

CAN E-ASSESSMENT BECOME MAINSTREAM?

Helen Ashton, Cliff Beevers and Ruth Thomas

Can e-Assessment Become Mainstream?

Helen Ashton¹, Cliff Beevers¹ and Ruth Thomas²

¹ Heriot-Watt University, Edinburgh,
H.S.Ashton@hw.ac.uk

² JeLSIM Partnership, www.jelsim.org

Abstract

For over twenty-five years the field of e-Assessment has grown steadily at the margins of assessment practice. In the past few years, there has been consideration of the form it might take in mainstream assessment and of the barriers and drivers affecting its uptake (e.g. QCA (2005), JISC (2006, 2007)). In this presentation, we identify the issue of pedagogic validity of e-Assessment as central in its move to the mainstream and look at what this means in terms of the stages in e-Assessment growth proposed by Bennett (1998). Achievement of pedagogic validity is a necessary precursor to mainstream acceptance, but on its own it is not sufficient. It also requires well-trained and informed practitioners who are able to understand and address issues relating to use of e-Assessment: knowledge and understanding of assessment practice and technical skills to devise or create appropriate tests. The embryonic e-Assessment Association (<http://www.e-assessment.com>) has recently emerged as a result of this need, and we will outline the main aims of this organisation.

Introduction

When we talk about “mainstream” uptake of e-Assessment, what do we mean? To be mainstream, e-Assessment does not need to be the assessment technique of choice for all subjects at all levels, but the implication is that: it will be the technique of choice of a large number of practitioners; that large numbers of people will sit e-Assessments; and there will be appropriate e-Assessment available for a wide range of subject areas and that a majority of educational objectives (including higher order skills) can be tested.

Issues identified in the QCA (2005) and JISC (2006, 2007) studies as potential barriers to uptake of e-Assessment include pedagogy (development of best practice – authoring, marking, reporting), validity and reliability, technical (hardware/software reliability and security), examination procedures, staff skills and understanding, management attitude and awareness, technical, administrative and specialist support and legal matters (plagiarism, data protection and intellectual property rights). The QCA (2005) report identifies the top three drivers as cost savings, government policy and immediate feedback.

Responding only to a driver to reduce costs without paying attention to quality of assessment would be short sighted. A reduction in the quality of assessment will have a profound impact on education in general and, in the long term, on the acceptability of e-Assessment. To win the hearts and minds of all stakeholders: teachers, government, awarding bodies and the public, e-Assessment must have pedagogic validity. It must begin to improve standards and become part of an educational system in which teaching learning and assessment are truly aligned. E-Assessment must be pedagogically acceptable at all stages of the education process. When this is achieved some barriers to mainstream adoption will be removed, as it will be seen as an acceptable testing method for accreditation by all stakeholders and the public (including the media). All will perceive it as valid, reliable, appropriate and fair.

In this presentation we will focus on pedagogic acceptability as the key to long term mainstream e-Assessment. In the past we have spoken about the pedagogical acceptability of e-Assessment in general terms, arguing for the need for constructive alignment of teaching, learning and assessment as well as for authentic assessment to cover all the learning objectives of a course (Thomas et al 2004). In this present article we want to focus specifically on what this might mean in the longer term as e-Assessment technology and techniques develop.

Bennett (1998) predicted three generations of e-Assessment:

1. Traditional questions converted to electronic format;
2. An increasing use of interactive multi-media to open up new possibilities; and
3. Breaking down the barriers between assessment and learning – students assessed as they learn through complex interactive environments.

The stages described by Bennett have been further endorsed in the review articles by Ridgway et al (2004) and its recent update by Ripley (2007). These latter authors also cite the use of white boards and personal response systems in the changing classroom practice and the introduction of e-portfolios as important drivers towards a more e-enabled future.

In this presentation we will look at the implications for pedagogic acceptability and other factors at each of Bennett's stages if the move of e-Assessment to the mainstream is to be accomplished.

Moving to the mainstream with each generation of e-Assessment

1. First generation

In this generation of e-Assessment the early effects of innovation, customer service and productivity emerge, but fundamentally, the tests are no different

from traditional paper and pen tests. These assessments measure the same skills, use the same behavioural designs and depend primarily on similar types of tasks. This first generation is therefore typified by a translation of existing practices into electronic format.

We routinely make the assumption that current assessment techniques are fair, valid and reliable. If this is not true, then in converting traditional assessments into electronic format we will take these problems with us. However, even if we assume acceptability of our traditional questions, then the process of converting these into online assessments may invalidate this situation, particularly if we are not able to duplicate the traditional format of the question.

The availability of e-Assessment technology has a huge influence on the pedagogic acceptability of e-Assessment at this stage. Many current assessment systems focus on *selected* response questions (i.e. Multiple Choice), as authoring tools are straightforward to create and the questions are typically easy to mark. However, many traditional questions are *constructed* response questions. The creation of a selected response question as a replacement may not fulfil the same learning objectives. Objectors to e-Assessment often refer to the inability of current technology to allow the assessment of higher order skills and doubt whether it can provide questions to test the full range of educational objectives outlined in taxonomies such as those defined by Bloom et al (1956). Whilst it is possible to write selected response questions which test higher order skills, it is by no means easy. The creation of good selected response questions has many well documented pitfalls – creating plausible distracters, ensuring the use of wording which does not give away the answers, avoiding “working backwards”, failing to take account of guessing and so on.

Question writers require an in depth understanding of these issues. Without this awareness and training, authors may produce questions that are inappropriate and of poor quality. This has an impact on the perception of e-Assessment - if the only exposure you have of e-Assessment is to poor quality selected response questions, then it is not a surprise if this becomes a poor perception of e-Assessment rather than the quality of the questions themselves.

However, there are e-Assessment systems which support traditional constructed response questions (for example, Ashton et al (2003), Sangwin (2004)). There are still avenues to be developed further in this first generation, for example, improving technology in the marking of textual responses (e.g. Intelligent Assessment Technologies (2006)). There are also technological restrictions on the variety of automatically marked question types available, but as technology improves these will reduce even further.

In this period of first generation e-Assessment, where students may be assessed electronically and traditionally, it is important to be sure we are creating an equivalent assessment environment. In order to explore this further, a number of researchers have investigated the comparability of paper

and electronic assessment with positive results (Beevers et al (1995), Greenwood et al (1998), Sandene et al (2005), Ashton et al (2005, 2006)).

There are already compromises being made in traditional assessments, and at the end of this generation we have carried through these compromises. It is likely that there are areas which these traditional practices do not assess appropriately, and for which we need to explore new assessment practices.

2. Second generation

In the second generation of e-Assessment, new question formats will emerge making extensive use of multimedia. These new assessment types can be used to extend the assessment of traditional skills (e.g. Watch this video clip and answer the questions which follow) and assess new sets of skills (e.g. Take this information and create a web page which meets a specific set of criteria). New opportunities for adaptive testing and sophisticated marking will also appear. This second generation will be typified by a move away from traditional assessment to more authentic 'performance' assessment that is impractical before the advent of new technology.

It will be necessary to rethink curricula and learning objectives as second generation e-Assessment becomes more widespread. For example, the use of sound and video (in addition to text and images) can make questions more realistic and appropriate (e.g. the computer based examination of "Moving Image Arts" in Northern Ireland (CCEA) where students are required to discuss concepts in real-life video clips which are presented online). Increased use of media may reduce the emphasis on language skills; this could change the construct validity of the test (whether it would increase it or decrease it depends on the original purpose of the test).

Paper-based assessments are becoming increasingly less authentic as classroom and professional practices embrace the use of computing and IT. Computer skills such as formatting a document properly or correct use of a spreadsheet can only properly be assessed by using a computer. Students may use computer-based tools in the classroom that they will later use in their professional life (for example, in the USA, students have an option of using a Computer Algebra System when taking the College Board's Advanced Placement Calculus test). These types of skills can only be tested using an application of second generation e-Assessment principles.

During this phase of e-Assessment growth it will be possible to pose questions which could not otherwise be asked. The work of Thomas et al (2003, 2004) provides a better understanding of future possibilities: interactive multi-media elements have combined with an assessment engine to deliver simulation exercises as predicted by stage 2 of Bennett's roadmap.

3. Third generation.

In the third generation, assessment will be reinvented, and the barriers between assessment and learning will become less obvious. Students will be assessed as they learn, through complex interactive environments such as simulations and microworlds. Intelligent tutoring systems will collect information about a student's performance and automatically adapt the learning content which is delivered to them – to reinforce existing knowledge, to build on strengths, address shortcomings and suggest new avenues for exploration. This generation of assessment is typified by a high degree of personalisation of the learning and assessment process.

Integrated assessment that provides feedback to students as they learn signals the end of the assessment engine as a monolithic software application. Assessment will not be a distinct activity; it will become integrated within teaching and learning, available on demand at any time, related to the student's personal needs and learning context. This will clearly demand an even greater rethink of education than the advent of the second generation technology. If these goals could be achieved, e-Assessment could truly become mainstream as it would be fully integrated with the learning process.

Current technology is barely scratching the surface of this generation of e-Assessment. The paper by Thomas et al (2003) touched on the idea of assessing students as they learnt whilst using a simulated practical experiment, Maughan and Mackenzie (2004) have shown how a simulated microscope environment can be used for the assessment of outcomes and process in Biology, and simulated environments have been used to assess skills in using computer applications to solve specified tasks (for example the ICT key stage 3 assessment (NAA)).

Other factors

Although achievement of pedagogic validity is a necessary precursor to mainstream acceptance of e-Assessment, it is not on its own sufficient, it also requires well trained and informed practitioners who:

- are able to understand and address issues relating to the use of e-Assessment;
- have knowledge and understanding of assessment best practice;
- possess the technical skills to devise or create valid, reliable tests;
- have access to ubiquitous and easy to use tools or resources; and
- are able to take advantage of high quality continuing professional development courses to allow them to become familiar with the principles and best practice of e-Assessment as it develops.

Building this strength and depth of e-Assessment expertise requires the development of a strong professional organisation to support this. It is possible that the embryonic e-Assessment Association (eAA) (See Appendix)

could be that body in the UK. The eAA, an independent organisation for e-Assessment, aims to provide support for professionals in the field, collect best practice guidelines and communicate the positive aspects of e-Assessment. Together with improvements to the pedagogic validity of e-Assessment, it could help with the transition of e-Assessment to the mainstream and be an important stepping stone to its widespread adoption.

References

Ashton H.S., Beevers C.E., Korabinski A.A. and Youngson M.A., (2005) Investigating the Medium Effect in School Chemistry and College Computing National Examinations, *British Journal of Educational Technology*, 36(5), 771-787.

Ashton H.S., Beevers C.E., Korabinski A.A. and Youngson M.A., (2006) Incorporating partial credit in computer-aided assessment of Mathematics in secondary education, *British Journal of Educational Technology*, 37(1), 93-119.

Ashton H.S., Schofield D.K., Woodger S.C., (2003) Piloting Summative Web Assessment in Secondary Education, *Proceedings of the Seventh International CAA Conference*, Loughborough University, UK

Beevers, C., McGuire, G., Stirling, G., and Wild, D., (1995) Mathematical Ability Assessed by Computer, *J Comp Educ.* 25, pp.123-132.

Bennett, R., (1998) Reinventing Assessment: Speculations on the Future of Large-Scale Educational Testing, *Educational Testing Service*, <http://www.ets.org/research/pic/bennett.html>

Bloom, B.S., Englehart, M.D., Furst, E.J., Hill, W.H., and Krathwohl, D.R., (1956) *Taxonomy of Educational Objectives: Handbook 1: Cognitive Domain*, Longman, New York.

CCEA Moving Image Arts: <http://www.ccea.org.uk/movingimagearts/>

e-Assessment Association: <http://www.e-assessment.com/>

Greenwood, L., Cole, U.M., McBride, F.V., Morrison, H., Cowan, P., and Lee, M. (1998) Can the same results be obtained using computer-mediated tests as for paper-based tests for National Curriculum assessment? *Int Conf Mathematics / Science Educ and Tech*, Vol. 2000, Issue. 1, 2000, pp.179-184.

Intelligent Assessment Technologies (2006) E-Assessment of Short-Answer Questions, http://www.intelligentassessment.com/pdf/IAT_Free-Text_White_Paper.pdf

JISC (2006) Roadmap for eAssessment <http://www.jisc.ac.uk/media/documents/programmes/elearningpedagogy/roadmapeassessment.doc>

JISC (2007) Effective practice with eAssessment <http://www.jisc.ac.uk/media/documents/themes/elearning/effpraceassess.pdf>

Maughan S Mackenzie D (2004) Bioscope: The Assessment of Process and Outcomes using the TRIAD System, *Proceedings of Eighth International CAA Conference*, Loughborough University, UK

NAA Key Stage 3 ICT assessments:
<http://www.naa.org.uk/naaks3/default.asp> (with further reports at
<http://www.naa.org.uk/naaks3/384.asp>)

QCA (2005) Adopting e-assessment for UK awarding bodies: Drivers and barriers,
http://www.ofqual.gov.uk/files/pdf-05-1671_Barriersanddriverstoeassessment.pdf

Ridgway J., McCusker S. and Pead D. (2004) E-assessment – an update on research, policy and practice, Futurelab e-assessment literature review.

Ripley M., (2007) The four changing faces of e-assessment 2006-2016, Futurelab e-assessment literature review.

Sandene B., Horkay N., Bennett R.E., Allen N., Braswell J., Kaplan B. and Oranje A. (2005) Online Assessment in Mathematics and Writing: Reports from the NAEP Technology-Based Assessment Project, Research and Development Series,
<http://nces.ed.gov/nationsreportcard/pdf/studies/2005457.pdf>

Sangwin C.J., (2004) Assessing mathematics automatically using computer algebra and the internet, Teaching Mathematics and its Applications, 23(1):1–14, <http://web.mat.bham.ac.uk/C.J.Sangwin/Publications/tma03.pdf>

Thomas R.C. and Milligan C.D., (2003) Online Assessment of Practical Experiments, Proceedings of the Seventh International CAA Conference, Loughborough University, UK, pp 421-430.

Thomas R.C., Ashton H.S., Austin W.J., Beevers C.E., Edwards D., and Milligan C.D., (2004) Assessing Higher Order Skills Using Simulations, Proceedings of the Eighth International CAA Conference, Loughborough University, UK, pp 417-427.

Appendix: The e-Assessment Association

The eAA is a non-profit making independent professional body that aims to promote the effective use of e-assessment in support of learning: by seeking to advance understanding of the concepts and practices of e-assessment, by

- encouraging individuals and organisations to make good use of e-assessment;
- explaining the issues to the public; and
- influencing educational thinking and policy.

In pursuit of these aims, the eAA's main objectives are to:

1. Provide members with professional support;
2. Work collaboratively with national organisations to communicate the positive contribution that technologies can make to assessment;
and
3. Produce a statement of good practice for commercial vendors of e-Assessment.