

**FINDING APPROPRIATE METHODS TO
ASSURE QUALITY COMPUTER-BASED
ASSESSMENT DEVELOPMENT IN UK
HIGHER EDUCATION**

Andrew Boyle and David O'Hare

Finding Appropriate Methods to Assure Quality Computer-Based Assessment Development in UK Higher Education

Andrew Boyle,
Assessment and Measurement Department,
National Foundation for Educational Research,
The Mere,
Slough,
Berkshire, SL1 2DQ.
Telephone: 01753 574 123
Email: a.boyle@nfer.ac.uk
Web site: <http://www.nfer.ac.uk/>

Dr Dave O'Hare,
Centre for Interactive Assessment Development,
University of Derby,
Kedleston Rd,
Derby, DE22 1GB.
Telephone: 01332 591 741
Email: d.ohare@derby.ac.uk
Web site: <http://www.derby.ac.uk/ciad/>

Abstract

Over the past decade Computer-Based Assessment (CBA) has become widely used in Higher Education (HE) in the UK (Bull and Hesketh, 2001). Most implementations of CBA to date have been item-based (Zakrewski and Bull, 1998). Item-based assessment demands different types of Quality Assurance (QA) to traditional essay- and performance-based assessment. Specifically, item-based assessment forefronts QA in the assessment development phase.

Appropriate agencies have stipulated the necessary quality standards for assessment in HE (see QAA, 2000). But, in contrast to other educational sectors in the UK, mandatory assessment development procedures are not set out in detail (cp. QCA *et al.*, 2002). Standards for assessment development have also been published in other countries (AERA *et al.*, 1999), and guidelines specifically aimed at CBA have been published (ATP, 2000). A set of core tasks to assure quality in assessment development can be discerned from these diverse sources.

The particular context of UK Higher Education Computer-Based Assessment may mean that it is not possible to carry out all the core tasks for quality assessment development. However, situational factors do not make quality assessment development impossible. In this paper, we will outline factors that mean that developers of CBAs in HE must modify widely-advocated QA procedures, and suggest alternative but equivalent ways to assure assessment quality.

The use of Computer-Based Assessment

Computer-Based Assessment is a relatively new innovation in United Kingdom Higher Education. It is believed that the use of CBA has grown significantly since the early 1990s. Two surveys conducted in the late 1990s addressed Computer-Assisted Assessment (CAA). CAA is a wider concept than CBA, since it includes assessments conducted on paper, but whose processing is significantly assisted by computers. For example, there are multiple-choice assessments in which answers are entered onto Optical Mark Reader (OMR) sheets and then scanned to calculate the number of correct answers in the assessment. Although the two surveys included such OMR assessments, they also provide the only available empirical information on the uptake of Computer-Based Assessment.

The surveys do not demonstrate CAA's growth unequivocally. In January 1997 Stephens and Mascia published the results of a survey of 445 questionnaires returned by academics (Stephens and Mascia, 1997). Their data showed that 300 (67 per cent) of respondents were using CAA in some form. They concluded that the use of CAA in UK HE was growing, and that this delivery method would 'prevail' (1997, p. 26). However, they found that CAA was mainly used in science-related disciplines; and much less in Humanities and Arts (*ibid.* at p. 18, and see also McKenna (2000) summarising findings from the CAA National Survey relevant to Humanities).

The survey that Stephens and Mascia reported in 1997 was followed by the CAA Centre National Survey. McKenna and Bull (1999) report that 44 per cent of the 754 respondents who completed questionnaires (332 actual cases) were using CAA in one way or another.

A naïve interpretation of these findings would be that the uptake of CAA had declined between the 1997 and 1999 surveys. But there were differences in the size, return rates and purposes of the two surveys, which make comparisons difficult. Also, a decline in uptake seems unlikely, given the increasing use of computers for personal and professional purposes in that era (see Bennett, 2001, p. 3), and assertions of the increasing importance of computerised assessment in the literature (for example, Brown *et al.*, 1999, p. 1). Perhaps the best interpretation is that CAA has been shown to be used by a significant number of academics in UK HE, but that a rise in the use of CAA cannot be construed from these surveys. The growth of computerised assessment will also have been aided by the use of Virtual Learning Environments (VLEs) in HE, as these often have simple CBA built in. Jenkins *et al.* (2001) conducted a survey of VLE usage (covering around 70 institutions) and observed that 80 per cent of respondents said that VLEs were being used in their institutions.

The item-based nature of CBA

Traditional assessment practice in UK HE has involved the use of 'essay and problem-type final examinations and similarly constructed coursework' (Elton and Johnston, 2002, p. 7). However, most implementations of CBA to date have used discrete items (that is, assessments consist of a number of short tasks, each of which has a single right answer). Zakrewski and Bull (1998, p. 148) note the time lecturers must devote to developing such items, and assert that designing 'objective' assessments to target 'higher-order cognitive abilities' requires a good deal of 'time, skill and care' (*ibid.* at p. 151). Boyle *et al.* (2002) suggested that the development of high-quality multiple-choice items was a difficult skill to acquire (2002, p. 279).

In so far as the use of 'discrete' or 'objective' items represents a departure for HE, it implies a shift in QA practice. Whilst in essay- or performance-based assessment much of the work goes on after the administration of the assessment (double marking, moderation, etc.), in item-based assessments, the development phase becomes crucially important. This suggests a fundamental change in the process of assessment development in Higher Education, which may have significant implications.

Furthermore, the introduction of new assessment methods in HE may be problematic, as several authors have already noted the need for changes in existing assessment practices (e.g. Race, 1993). Such changes include not only the development of staff involved in assessment (Yorke *et al.*, 2000), but also the management of the assessment process itself (Yorke, 1998). Thus, there is already a clear need for improvement in the conduct of assessment in HE.

Quality assurance in UK HE assessment

The Quality Assurance Agency (QAA) oversees quality and is responsible for maintaining academic standards in HE. Section six of the QAA's Code of Practice for the assurance of academic quality and standards in Higher Education deals with assessment (QAA, 2000). The Code of Practice for the assessment of students consists of *precepts*, and guidance thereon. The *precepts*:

'... identify those key matters that the [QAA] expects an institution to be able to demonstrate it is addressing effectively through its own Quality Assurance mechanisms.' (QAA, 2000, p. 2)

The QAA expects all Higher Education Institutions (HEIs) to demonstrate adherence to the *precepts* (QAA, 2000, p. 3). But the *precepts* are general statements, not specific prescriptions for implementation. In determining adherence to the *precepts*, the QAA looks for 'equivalence of effect' (p. 4), rather than assessment practices found on an exhaustive list.

Nonetheless, the *precepts* are clear. For example, institutions should ensure that:

‘ ... assessment policies and practices are responsive and provide for the effective monitoring of the validity, equity and reliability of assessment.’ (QAA, 2000, p. 6)

A number of documents describe Quality Assurance procedures in UK HE in more detail. The ‘Handbook for Institutional Audit for England’ (QAA, 2002) gives an operational description of the external review process for HE. This document describes ‘institutional audit’, a new process effective from 2002 – 03. It aims to reassure the public that HEIs are:

‘ ... providing Higher Education, awards and qualifications of both an acceptable quality and appropriate academic standard.’ (2002, p. 2)

The *Audit Handbook* stipulates the scrupulous use of external examiners in summative assessment (2002, p. 4), and states that audit visits may involve scrutiny of students’ work (p. 7).

The *Audit Handbook* states that audit trails will also refer to *Subject Benchmark Statements* (QAA, 2003), the *QAA Code of Practice* and the *Framework for Higher Education Qualifications* (QAA, 2001). Whilst this brief paper only mentions these complex documents in passing, the following can be said. None of the QA documents mandate universities to carry out specific steps in assessment development (cp. QCA *et al.* 2002, described below). Furthermore, none of the documents refer specifically to CBA as a particular assessment method whose QA demands are quite different to traditional assessment.

Quality assurance in other UK educational sectors

The Qualifications and Curriculum Authority (QCA)’s mission is as follows:

‘QCA is a guardian of standards in education and training. We work with others to maintain and develop the school curriculum and associated assessments, and to accredit and monitor qualifications in schools, colleges and at work.’ (QCA, 2003)

To support its mission, the QCA (and sister organisations in the UK’s constituent nations) have published a mandatory Common Code of Practice, which applies to high-stakes examinations taken by school students and others. The *QCA Code* specifies:

‘ ... the processes and procedures required to ensure high-quality, consistency and rigorous standards in assessment and awarding across all qualifications’ (QCA *et al.*, 2002, p. 3)

In its specificity and rigour, the *QCA Code* contrasts sharply with the more permissive approach adopted in the QAA *precepts*. This is particularly evident in the enforcement of the *QCA Code*. The ability to operate as an awarding body for the qualifications covered by the *QCA Code* is dependent upon acceptance of, and the ability to implement, the Code. The implementation of the Code is monitored annually by QCA, shortcomings are

identified and assurances sought that they will be rectified (Woodthorpe, personal communication).

The QCA Code contains detailed prescriptions that must be followed in the preparation of assessments and mark schemes (QCA *et al.*, 2002, p. 11 *et seq.*). Assessment institutions are required to have staff performing a number of precisely defined roles (p. 11). The staff have specific duties in the assessment development process. Chief examiners, chairs of examiners and awarding body officers have key roles, and the latter are required to report any issues of concern to the awarding body's senior management (p. 14).

Quality assurance documents from other countries

The 'Standards for Educational and Psychological Testing' have been developed by the American Educational Research Association, the American Psychological Association and the National Council on Measurement in Education (AERA *et al.*, 1999) (hereafter the *Joint Standards*). The purpose of the *Joint Standards* is:

' ... to provide criteria for the evaluation of tests, testing practices and the effects of test use ... ' (AERA *et al.* 1999, p. 2)

The *Joint Standards* seek to encourage developers to provide information about their quality control procedures, to inform public debate, rather than prescribing the technical characteristics of good assessments (*ibid.*). The *Joint Standards* are not mandatory. Rather, it is hoped that responsible organisations will adopt them as good practice (*ibid.*).

The *Joint Standards* (p. 37 *et seq.*) outline an assessment development methodology. Great emphasis is placed on assessment development with a considerable number of qualitative and quantitative control stages. The former range from assessment specification as an *a priori* (p. 38) to item review by skilled and interested parties (p. 39). Statistical analyses are initially based on field testing (sometimes referred to as 'pre-testing' or 'trialling'), and address issues of fairness, and item and instrument quality (*ibid.*).

The clear focus of the *Joint Standards* dictates that reputable assessment developers should provide information about their procedures to assessment users, so that the latter can evaluate the quality of the assessment.

Quality assurance documents that relate specifically to CBA

The *Joint Standards* make only three specific references to Computer-Based Assessment (Olsen, 2000). Therefore, several documents specific to CBA have been published. The Association of Test Publishers (ATP) in the United States has developed Guidelines for Computer-Based Testing (ATP, 2000). The *ATP Guidelines* are intended to 'supplement, extend and elaborate on' the *Joint Standards* (Olsen, 2000, p. 2). In common with the *Joint Standards*, the *ATP Guidelines*:

' ... are written to inform and improve professional practice and judgment, but are not to be viewed as a set of explicit checklists or standards that must be followed in minute detail.' (Olsen, 2000, p. 2)

The *Guidelines* note that planning for a Computer-Based Assessment is similar to planning for other kinds of assessments (Olsen, 2000, p. 6). In planning for the assessment, the purpose must be ascertained (*ibid.*). To ascertain this purpose, needs analysis may be conducted. The first step in the needs analysis would be to carefully describe the assessment-taking population (p. 7). In addition to normal demographic characteristics, the computer literacy of assessment takers should be established (*ibid.*). The degree of computer literacy in a population would be a factor in determining the complexity of question formats, and the nature and extent of help and practice questions for users (*ibid.*).

Assessment developers should have a validation plan. This should address a number of areas; such as, 'construct irrelevant variance', that is, scores must represent the assessment takers' knowledge or skills in the construct being assessed, and not their computer skills (*ibid.*).

The British Standards Institution has introduced a new Standard: BS 7988: 2002 - Code of Practice for the Use of Information Technology in the Delivery of Assessments (BSI, 2002). The Standard introduces guidelines and minimum requirements for organisations using computers to conduct assessments and establishes a Code of Practice for delivering them. However, the preparation of assessment content is outside the scope of the Standard (2002, p. 8); since the main thrust of the Standard is assessment delivery. Thus a poor assessment, delivered appropriately, would conform to BS 7988.

The International Test Commission (ITC) is currently developing draft Guidelines on Computer-Based and Internet-Delivered Testing. The *ITC Guidelines* seek to synthesise earlier QA documents in assessment (ITC, 2003, p. 2) to produce an internationally-recognised set of Guidelines that describe good practice in CBA (*ibid.*). The *ITC Guidelines* do not deal with QA issues that could apply in any assessment medium (*ibid.*), but they do make a number of recommendations relevant to QA in CBA development. Amongst other things, developers are required to document constructs to be measured (p. 10), and to understand current professional and ethical issues (*ibid.*). As in the case of the *ATP Guidelines*, assessment developers should avoid construct irrelevant variance (p. 12), and make appropriate use of psychometric theories and models (*ibid.*).

The Scottish Qualifications Authority (SQA) has produced Guidelines for Online Assessment for Further Education – *the SQA Guidelines* (SQA, 2003). This document advises potential users of online assessment to ensure institutional commitment, to define reasons for using online assessment and to select appropriate delivery methods (pp. 1 – 13). With specific respect to assessment development, the *SQA Guidelines* suggest that the developer should choose item types that are 'appropriate to the knowledge or skill being tested' (p. 14). The *SQA Guidelines* give tips for the development of valid items (pp. 17 – 19), suggest the collection of metadata (p. 19), and the analysis of item quality (p. 21). The *SQA Guidelines* make a clear

recommendation in favour of item banks, sharing questions between institutions and/or re-using items in a number of examination sessions in order to improve item and assessment quality (p. 24). An item bank is a structured set of test items. Although a folder containing pieces of paper can, theoretically, be described as an item bank, most modern systems use sophisticated database software. (For further details on item banks see Umar, 1999.)

Core tasks to assure quality in CBA development

From the documents précised above it is possible to posit a set of steps that must be undertaken if Computer-Based Assessment development is to have the distinguishing characteristics of high quality. This set of stages is listed below:

- An analysis of the need for and a description of the purpose of the assessment;
- Detailed documentation of the instrument(s) to be produced;
- An iterative process to produce assessment content;
- One or more field-testing phases to assure quality prior to live administration;
- Assessment administration under controlled conditions;
- Review of assessment outcomes to consider the fairness and validity of actual outcomes.

Not all the quality control steps are mentioned in all the documents in our review. For example, the *QCA Code of Practice* does not mention a field-testing phase. This may reflect the traditional aversion to pre-testing high-stakes examinations in the UK due to security concerns.

Conversely, not all QA phases from the reviewed documents are included in the minimal list. For example, the *Joint Standards* and the *SQA Guidelines* both refer to item banking as an important element of QA. Whilst item banking can indeed so be, it is not appropriate for all types of assessment and may make specific demands that not all developers would either be able to, or wish to, meet.

Inhibitors of the core tasks

There are several characteristics of assessment practice in UK HE that may inhibit implementation of the core tasks.

Assessment developers' lack of training

All university lecturers have a responsibility to produce assessment. However, commentators have alleged that the standard of training in assessment is low. Race suggests new staff are 'thrown in at the deep end' (1993, p. 7). Elton and Johnston claim many assessors lack key skills, such as the ability to write essay prompts of comparable difficulty (2002, p. 9). This should be contrasted with mandatory training at large-scale institutions that produce assessments governed by the *QCA Code of Practice* (QCA *et al.*, 2002, p. 11).

Inexperience at writing ‘objective’ assessment items

McKenna conducted focus groups and structured interviews, as part of the CAA National Survey. She reports that the shift to new question formats was ‘time consuming initially’ (2001, p. 315), and cites a respondent who said:

‘[Writing multiple-choice questions] is so difficult that you almost have to be a professional at creating these questions.’ (*ibid.*)

The authors of the *QCA Code of Practice*, among others, would maintain that you **do** have to be a professional at producing items, particularly if your assessment has an important impact on students’ lives. Indeed, as assessment becomes more accountable, the lack of professional item-writing skills may become a matter for the courts (cf. Euben, 2001).

CBA has been an *ad hoc* innovation

Often, the impetus for CBA development has come from enthusiastic individuals working alone or in small groups (McKenna and Bull, 2000, p. 24). This *ad hoc* development may have disadvantages. For example, it may be more difficult to embed CBA in a university’s wider QA practices (*ibid.*), and developments may proceed in ‘anarchic fashion’ (*ibid.* at p. 27), and it may not be possible to disseminate the innovation effectively (*ibid.*).

Hindrances to field testing

The set of core tasks includes a field-testing phase, whose purpose is to evaluate (and consequently improve) item and assessment quality prior to live administration. The research methods used in pre-testing are most often quantitative; involving the calculation of standard statistical indices. But, particular features of CBA in UK HE may inhibit the use of pre-testing. Firstly, the fears about security that militate against pre-testing in UK high-stakes school examinations may also exist in HE. Further, the use of statistical techniques depends upon the existence of stable assessment programmes and large groups of students to provide data for analysis. Knight (1999, p. 102, n4) contrasts the position in the USA, which has ‘massive data sets’ that permit sophisticated analysis with the UK, where work takes the form of ‘small-scale craft activities’ (1999, p. 102).

Possible solutions to QA problems

Several ways to address inhibitors to quality are described below.

Institutional support

Several writers agree that provision of institutional support is crucial to the effective implementation of CBA. Stephens *et al.* (1998, p. 287) describe the CAA Support Unit at Luton and the Flexible Learning Initiative at Loughborough. Also, the University of Derby has established the Centre for Interactive Assessment Development (CIAD) – a dedicated unit to advise on all aspects of CBA production (Mackenzie, 2000).

King *et al.* (1998) describe early experiences with a CAA protocol at the University of Portsmouth. Zakrewski and Bull describe procedures at the University of Luton for the running of CBAs. However, the list of 'detailed procedures: pre-examination day' (Zakrewski and Bull, 1998, p. 146) does not refer to the in-depth QA steps for assessment development outlined in the current paper. It is our contention that any protocol for CBA must pay full attention to assessment development and the training of staff involved in the assessment production.

Training may also only be one part of the solution, in that institutions may have to consider using internal (or external) item-based assessment experts to oversee the development of the assessment. This may be problematic, in that it would remove the ownership of the assessment from the subject expert. Thus, such an innovation would require a major shift in institutional culture.

Staff training

We also contend that the development of 'objective' question styles is a professional skill different to teaching or conducting research. It is therefore essential that where CBA is practised, it is accompanied by training. King (1997, p. 24) reports that such training is carried out by means of small focus groups. McKenna and Bull (2000) note that academics designing tests for the first time will require 'support'. Indeed, the majority of publications covering implementation of CAA still regard staff training as voluntary and optional. Given the impact of university examinations on students' future careers, mandatory training and certification of staff involved in CAA production may be required.

The University of Derby now intends to institute a more formal system. A proposal currently under development envisages the creation of a Continuing Professional Development module in Assessment; successful completion of which would provide a credit towards a Masters degree in Education (M.Ed). This would be open to all staff and mandatory for new staff as part of the Post-Graduate Programme in Teaching and Learning.

The Scottish Qualifications Authority is currently developing an Advanced Certificate in E-assessment. A national qualification such as this would be of great value in ensuring homogeneity of standards.

Such developments are to be applauded, as they are potentially of great value in assuring the quality of assessment. However, unless certification of those involved in CBA becomes mandatory, their impact may be reduced.

Sharing materials

The sharing of good quality assessment materials within and between departments may be a way to avoid duplication of effort and ensure use of only the best items. Stephens *et al.* describe a discipline-level system, which maintains databases of available material (1998, p. 290).

In their 'Guidelines for implementing CAA in the FE sector in Scotland', Herd and Clark (2003, p. 1) strongly favour item banking:

'What will really make CAA work ... is the development of large assessment item banks ... where colleges can combine their efforts to create high quality, peer-reviewed questions.'

However, caution is needed, as some have a different interpretation of item banking. Some writers have suggested that items supplied with textbooks could provide a ready means of producing a usable item bank (Dalziel, 2000). But such an approach has been noted elsewhere to be inappropriate (Boyle *et al.*, 2002).

Beeston (2000) describes the use of a sophisticated item-banking system at an examination board. In Beeston's description the technology of the item bank is used to embed QA stages into working practices (2000, p. 8). For example, assessment materials can only move to the next part of the bank (moving ultimately to the 'live test bank') when all necessary QA stages have been carried out. The approach to item banking described by Beeston shows how strict procedures are essential if this technology is to improve assessment quality. If a large set of items were unthinkingly downloaded from a source of unknown repute, loaded into an item bank, and then added into assessments as a 'quick fix', quality would likely be negated, rather than improved.

Beeston also describes how the assembly and re-assembly of assessment forms from a bank of interchangeable items is dependent upon Item Response Theory (IRT) calibration. IRT analysis is necessary to provide the context-independent item parameters that permit the re-assembly of assessment forms with known measurement properties (2000, p. 9). Herd and Clark advocate the use of IRT as a long-term aim (2003, p. 22). But until IRT is successfully deployed, the value of item banking as a guarantor of quality will be limited.

Little work has been done to date to validate the use of item banking in HE. Wellington *et al.* (2001) describe their attempts to populate an item bank in electrical engineering using a team of subject experts, producing items that relate to the QAA Subject Benchmarks. There was a good degree of agreement about the quality of the items, their standard and the outcome they assessed. Unfortunately (to the current authors' knowledge) no item performance statistics were compiled. Thus, the validity of this application of item banking can only be supported with reservation.

One must also question the degree to which institutions would be willing to share items. Previous attempts (organised by Computers in Teaching Initiative (CTI) and Learning Technology Support Network (LTSN)) have largely foundered due to pragmatic issues such as the relative contributions of partner institutions. Thus, given practical issues, and the limited use of IRT in HE, it is our view that item banking (at least in the short term) is unlikely to guarantee assessment quality.

Qualitative research to assure quality

We noted above that in HE CBA there were a number of inhibitors to the use of statistical analysis techniques as commonly employed by large-scale assessment institutions. Knight has written about 'small-scale craft activities' (1999, p. 101). Knight stresses that he does not use this term pejoratively. Whilst large-scale statistical analyses may not be possible in UK HE, a range of qualitative techniques could be used to support quality. The focus groups of 'academics-assessment developers' advocated by King (1997) could be conceived of as an application of qualitative research methodology. Equally, research could be conducted with students to assure quality. This might take the form of questionnaires used to evaluate students' opinions (O'Hare, 2001) or researchers could work with small numbers of students using Verbal Protocol Analysis (Ericsson and Simon, 1984/1993) to ascertain reactions to specific items.

Conclusions

There is a clear requirement for QA procedures for CBA. The present approach can only lead to the delivery of unverified items and potentially poor assessments. Thus, there is a need for the UK CBA community to produce a clear set of mandatory guidelines for the assurance of quality in assessment development. The authors recommend the following steps as measures to aid the QA process:

- Mandatory training and certification of staff involved in CBA production, preferably via nationally recognised qualifications;
- Expert panels to review assessments, membership thereof to include assessment development experts;
- Piloting of new items in assessments (with scores not contributing to grades);
- Full evaluation of CBA within institutions, using quantitative and qualitative approaches as appropriate;
- Subject networks (e.g. LTSN) to discuss methods and arrangements for sharing validated items within national assessment banks;
- Wider dissemination of QA procedures within the CBA community.

The authors would welcome discussion of these issues in the wider CBA community.

Acknowledgements

NFER

Robert Ager	Research Officer
Pauline Benefield	Senior Research Librarian
Dr Adrian Woodthorpe	Principal Research Officer, Project Director

CIAD

Professor Don Mackenzie Director of CIAD

References

American Educational Research Association (AERA), American Psychological Association (APA) and National Council on Measurement in Education (NCME) (1999) *Standards for Educational and Psychological Testing* Washington, DC: American Educational Research Association.

Association of Test Publishers (ATP) (2000) *Guidelines for Computer-Based Testing* Washington, DC: Association of Test Publishers.

Beeston, S. (2000) *The UCLES EFL item banking system*. Research Notes 2 8-9 <http://www.cambridge-efl.org/rs_notes/0002/rn2.pdf> (14 May 2003).

Bennett, R.E. (2001) *How the internet will help large-scale assessment reinvent itself*. Education Policy Analysis Archives 9 (5) <<http://epaa.asu.edu/epaa/v9n5.html>> (14 May 2003).

Boyle, A., Hutchison, D., O'Hare, D. and Patterson, A. (2002) *Item selection and application in higher education* in Danson, M. (ed.) Sixth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, 9th and 10th July 2002 www.lboro.ac.uk/service/ltd/flicaa/conf2002/pdfs/ohare_d1.pdf (9 May 2003).

British Standards Institution (BSI) (2002) *BS 7988: 2002 - Code of Practice for the Use of Information Technology in the Delivery of Assessments* London: British Standards Institution.

Brown, S., Race, P. and Bull, J. (eds.) (1999) *Computer-Assisted Assessment in Higher Education* London: Kogan Page.

Bull, J. and Hesketh, I. (2001) *Computer-assisted assessment centre update* in Danson, M. and Eabry, C. (eds.) Fifth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, 2nd and 3rd July 2001 <<http://www.caaconference.com>> (11 March 2003).

Dalziel, J. (2000) *Integrating CAA with textbooks and question banks: options for enhancing learning* in Cooper, H. and Clowes, S. (eds.) Fourth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, 21st and 22nd June 2000 <<http://www.lboro.ac.uk/service/ltd/flicaa/conf2000/pdfs/dalzielj.pdf>> (23 May 2003).

Elton, L. and Johnston, B. (2002) *Assessment in Universities: a critical review of research* <http://www.ltsn.ac.uk/application.asp?app=resources.asp&process=full_record§ion=generic&id=13> (9 May 2003).

Ericsson, K.A. and Simon, H.A. (1984/1993) *Protocol Analysis: verbal reports as data* Cambridge, MA: Harvard University Press.

Euben, D.R. (2001) *Who Grades Students? Some legal cases, some best practice* Michigan, MI: University of Michigan, American Association of University Professors (AAUP) <<http://www.aaup.org/Com-a/aeuben.HTM>> (23 May 2003).

Herd, G. and Clark, G. (2003) *Guidelines for Implementing CAA in the FE Sector in Scotland* Glenrothes: Glenrothes College/Scottish Further Education Funding Council <<http://www.onlineassessment.org/Glenrothes%20Report-final.pdf>> (14 May 2003).

International Test Commission (2003) *International Guidelines on Computer-Based and Internet-Delivered Testing: Draft Version 0.3, March 2003* <<http://www.intestcom.org/Downloads/ITC%20Guidelines%20on%20Computer%20-%20draft%200-3.doc>> (9 May 2003).

Jenkins, M., Brown, T. and Armitage, S. (2001) *Management and Implementation of Virtual Learning Environments: a UCISA funded survey.* <<http://www.ucisa.ac.uk/TLIG/vle/VLEsurvey.pdf>> (23 May 2003).

King, T. (1997) *Recommendations for managing the implementation of computer aided assessment.* *Active Learning* **6** 23-7.

King, T., Billinge, D., Callear, D., Wilson, S., Wilson, A. and Briggs, J. (1998) *Developing and evaluating a CAA protocol for university students* in Danson, M.R. and Wade, W.B. (eds.) *Second Annual Computer Assisted Assessment (CAA) Conference Proceedings*, Loughborough University, 17th and 18th June 1998 <<http://www.tech.port.ac.uk/~kingt/research/sacaa.html>> (14 May 2003).

Knight, P.T. (1999) *Get the assessment right and everything else will follow?* *Quality in Higher Education* **5** (2) 101-105.

McKenna, C. (2000) *Using computers to assess humanities: some results from the national survey into the use of Computer-assisted Assessment (CAA).* *Computers & Texts* **18/19** 6-7
<<http://users.ox.ac.uk/~ctitext2/publish/comtxt/ct18-19/supp06mckenna.pdf>> (14 May 2003).

McKenna, C. (2001) *Academic approaches and attitudes towards CAA: a qualitative study* in Danson, M. and Eabry, C. (eds.) *Fifth International Computer Assisted Assessment (CAA) Conference Proceedings*, Loughborough University, 2nd and 3rd July 2001
<http://www.lboro.ac.uk/service/ltd/flicaa/conf2001/pdfs/k2.pdf>> (13 May 2003).

McKenna, C. and Bull, J. (1999) *The CAA Centre National Survey of Computer-assisted Assessment.* Paper presented at the Association for Learning Technology Conference, Bristol, September 1999.

McKenna, C. and Bull, J. (2000) *Quality assurance of computer-assisted assessment: practical and strategic issues*. *Quality Assurance in Education* **8** (1) 24–31.

Mackenzie, D.M. (2000) *Production and delivery of TRIADS assessments on a University wide basis* in Cooper, H. and Clowes, S. (eds.) Fourth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, 21st and 22nd June 2000 <<http://www.lboro.ac.uk/service/ltd/flicaa/conf2000/pdfs/mackenzied.pdf>> (22 May 2003).

O'Hare, D. (2001) *Student views of formative and summative CAA* in Danson, M. and Eabry, C. (eds.) Fifth International Computer Assisted Assessment (CAA) Conference Proceedings, Loughborough University, 2nd and 3rd July 2001 <<http://www.lboro.ac.uk/service/ltd/flicaa/conf2001/pdfs/m2.pdf>> (14 May 2003).

Olsen, J.B. (2000) *Guidelines for Computer-based Testing* <<http://www.isoc.org/oti/printversions/0500olsen.html>> (9 May 2003).

Qualifications and Curriculum Authority (QCA) (2003) *QCA's Mission and Business* <<http://www.qca.org.uk/about>> (9 May 2003).

Qualifications and Curriculum Authority (QCA), Qualifications, Curriculum and Assessment Authority For Wales (ACCAC), and Council For Curriculum, Examinations and Assessment For Northern Ireland (CCEA) (2002) *GCSE, GCSE in Vocational Subjects, GCE, VCE and GNVQ Code of Practice 2002/3* <http://www.qca.org.uk/nq/scripts/code_of_practice.pdf> (9 May 2003).

Quality Assurance Agency for Higher Education (QAA) (2000) *Code of Practice for the Assurance of Academic Quality and Standards in Higher Education: Section 6: Assessment of Students* http://www.qaa.ac.uk/public/COP/COPaosfinal/COP_AOS.pdf (14 May 2003).

Quality Assurance Agency for Higher Education (QAA) (2001) *The Framework for Higher Education Qualifications in England, Wales and Northern Ireland* - January 2001 <http://qaa.ac.uk/crntwork/nqf/ewni2001/ewni2001_textonly.htm> (14 May 2003).

Quality Assurance Agency for Higher Education (QAA) (2002) *Handbook for Institutional Audit: England* <http://www.qaa.ac.uk/public/inst_audit_hbook/audit_handbook.pdf> (14 May 2003).

Quality Assurance Agency for Higher Education (QAA) (2003) *Benchmarking Academic Standards* <<http://www.qaa.ac.uk/crntwork/benchmark/benchmarking.htm>> (15 May 2003).

Race, P. (1993) *Quality of assessment* in Race, P. Never Mind the Teaching Feel the Learning, SEDA Paper 80
<<http://www.lgu.ac.uk/deliberations/seda-pubs/Race.html>> (30 May 2003).

Scottish Qualifications Authority (SQA) (2003)
SQA Guidelines on Online Assessment for Further Education
<[http://www.sqa.org.uk/files_ccc/GuidelinesForOnlineAssessment\(Web\).pdf](http://www.sqa.org.uk/files_ccc/GuidelinesForOnlineAssessment(Web).pdf)> (9 May 2003).

Stephens, D. and Mascia, J. (1997) *Results of a Survey into the Use of Computer-Assisted Assessment in Institutions of Higher Education in the UK January 1997*.
<<http://www.lboro.ac.uk/service/ltd/flicaa/downloads/survey.pdf>>
(23 May 2003).

Stephens, D., Bull, J. and Wade, W. (1998) *Computer-assisted assessment: suggested guidelines for an institutional strategy*. *Assessment and Evaluation in Higher Education* **23** (3) 283–94.

Umar, J. (1999) Item Banking in Masters, G.N. and Keeves, J.P. (eds.) *Advances in Measurement in Educational Research & Assessment*. London: Pergamon.

Wellington, S.J., White, S. and Davis, H.C. (2001) *Populating the testbank: experiences within the electrical and electronic engineering curriculum* in Danson, M. and Eabry, C. (eds.) *Fifth International Computer Assisted Assessment (CAA) Conference Proceedings*, Loughborough University 2nd and 3rd July 2001
<http://www.lboro.ac.uk/service/ltd/flicaa/conf2001/pdfs/k2.pdf>> (13 May 2003).

Yorke, M. (1998) *The management of assessment in Higher Education*. *Assessment and Evaluation in Higher Education* **23** (2) 101-16.

Yorke, M., Bridges, P. and Woolf, H. (2000) *Mark distributions and marking practices in UK Higher Education*. *Active Learning in Higher Education* **1** 7-27.

Zakrewski, S. and Bull, J. (1998) *The mass implementation and evaluation of computer-based assessments*. *Assessment and Evaluation in Higher Education* **23** (2) 141–52.