THE FORMATIVE USE OF E-ASSESSMENT: SOME EARLY IMPLEMENTATIONS, AND SUGGESTIONS FOR HOW WE MIGHT MOVE ON

Andrew Boyle

The Formative Use of e-Assessment: Some Early Implementations, and Suggestions for How We Might Move On

Andrew Boyle,
Assessment Research team,
Regulations and Standards division,
Qualifications and Curriculum Authority (QCA),
83 Piccadilly,
London W1J 8QA.
0207 509 5349
BoyleA@qca.org.uk
http://www.qca.org.uk/

Abstract

This paper reviews research into the formative use of e-assessment. The review groups implementations into three areas, and then suggests areas for further research in each area. There are nine areas for further research in total.

The discussion section examines the areas for further research to establish commonalities between them. By this process, it proposes four key issues to inform the future of formative e-assessment research.

The key issues are:

- Better defining those instances where formative e-assessment provides particular benefit over and above benefits that would accrue from the use of formative assessment in any medium.
- Being aware of – and attempting to avoid – formative e-assessment implementations that represent a reduced or impoverished conception of formative assessment.
- Being aware of circumstances in which the introduction of formative e-assessment could lead to increased burdens on classroom practitioners.
- The need to understand how students will be required to adopt novel roles (e.g. different ways of working and communicating) when using formative e-assessment.

Introduction

Early e-assessment soothsayers made several predictions. An important one was that e-assessment would facilitate a lowering of barriers between assessment and learning. It so happens that the early years of e-assessment implementation have coincided with a heightened interest in formative assessment (FA).

Thus, it is felt timely to conduct a literature review into the formative use of e-assessment. This review looks across studies and attempts to spot frequent implementations of formative e-assessment (eFA), then group and present them to give an insight into what has been done most frequently in this field.

However, this is also a critical review. As well as constructing categories of frequently used implementations of eFA, the review points out issues that are not adequately resolved and suggests further research to rectify omissions or misunderstandings that currently exist. Building upon those suggestions for further research, the review concludes by proposing four key issues for improving the body of research into the formative use of e-assessment (eFA).

Definitions

Formative assessment

Black & Wiliam (1998a) define formative assessment as follows:

> [FA] encompasses all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged.

Other terms have been used to refer to formative assessment, including 'Assessment for Learning' (AfL) (Black & Wiliam, 1998a) and 'classroom evaluation' (Crooks, 1988).

Formative assessment is often contrasted with summative assessment. Summative assessment is assessment that summarises learning, and which is used for recording and reporting the amount of learning but not for feeding back into learning (Harlen, 2005, p. 208).

E-assessment

E-assessment includes tests that are delivered on-screen, as well as other assessment instruments – in particular e-portfolios. Also, the review encompasses e-learning technologies (such as virtual learning environments – VLEs – and components thereof such as electronic discussion boards, forums and so on).

Cognate terms for e-assessment are included in this review, including: computer-based assessment (CBA) and computer-assisted assessment (CAA). Further, some articles included in the review might not talk about e-

assessment at all. They refer to e-portfolios or e-learning courses, and the use of these technologies for FA purposes.

Research aims

Thus, the aims of this research are:

- To identify types of implementations that have been used frequently by researchers working in the field.
- Having described key features of implementation types, to suggest ways in which the body of research evidence might be expanded.

In describing eFA implementations and areas for potential further study, an underlying aim is to delineate those areas where eFA provides a distinctive input when compared to formative assessment research more generally[1].

Method and scope

This review is inclusive, rather than excluding. It attempts to provide a 'best evidence synthesis' and results that are authentic, faithful and convincing (Black & Wiliam, 2003, p. 629), rather than complying with one or more 'objective' criteria.

This is a thorough review of eFA literature. It is backed up by a selective review of formative assessment literature. It is not a general review of e-assessment[2].

Data

109 papers have been considered for this review. Their distribution between eFA and 'plain' FA is shown in the table below:

Formative use of e-assessment	73
Plain formative assessment	25
General policy of e-assessment	11
Total	**109**

Table 1: Number of papers in review of different types

The clear majority of the papers related to eFA. A substantial minority described issues in 'plain' FA research. A third category of 11 papers was also discerned (see, for instance: Bennett, 1998; Bennett, 2002; Wainer,

[1] In implementing this 'background aim', regard is had both to those thinkers who suggest that e-assessment will be a necessarily transformative technology (e.g. Bennett, 1998 and 2002), and to more sceptical commentators – who point out how supporters of new technologies have often overstated their potential, and that adoption of such has often led to unexpected consequences (Cuban, 2001).

[2] There are several comprehensive reviews of e-assessment: Ridgway et al, 2004; Sim et al, 2004; Conole & Warburton, 2005.

2000). These papers were early works discussing the potential of e-assessment to transform education; in particular, to facilitate a closer link between assessment and learning.

Background findings on formative assessment

Black and Wiliam (1998b) summarises a fuller description of a comprehensive literature review (Black and Wiliam, 1998a). It poses, and then answers, some questions, including:

- Is there evidence that improving formative assessment raises standards?
- Is there evidence about how to improve formative assessment?

Black and Wiliam (1998b) concludes that there is evidence of substantial learning gains from formative assessment. Further, FA is particularly effective at helping lower-achieving pupils.

Elwood has questioned whether claims for formative assessment's effectiveness in improving learning have been overstated. She suggests that learning gains may be partly accounted for by error variance in test scores, and that gains of learners in FA studies may result from sources other than the FA intervention (Elwood, 2006, p. 227).

Black and Wiliam (1998b) describe how to improve FA practice:

- Feedback to any pupil should be about the particular qualities of his or her work, with advice on what he or she can do to improve, and should avoid comparisons with other pupils.
- For formative assessment to be productive, pupils should be trained in self-assessment so that they can understand the main purposes of their learning and thereby grasp what they need to do to achieve.
- Opportunities for pupils to express their understanding should be designed into any piece of teaching, for this will initiate the type of interaction in which formative assessment aids learning.
- The dialogue between pupils and a teacher should be thoughtful, reflective, focused to evoke and explore understanding, and conducted so that all pupils have an opportunity to think and to express their ideas.

Thus, formative assessment has several aspects – concerning the nature of classroom interactions between teachers and learners (including the way that questions are asked and answered), peer- and self-assessment and the nature of written feedback.

Feedback is a central issue in FA (Sadler, 1989; Sadler, 1998). This includes both the way that teachers interact with pupils in speech, and the nature of written feedback. Written comments are more effective when they are

specific (e.g. not just saying 'an excellent effort') and when they permit a pupil to 'close the gap' between current and desired performance.

There is controversy as to whether written feedback should contain a mark or grade. Black and Wiliam (1998a; 1998b) state that written comments should not contain a mark or grade. Effectively implemented 'comment-only' marking is more likely to give pupils the necessary information to close the learning gap, whereas recipients are more likely to focus on marks or grades at an emotive level (as a comment on their personal worth) rather than as providing a spur to improve work.

Smith and Gorard (2005) cautiously reported an implementation of comment-only marking that did not work as Black and Wiliam would have predicted[3]. In Smith and Gorard's small study, pupils receiving comment-only feedback made inferior progress to that of other classes.

Most FA research has been about a range of classroom practices rather than evaluating assessment instruments and questions. However, Wiliam (2005) proposed that good FA items might have the following properties, which are different to those for good summative assessment items:

- Can have more (or less) than one correct answer
- Items need to be generative
  - of learning
  - of insights into learning
  - of insights into how to promote learning
- Distractors must be explicitly connected to incorrect or incomplete conceptions (facets)
- Item responses must provide clues to effective action

Thus, FA research has examined an area of interest in some depth, and has established some fairly clear principles. There are some reservations about the extent to which reported gains represent genuine effects and a feeling that there needs to be a deeper understanding of the effects of error variance in assessment scores; this is quite a common concern in assessment research. Also, the ways in which clear principles are interpreted when rolled out across an educational sector remain worthy of further study.

These two *caveats* are worth bearing in mind when considering eFA research.

[3] Black et al (2005) attempted to rebut Smith and Gorard's tentative findings.

Review findings: eFA implementations and areas for further research

In the following part of the paper, common implementations of eFA are presented and elaborated. Then, they are critiqued and suggestions for further research are made.

Finding 1

Electronic technologies provide a range of new tools that classroom teachers can use to create formative assessments to suit their and their students' needs.

Many eFA implementations used different task or item types and varied assessment designs. These include:

- Variations on the theme of multiple-choice questions (MCQs):
  - 'formative quizzes' (Cassady & Grindley, 2005)
  - MCQ-based tests made available to students for frequent use (Baggott & Rayne, 2001; Peat & Franklin, 2002)
  - MCQ tests provided for students to allow them to practise the format of the final exam (Cassady et al, 2001; Peat et al, 2005) or as revision (Irving et al, 2000)
  - MCQs adapted to allow students to indicate how confident they are in a particular answer before giving it (Farrrell et al, 2005; Gardner-Medwin & Gahan, 2003)

- More advanced or 'sophisticated' (Boyle, 2005) e-assessment tasks – including those rich in interactivity and multimedia:
  - Scenario-based assessments (Crisp & Ward, 2005)
  - Simulation-based assessments (Young & Cafferty, 2003)
  - Concept maps used for formative assessment of collaborative problem solving (Hsieh & O'Neill, 2002)

- Test designs that are specific to e-assessment[4]:
  - Computer Adaptive Testing (Lilley et al, 2004; Lilley et al, 2005; Yong & Higgins, 2004)

- The use of e-portfolios to facilitate closely integrated formative and summative assessment (McGuire et al, 2004; McGuire, 2005; Woodward & Nanlohy, 2004).

[4] Or at least can be done much more efficiently electronically.

- The use of communications tools such as electronic discussion boards and forums for self- and peer feedback in e-learning courses (Keppell, & Carless, 2006; Keppell et al, 2006; Lin et al, 2001).

These examples perhaps support Bennett's (2001) contention that e-assessment will give rise to mass customisation of assessment products; that is, the ability of educational practitioners to use technologies to provide assessment solutions to suit their particular teaching and learning needs.

However, the early usage of e-assessment instruments for formative purposes also gives rise to areas in need of clarification. These are set out below.

Finding 1: Area for further research (a)

'Plain' FA research has suggested that formative and summative test questions may have different ideal characteristics. However, existing implementations of eFA have tended to take item and task types that originate from summative assessment. New research should attempt to establish the ideal characteristic of items and tasks used for eFA.

'Plain' FA research has not focused much on the nature of test instruments used. eFA provides a range of instruments that practitioners may find useful. However, many early implementations have simply applied summative test and question designs to the formative arena. This may be appropriate, but an interesting new strand of research might build upon Wiliam's contrasting of different properties of good formative and summative items and suggest distinctive features of good eFA items.

Finding 1: Area for further research (b)

eFA implementations have not sufficiently distinguished notions of 'formative assessment' from 'exam revision' or 'becoming acquainted with summative test formats'. Future research should make that distinction more clearly.

The body of FA knowledge has a range of facets. However, several of the eFA papers equate exam revision or practice testing with FA. This is not to say that exam revision is a bad thing; it has a role to play in decreasing students' test anxieties (Cassady & Gridley, 2005) and frequent use of e-assessment quizzes can help students learning from distance to remain motivated and focused (Baggott & Rayne, 2004). Nonetheless, the danger of equating eFA with exam revision is that it will represent a reduced notion when compared to the complete body of formative assessment research.

Finding 1: Area for further research (c)

Early implementations of eFA tended to involve innovators developing their own questions. Further research should investigate whether it is realistic for all teachers to write test questions for eFA or whether – if

teachers merely select from a bank of questions – anything is lost by that process.

Early enthusiasts have developed eFA systems by writing their own questions. It is debatable whether the wider body of teachers would have the necessary time, motivation and skills to write large numbers of high quality test questions.

If teachers using eFA do not write their own questions, an alternative might be for them to use products that contain pre-written questions. Further research might fruitfully investigate the implications of using such eFA products. For example, would the use of a pre-written bank decrease a teacher's ability to tailor questions to suit the needs of learners in her own class?

Finding 2

e-assessment functionality permits formative feedback to be given in a variety of ways that is not possible in 'plain' FA.

Developers of eFA systems have found a range of ways to deliver formative feedback, including:

- Formative feedback given differentially for entirely correct, partially correct and entirely incorrect answers (Wood and Burrow, 2002)
- Feedback as references to textbook chapters (Buchanan, 2000)
- Feedback realised as rich multimedia (Mackenzie, 2000)
- Feedback as references to web sites (Mackenzie, 2003; Clarke et al, 2004)
- Feedback delivered within questions (CIAD, 2005) after each question, or at the end of each timed session (Baggott and Rayne, 2001)
- Rich-media feedback as a stimulus to peer-to-peer discussion of content (Mackenzie, 2003)

Advocates of e-portfolio systems have suggested several advantages that can accrue when e-portfolios are used to provide feedback. These include:

- e-portfolio authoring encourages teachers and students to view drafts of work, and interact about them. The process of generating work is forefronted, rather than merely concentrating on the final product (Twining et al, 2006, p. 55).
- Tools in e-portfolios can allow teachers to ask students to upload materials at significant stages, thus illustrating what the students believe to be progress (an important element of self-assessment) (Twining et al, *ibid.*; McGuire et al, 2004, p. 4).
- Communications tools associated with e-portfolios can allow for the provision of varied feedback with respect to: authors (fellow students or teachers), formality, and mode of communication – writing or speech (McGuire, 2005, p. 267). Such variety can be useful for facilitating

learning by students of different dispositions, experiences and cognitive styles.

Researchers have reported their uses of online e-learning technologies – often on distance learning courses. They have described how communications technologies (such as message boards and discussion forums) have allowed them to provide innovative feedback to assist learning, including:

- Students taking part in online discussions, and being required to submit a specified number of contributions (Goodfellow & Lea, 2005) – a form of peer feedback
- Students keeping a reflective journal (Keppell & Carless, 2006) – feedback to oneself or self-assessment
- Students rating peers' work quite formally – including giving marks (Bhalerao & Ward, 2001) or less formally (Lin et al, 2001), including taking part in collaborative group activities (MacDonald, 2004)

Thus, practitioners have used a range of e-assessment technologies to provide feedback to students. However, there also remain questions arising from these implementations, which may allow researchers to theorise the use of eFA to provide feedback more comprehensively.

Finding 2: Area for further research (a)

Where teachers use extensions to e-test delivery systems to provide feedback to students, further research should establish principles for the design of such feedback so as to optimise students' learning opportunities.

Several researchers have attempted to systematise understanding of the qualities of effective feedback when using e-tests (e.g. Hanson et al, 2001; Hseih and O'Neill, 2002; Clarke et al 2004; Brettell et al, 2005). However, questions remain to be resolved, including:

- Does the stricture from 'plain' FA that feedback should be made up of comments but not grades apply when e-tests are used? If so, does this disable one of the most obvious uses of an e-assessment system for formative purposes?
- To what extent is engagement with rich media or interactive feedback synonymous with deep learning? Or are there circumstances where varied media or interactive possibilities distract learners and lead to superficiality (e.g. clicking through links without truly processing the content of web pages – see Clarke et al, 2004)?

- Is the impact of feedback related to students' learning styles? For example, the work of Brettell et al (2005) to distinguish responses to feedback of 'deep', 'strategic' and 'surface apathetic' learners could profitably be extended.

Finding 2: Area for further research (b)

Where e-portfolios are used with the aim of facilitating the giving of feedback (teacher-to-student; student-to-other-student and student-to-self), logistical or ergonomic studies should be conducted to make sure that users find it practical to give feedback via the portfolio tools.

McGuire (2005) noted that e-portfolios were not 'an easy option', but asserted that they were worthwhile in that they allowed the giving of rich feedback. It will be important to ensure that this potential is not lost; teachers can find it burdensome to provide comments of sufficient quality on students' work (Smith & Gorard, 2005). ICT elements of portfolios should reduce this burden, and thus facilitate the giving of high-quality feedback.

Finding 2: Area for further research (c)

Where online tools such as discussion boards and electronic forums are used to facilitate feedback, research should investigate the impact of cultural factors on students' ability to give peer feedback.

Students giving feedback via electronic tools may suffer if they do not understand cultural norms relating to the giving of feedback. This may have two facets; many online distance learning courses will involve students from different parts of the world. Such students may have differing prior assumptions about commenting on colleagues' work. This may be accentuated when they are working remotely and thus have fewer opportunities to interact face-to-face with peers and/or teachers.

Misunderstanding cultural norms can occur when students are from different countries. However, it can also occur when students have not internalised the norms associated with academic discourses. In particular, early thinking on electronic communication asserted that new communication forms blurred the boundaries between writing and speech – e.g. writing with reduced formality and increased interactivity would be more like speech (Lawler & Dry, 1998). However, giving written feedback on peers' work in an electronic environment is a novel discourse form, and its relationship to formal academic writing remains to be established (Russell et al, 2006). Further research could set out similarities and differences in these two ways of writing and help students to effectively switch between the two.

Finding 3

eFA applications can be used remotely in time (asynchronously). This facility of electronic tools provides a resource which is not easily replicated via pencil-and-paper materials.

Some papers in the review present implementations in which students have been able to go away and use formative assessment materials. Many of the reported studies involved Higher Education classes – often those with new undergraduates. The asynchronicity afforded by electronic materials was said to have the following advantages:

- The use of remote self-access formative assessment materials was associated with reduced examination stress (Baggott and Rayne, 2001; Cassady et al, 2001; Cassady & Gridley, 2005).
- The eFA materials were popular with students and motivating (Blayney & Freeman, 2003).
- The provision of eFA materials freed up teachers' time and thus facilitated courses with high student:teacher ratios (Peat et al, 2005).
- The use of self-assessment eFA materials allowed students to increase their self-regulation (Brettell et al, 2005), in particular to get used to learning independently in tertiary study (Peat et al, 2005).
- The asynchronous aspect of online discussions, added to the fact that evidence of discussion content could be reviewed (e.g. by looking at 'threads' of groups on a web site), facilitated participants' enhanced reflection (Russell et al, 2005).

However, some researchers have noted areas that require clarification.

- There appears to be some relationship between learners' cognitive styles and or their motivations and their use of electronic self-assessment materials. In particular, those who are already skilled in self-regulation may get more benefit from the materials than those who are not (Lin et al, 2001). Also, usage patterns may differ between those learners who are intrinsically interested in learning for its own sake and 'pragmatists' (Keppell & Carless, 2006).
- There are varying results with respect to usage patterns of asynchronous eFA materials. Some researchers report that students used the materials throughout their courses (Bryan et al, 2005), whilst others found usage was concentrated in the period running up to the summative assessment (Pitt & Gunn, 2004).

In addition to those reservations about the corpus of research evidence on the asynchronous use of eFA materials, the current review adds two further areas that should be clarified so that research evidence is more complete.

Finding 3: Area for further research (a)

Although several studies have claimed that use of eFA materials is associated with learning gains, the bases on which they do so are generally not well founded. If a claim is to be made that eFA provides enhanced learning gains over and above 'plain' FA, then better designed studies need to be conducted.

A substantial number of the eFA papers in this review (especially those that reported on the asynchronous/self-access use of formative materials) claimed that students who used the materials had an attainment benefit. However, in almost all cases these claims were undermined by an aspect of the research design. For instance, studies were conducted with small cohorts, or the difficulty of two years' tests was not properly equated or studies confounded variables (e.g. did the students using eFA score more highly because it was an eFA intervention, or did they score more highly because they worked harder?).

Thus, an important claim of the eFA literature has not been robustly established. That 'plain' FA is associated with learning gains is an important tenet of that literature, but it might be interesting for researchers to design studies that build from the work of plain FA researchers and show particular ways in which eFA supports enhanced attainment.

Finding 3: Area for further research (b)

The equating of eFA with self-assessment is strongly associated with patterns of learning in tertiary education. It would interesting to see whether the self-access paradigm could be imported into secondary or primary education.

The literature reporting the asynchronous use of eFA materials is strongly associated with tertiary education. Taking online quizzes and the like is seen as a way to encourage new undergraduates to manage their study in an environment where they were expected to take more responsibility than at school.

It would be useful to see what issues would crop up if e-self-assessment materials were widely used by school-age students. For example, school teachers may feel a greater obligation to moderate feedback (e.g. to avoid students receiving potentially demotivating critical feedback). Other issues not apparent in the tertiary sector might also arise (e.g. the role of parents in supporting their children's online learning).

Discussion

The aim of this review was to map implementations of eFA, and to suggest areas for further research. In doing so, the intention was also to describe those areas where the use of e-assessment for formative purposes provided a distinctive contribution; different to anything that came from the wider body of formative assessment research.

Starting from implementation has the virtue of being a 'reality check'; giving an overview of the state of the art at a particular point. It affords the possibility of description of actual practice. Evaluation of that practice can then suggest the extent to which implementations have fulfilled aspirations for eFA. It can also facilitate a re-focusing on areas that need increased attention; especially if such areas are unexpected.

However, it may be that working from implementations can give a somewhat fragmented picture of the unique features of formative e-assessment. For that reason, attention has been paid to the nine 'areas for further research' that have been proposed in this review. These have been examined to search for commonality between them.

In fact, there does appear to be some commonality between the nine areas for further research, and so it is possible to propose four 'super categories' or key issues that might guide future eFA research.

Key issue 1

eFA research needs to better define the ways in which the electronic element provides added benefit above and beyond 'plain' FA use.

This key issue requires thinking about eFA to demonstrate its added value beyond plain FA. Also, however, it would critique eFA implementations that simply adopted summative e-assessment designs without showing their suitability for the formative purposes. The key issue arises from the following areas for further research:

- 1a: use of e-assessment instruments by practitioners
- 2a: provision of feedback from e-assessment instruments
- 3a: need for better-designed studies to demonstrate attainment benefits

Key issue 2

Those promoting eFA implementations should ensure that eFA does not amount to a reduced or impoverished notion when compared to the full understanding of formative assessment.

This key issue is – in some senses – the converse of the first. However, it goes somewhat further; whilst key issue 1 imposes a positive duty on eFA to show distinctive benefit, this key issue notes the possibility that eFA can have negative consequences. It arises from the following areas for further research:

- 1b: equating FA with exam revision
- 2a: provision of feedback from e-assessment instruments

Key issue 3

Attention should be given to the danger that eFA might impose new burdens on teachers (and – to some extent – students).

ICT innovations are often touted as labour saving. However, if they are not well designed (or specifically fit for an educational purpose, Cuban (2001, p. 170)), they may not be as widely adopted as expected.

This key issue arises from the following areas for further research:

- 1c: requirement for teachers to write their own test questions
- 2b: need for e-portfolios to provide manageable systems for giving feedback

Key issue 4

Students using eFA applications will sometimes be required to take on novel roles. The ways in which students adapt to such novel roles should be monitored.

Students may need to work more independently than previously, or to communicate according to cultural or social norms which are alien to them. The extent to which they are successful in so adapting could be an important area of eFA research.

This key issue arises from the following areas for further research:

- 2c: cultural factors in the use of electronic communications tools
- 3b: strong element of independent working and self-assessment in eFA

Different sets of key issues may be arguable, but it is proposed that if eFA research were to focus on these four areas, then it would be stronger, and have a chance of leading to more principled implementations.

Bibliography

Baggott, G. & Rayne, R. (2001) *Learning support for mature, part-time, evening students: providing feedback via frequent, computer-based assessments* **in** Danson, M. (ed.) Fifth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, July 2001. http://www.caaconference.com/.

Baggott, G. & Rayne, R. (2004) *Student perceptions of computer-based formative assessments in a semi-distance module* **in** Ashby, M. (ed.) Eighth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, 6th and 7th July 2004. http://www.caaconference.com/.

Bennett, R.E. (1998) *Reinventing Assessment: speculations on the future of large-scale educational testing*. ftp://ftp.ets.org/pub/res/reinvent.pdf.

Bennett, R.E. (2001) *How the internet will help large-scale assessment reinvent itself*. Education Policy Analysis Archives, Volume 9 Number 5. http://epaa.asu.edu/epaa/v9n5.html.

Bennett, R.E. (2002) *Inexorable and Inevitable: The Continuing Story of Technology and Assessment*. The Journal of Technology, Learning and Assessment (JTLA), Volume 1, Number 1.

Bhalerao, A. & Ward, A. (2001) *Towards electronically assisted peer assessment: a case study*. Association for Learning Technology Journal (ALT-J) 9(1) pp. 26 – 37.

Black, P. & Wiliam, D. (1998a) *Assessment and classroom learning*. Assessment in Education: principles, policy & practice, 5(1), pp. 7 – 73.

Black, P. & Wiliam, D. (1998b) *Inside the Black Box: Raising Standards Through Classroom Assessment.* (London: King's College London School of Education).

Black, P. & Wiliam, D. (2003) *'In praise of educational research': formative assessment*. British Educational Research Journal, Volume 29, Number 5, pp. 623 – 637.

Black, P., Harrison, C., Hodgen, J., Marshall, B. and Wiliam, D. (2005) *The dissemination of formative assessment: a lesson from or about, evaluation*. Research Intelligence, 2005, p. 12.

Blayney, P. & Freeman, M. (2003) *Automated marking of individualised spreadsheet assignments: the impact of different formative self-assessment options* **in** Christie, J. (ed) Seventh International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, July 2003. http://www.caaconference.com/.

Boyle, A. (2005) *Sophisticated Tasks in E-Assessment: What are they? And what are their benefits?* **in** Danson, M. (ed) Ninth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, 5th and 6th July 2005. http://www.caaconference.com/.

Brettell, S., Durham, J. & McHanwell, S. (2005) *'Well nobody reads learning outcomes do they?' – An evaluation of CAA and its feedback on directed student learning* **in** Danson, M. (ed) Ninth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, 5th and 6th July 2005. http://www.caaconference.com/

Bryan, N. & Glasfurd-Brown, G. (2005) The SPRinTA project: supporting student assessment through a portal **in** Danson, M. (ed) Ninth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, 5th and 6th July 2005. http://www.caaconference.com/.

Buchanan, T. (2000) *The efficacy of a World-Wide Web mediated formative assessment*. Journal of Computer Assisted Learning Volume 16 Issue 3, p. 193.

Cassady, J.C. & Gridley, B.E. (2005) *The Effects of Online Formative and Summative Assessment on Test Anxiety and Performance*. Journal of Technology, Learning and Assessment (JTLA), Volume 4, Number 1, October 2005.

Cassady, J.C., Budenz-Anders, J., Pavlechko, G. & Mock, W. (2001) *The effects of internet-based formative and summative assessment on test anxiety, perceptions of threat, and achievement*. Paper presented at the Annual meeting of the American Educational Research Association (AERA) (Seattle, Wa., April 10 - 14 2001).

Centre for Interactive Assessment Development (CIAD) (2005). *TRIADS Functionality for Formative Assessment.* http://www.derby.ac.uk/ciad/formative.php.

Charman, D. & Elmes, A. (1998). *A computer-based formative assessment strategy for a basic statistics module in geography.* Journal of Geography in Higher Education, 22(3), pp. 381-385.

Clarke, S. Lindsay, K., McKenna, C. & New, S. (2004) *INQUIRE: a case study in evaluating the potential of online MCQ tests in a discursive subject*. Association for Learning Technology Journal (ALT-J), Volume 12, Number 3, September, pp. 249 – 260.

Condie, R. & Munro, B. (2007) The impact of ICT in schools – a landscape review.
http://publications.becta.org.uk/display.cfm?resID=28221&page=1835.

Conole, G. & Warburton, B. (2005) *A review of computer-assisted assessment.* Association for Learning Technology Journal (ALT-J), Volume 13, Number 1, March, pp. 17-31.

Crisp, V. & Ward, C. (2005) *The PePCAA project: formative scenario-based CAA in psychology for teachers* **in** Danson, M. (ed) Ninth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, 5th and 6th July 2005. http://www.caaconference.com/.

Crooks, T.J. (1988) *The impact of classroom evaluation practices on students*. Review of Educational Research, 58, pp. 438 – 481.

Cuban, L. (2001) *Oversold and underused: computers in the classroom.* Cambridge, Ma.: Harvard University Press.

Elwood, J. (2006) *Formative assessment: possibilities, boundaries and limitations*. Assessment in Education: Principles, Policy & Practice, Vol. 13, No. 2, July 2006, pp. 215–232.

Farrell, G., Farrell, V. & Leung, Y.K. (2005) *A comparison of Blackboard CAA and an innovative self-assessment tool for formative assessment* **in** Danson, M. (ed) Ninth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, 5th and 6th July 2005. http://www.caaconference.com/.

Gardner-Medwin, A.R. & Gahan M. (2003) *Formative and summative confidence-based assessment* **in** Christie, J. (ed) Seventh International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, July 2003. http://www.caaconference.com/.

Goodfellow, R. & Lea, M.R. (2005) *Supporting writing for assessment in online learning*. Assessment & Evaluation in Higher Education, Volume 30, Number 3 / June 2005, pp. 261 – 271.

Hanson, J., Millington, C. & Freewood, M. (2001) *Developing a methodology for online feedback and assessment* **in** Danson, M. (ed.) Fifth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, July 2001. http://www.caaconference.com/.

Harlen, W. (2005) *Teachers' summative practices and assessment for learning – tensions and synergies*. The Curriculum Journal, Vol. 16, No. 2, June 2005, pp. 207 – 223.

Hsieh, I.L.G. & O'Neil, H.F. Jr. (2002) *Types of feedback in a computer-based collaborative problem-solving group task*. Computers in Human Behavior, Volume 18, Issue 6, November, Pages 699 – 715.

Irving, A., Read, M., Hunt, A. & Knight, S. (2000) *Use of information technology in exam revision* **in** Danson, M. (ed.) Fourth International

Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, July 2000. http://www.caaconference.com/.

Keppell, M. & Carless, D. (2006) *Learning-oriented assessment: a technology-based case study*. Assessment in Education Vol. 13, No. 2, July 2006, pp. 179–191.

Keppell, M., Au, E. Ma. A. & Chan, C. (2006) *Peer learning and learning-oriented assessment in technology-enhanced environments*. Assessment & Evaluation in Higher Education, Volume 31, Number 4 / August 2006, pp. 453 – 464.

Lawler, J. & Dry, H.A. (2008) *Using computers in linguistics: a practical guide.* London: Routledge.

Lilley, M. Barker, T. & Britton, C. (2004) *The generation of automated student feedback for a Computer-Adaptive Test* **in** Ashby, M. (ed.) Eighth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, 6th and 7th July 2004. http://www.caaconference.com/.

Lilley, M. Barker, T. & Britton, C. (2005) *Automated feedback for a Computer-Adaptive Test: a case study* **in** Danson, M. (ed) Ninth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, 5th and 6th July 2005. http://www.caaconference.com/.

Lin, S.S.J., Liu, E.Z.F. & Yuan, S.M. (2001) *Web-based peer assessment: feedback for students with various thinking-styles*. Journal of Computer Assisted Learning, Volume 17 Issue 4 p. 420.

Macdonald, J. (2004) *Developing competent e-learners: the role of assessment*. Assessment & Evaluation in Higher Education, Volume 29, Number 2 / April 2004, pp. 215 – 226.

Mackenzie, D. (2000) Production and delivery of TRIADS Assessments on a university-wide basis **in** Danson, M. (ed.) Fourth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, July 2000. http://www.caaconference.com/.

Mackenzie, D. (2003) *Assessment for E-learning: what are the features of an ideal E-assessment system?* **in** Christie, J. (ed) Seventh International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, July 2003. http://www.caaconference.com/.

McGuire, L. (2005) *Assessment using new technology*. Innovations in Education and Teaching International, Volume 42, Number 3, pp. 265 – 276.

McGuire, L., Roberts, G. & Moss, M. (2004) *Final report to QCA on the eVIVa project*. http://210.48.101.74/images/Eviva%20Final%20Report.pdf.

Natriello, G. (1987) *The impact of evaluation processes on students*. Educational Psychologist, 22, pp. 155 – 175.

Peat, M. & Franklin, S. (2002) *Supporting student learning: the use of computer-based formative assessment modules*. British Journal of Educational Technology, v.33 n5, p. 515 – 523.

Peat, M., Franklin, S., Devlin, M. & Charles, M. (2005) *Revisiting the impact of formative assessment opportunities on student learning*. Australasian Journal of Educational Technology, 21(1), 102 – 117.

Pitt, S.J. & Gunn, A. (2004) *The value of computer based formative assessment in undergraduate biological science teaching*. Bioscience Education e-Journal, Volume 3: May 2004.

Prins, F.J., Sluijsmans, D.M.A., Kirschner, P.A. & Strijbos, J-W. (2005) *Formative peer assessment in a CSCL environment: a case study*. Assessment & Evaluation in Higher Education, Volume 30, Number 4 / August 2005, pp. 417 – 444.

Ridgway, J., McCusker, S. and Pead, D. (2004) *Literature review of e-assessment*. http://www.nestafuturelab.org/research/reviews/10_01.htm.

Russell, J., Elton, L., Swinglehurst, D. & Greenhalgh, T. (2006) *Using the online environment in assessment for learning: a case-study of a web-based course in primary care*. Assessment & Evaluation in Higher Education, Volume 31, Number 4 / August 2006, pp. 465 – 478.

Sadler, D.R. (1989) *Formative assessment and the design of instructional systems*. Instructional Science, 18, pp. 119 – 144.

Sadler, D.R. (1998) *Formative assessment: revisiting the territory*. Assessment in Education: Principles, Policy & Practice, 5(1), pp. 77 – 84.

Sim, G., Holifield, P. and Brown, M. (2004) *Implementation of computer assisted assessment: lessons from the literature.* Association for Learning Technology-Journal (ALT-J), **12** (3) 215 – 229.

Smith, E & Gorard S. (2005) *'They don't give us our marks': the role of formative feedback in student progress*. Assessment in Education: Principles, Policy & Practice, Volume 12, Number 1, pp. 21 – 38.

Twining, P., Broadie, R., Cook, D., Ford, K. Morris, D. Twiner, A. & Underwood, J. (2006) *Educational change and ICT: an exploration of Priorities 2 and 3 of the DfES e-strategy in schools and colleges (The current landscape and implementation issues).* http://partners.becta.org.uk/page_documents/research/educational_change_and_ict.pdf.

Wainer, H. (2000) *CATs: whither and whence*. Psicológica 21, pp. 121 – 133.

Wiliam, D. & Thompson, M. (2006) *Integrating assessment with learning: what will it take to make it work?* **in** Dwyer, C. A. (ed) The future of assessment: shaping teaching and learning. Mahwah, NJ: Lawrence Erlbaum Associates.

Wiliam, D. (2005) *Formative assessment and the regulation of learning*. Paper presented at UC Berkeley seminar, March 2005, Berkeley, CA.

Wood, J. & Burrow, M. (2002) Formative Assessment in Engineering Using 'TRIADS' Software **in** Danson, M. (ed) Sixth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, 5th and 6th July 2005. http://www.caaconference.com/.

Woodward, H. & Nanlohy, P. (2004) *Digital portfolios: fact or fashion?* Assessment & Evaluation in Higher Education, Volume 29, Number 2 / April 2004, pp. 227 – 238.

Yong, C-F. & Higgins, C.A. (2004) *Self-assessing with adaptive exercises* **in** Ashby, M. (ed.) Eighth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, 6th and 7th July 2004. http://www.caaconference.com/.

Young, A. & Cafferty, S. (2003) *Simulation as a tool for computer-assisted formative assessment: First aid as a case study* **in** Christie, J. (ed) Seventh International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, July 2003. http://www.caaconference.com/.

All web links were live on 20 February 2007.

Appendix 1: Sources consulted in research

Research databases and specialist search engines

Research databases

- Education Resources Information Center (ERIC) (http://www.eric.ed.gov/)
- Bibliography on Computer Based Assessment and Distance Learning (http://liinwww.ira.uka.de/bibliography/Misc/cba.html)
- EBSCO Host Academic Search premier

Specialist search engines

- http://scholar.google.com/
- http://www.scirus.com
- http://citeseer.ist.psu.edu/

Journals

Comprehensively handsearched journals

- Assessment in Education: Principles, Policy & Practice
- Assessment & Evaluation in Higher Education
- Journal of Computer Assisted Learning (JCAL)
- Association for Learning Technology Journal (ALT-J)
- Journal of Technology, Learning and Assessment (JTLA)
- British Journal of Educational Technology (BJET)
- Research Papers in Education
- British Educational Research Journal (BERJ)
- Curriculum Journal

Other journals that provided articles for this project include

- Australasian Journal of Educational Technology
- Bioscience Education e-Journal
- CAL-elaborate
- Cambridge Journal of Education
- Computers in Human Behavior
- Educational Psychologist
- Engineering Education
- Innovations in Education and Teaching International
- Innovations in Education and Training International
- Journal of Dental Education

- Journal of Educational Multimedia and Hypermedia
- Journal of Geography in Higher Education
- Learning and Teaching in Higher Education
- Measurement
- Psicológica
- Research Intelligence
- Review of Educational Research
- Studies in Continuing Education
- Teaching Mathematics and Its Applications
- The Internet and Higher Education

Conference archives

- Computer-assisted Assessment (CAA) conference (http://www.caaconference.com/)
- Association for Educational Assessment – Europe (http://www.aea-europe.net/)
- International Association for the Evaluation of Educational Achievement (IAEA) (http://www.iaea.info/index.php?option=com_conferences&Itemid=45)

Other conferences provided articles for the research, but they did not have comprehensive central archives of papers.

Other sources of information

Review articles that provided references

- Ridgway et al (2004)
- Conole and Warburton (2004)
- Sim et al (2005)

Lists of ‘key ‘plain formative assessment’ papers’ provided by:

- Bill Boyle, Centre for Formative Assessment Studies (CFAS), University of Manchester
- Paul Newton, Qualifications and Curriculum Authority.