are not. Usually a compromise is struck with reliability. For example, questions for a particular syllabus or specification often follow a well-tried pattern.

Manageability
This is why many public examinations are taken in a small number of sittings. The arrangements reflect the practical difficulties of setting, sitting, scoring and grading. For ICT based tests, the practicalities point in the other direction, towards continuous assessment or at least sessions spread out over many sittings.

Striking the Balance

UNIVERSITY of CAMBRIDGE
Local Examinations Syndicate

Business Agenda - Operating Costs/Resources/Skills/Time

Manageability

External tests on demand

Educational Agenda - Fairness, Contextualisation, Motivation

Legal Agenda - Objectivity, Accountability, Standards

OPTIMUM

Validity

Simulation, free response, projects

Reliability

MCQs/CATs

POLITICAL AGENDA

6. 3+ year parallel ICT testing trial

The use of ICT in the first category (improving efficiency) is more of an operational matter than an educational issue, provide that one avoids the risk we have mentioned of impoverishment of education.

The second category promises much more, but to use ICT to its full potential in teaching and learning will require radical changes to classroom practice and to the style of assessment. There will need to be full public debate before old practices can be widely abandoned in favour of new ones more in tune with the medium. The kind of information that will inform that debate could come from research in which results from the existing system are compared with results from an ICT based system over a suitable period of time.

Several factors point towards using ICT more continuously in a learning-monitoring partnership. This promises real educational gains whilst helping with the resource and manageability aspects, although there is an issue with teachers' ICT familiarity. We suggest a trial over three or more years in which ICT-based teaching and testing run in parallel with traditional GCSE / A-Levels. Clearly a great deal of detail needs to be worked out, but the strategic objective of such a trial is clear.

7. Other research opportunities

Here we will simply give a short list of possibilities:
- Oral exams via the Internet
- Collaborative learning studies (teamwork, problem solving etc)
- Try out the use of electronic paper type devices such as the I-Pad
- Investigate the problem of doing mathematics on screen
- Studies into Authenticity
- Explore the potential enriched nature of CBA (process over product)

8. Conclusion

Computer-based learning is clearly set to play an increasingly important role in education. The roles of computer-based learning and computer-based assessment are closely linked and progress in one depends on progress in the other. Support is also needed for teachers and lecturers as they learn how to take advantage of both roles. UCLES is ready to play its part to assist and encourage every part of this process.

© UCLES 2001