RETHINKING ASSESSMENT THROUGH LEARNING TECHNOLOGIES

Joanna Bull
Eduology, UNITED KINGDOM
Joanna.bull@eduology.com

Grainne Conole, H.C. Davis & Su White
University of Southampton, UNITED KINGDOM
g.c.conole@soton.ac.uk, hcd@ecs.soton..ac.uk, saw@ecs.soton.ac.uk

Myles Danson
University of Loughborough, UNITED KINGDOM
m.danson@lboro.ac.uk

Niall Sclater
University of Strathclyde, UNITED KINGDOM
n.sclater@strath.ac.uk

Abstract

This paper describes a range of computer assisted assessment (CAA) research and development activities and discusses how CAA is enabling a rethinking of broader learning, teaching and assessment strategies. It includes four mini case studies which illustrate the diverse range of activities in this area. The first mini case study describes a review of the work of the CAA centre in the UK and a description of its impact across the sector, the second outlines the development of a test bank of CAA questions across a subject domain, discussing some of the key issues encountered by the project, particularly in terms of quality assurance aspects. The third describes a strategic institutional implementation of CAA and the way in which this has been achieved and finally the fourth mini case study looks to future developments in terms of CAA research and is an outline of a project which is developing a suite of freely available software to create, store, exchange and deliver assessment content. These mini case studies illustrate the ways in which CAA has developed and matured - moving from being a peripheral to a central part of the learning and teaching process. The paper describes some of the successes and failures of developing and implementing CAA systems. Evaluation of these activities has revealed insights about the potential of CAA, as well as identifying a number of issues concerning the future development of CAA and role of funded projects. The conference presentation will complement this and provide further in depth analysis of the results.

Keywords

Computer Assisted Assessment, Evaluation, Implementation, Development, Questionbanks

Introduction

Assessment content is *the* key component of online learning. It transforms text-based materials into an enriching learning experience. There is a serious lack of quality and timely feedback available to FE and HE students in the UK. Formative objective tests can be used to enhance interactivity within Web-based courses, providing feedback to tutors and learners on their progress and the effectiveness of the learning materials. The processes of (a) presenting a question or problem to the learner, (b) obtaining a response from them via the computer, (c) evaluating this response, (d) providing a mark and (e) responding to the

learner with feedback all require the provision of reliable systems. The databases behind these systems store the test items (questions and responses), the tests into which these items are incorporated, and details about users, their responses and results.

Computer Assisted Assessment (CAA) is used in a wide range of educational contexts which can broadly be summarised as: *diagnostic* – to ascertain the level of a student's (or group's) knowledge, carried out before or after learning activities take place; *formative* – carried out during a learning programme after a learning activity has been carried out to determine its effectiveness; and *summative* – carried out at the end of the learning programme or at a number of specific points during it to provide a mark which affects the level of qualification achieved. Assessment used for summative purposes acts as an *extrinsic* motivator on most students (Lepper, 1988). The quality of learning however can often be low when extrinsic factors are the main ones operating on students (Ramsden, 1992). However, when used for formative purposes within an online course, CAA can provide the interactivity necessary to engage students with learning materials. Otherwise drab page-turning modules can be brought alive by allowing the student periodically to assess their knowledge of the materials. Each of the four mini case studies described in this paper present different facets of CAA research and development activities from a mechanism for encouraging take up in the use of CAA and the cascading of good practice, through to more practical implementations of CAA systems or development of shared resources.

Mini Case Study 1 – Evaluation of the UK CAA centre

This section describes the findings of an evaluation of the UK CAA centre. A more detailed outline of the study and the methodology employed is reported elsewhere (G. Conole & Bull, 2002). The centre aimed "to disseminate good practice, guidelines and models of implementation and evaluation" (Bull, 2002). In particular the evaluation explored the impact of integrating CAA within learning and teaching and the effect at both the local and the institutional level. CAA encompasses a range of activities, including computer-based testing, computerised assessment, computer-aided assessment and Web-based assessment. Previous research has shown that there are a range of motivations for implementing CAA within a course and it is often a combination of factors which result in CAA being used (Bull & McKenna, 2001) and the evaluation aimed to explore this in more detail.

Overall evaluation feedback was very positive about the centre and its work. It was felt that it had had a real impact both institutionally and nationally. The centre produced a 'CAA blueprint' (Bull & McKenna, 2001) which was highlighted by the community as a valuable resource, which provided an authoritative text on the area. Networking and its value in projects of this kind, both in terms of support and access to expertise, was also cited by many in the evaluation as important. As has been reported for other learning technologies, CAA activities appear to act as a catalyst for discussing more generic learning and teaching issues. It was also reported that the project was timely within individual institutions and that the CAA project helped to drive forward other agendas and institutional initiatives. The project also afforded the opportunity to consider underlying theoretical questions about the nature and role of CAA and its use. It provided a snapshot of where institutions were at a given moment in time both in terms of CAA and other learning technology developments, which could then be used both as a historical record and a benchmark against which to contrast future developments. More generally it raised issues about the relationship between teaching, learning and assessment and existing practices.

It is clear from the evaluation that a range of strategies are used by staff when trying to implement and integrate CAA into existing learning and teaching programmes. This mirrors other research findings of the use and integration of ICT (Davis, Oliver, & Smith, 2001; Harvey & Oliver, 2001). Some of those interviewed doubted the impact of institutional strategy at ground level and commented that innovation in learning and teaching was left primarily to the individual. However, a number of key strategies and drivers did emerge for getting staff started in and then continuing to use CAA. One interviewee mentioned the importance of peer support and the role of colleagues in terms of providing advice and support. She commented that in her case she worked very closely with a work colleague and used him to 'bounce ideas off', going on to say that in a sense he acted as a driving force. Another stated that discussions within a committee about the increased problems of plagiarism helped her to make a mental note of thinking how she might use CAA in her own practice to reduce such dangers. Personal motivation and interest was also cited, and were clearly a major factor for many of those interviewed, particularly the enthusiasts. Institutional support was cited as critical and one interviewee commented that the perceived

lack of support for CAA within his institution was demoralising. Institutional support was seen as including explicit tie-ins to relevant university strategies (such as the learning and teaching strategy), resource allocation (in terms of adequate equipment, software purchase and technical support), as well as investment in the support mechanisms, staff development and training.

It was evident from the interviews that staff recognised that students are becoming increasingly proficient in the use of technology to support learning, but that equally they were critical of the use of technology for its own sake and looked to see evidence of the enhanced benefits for learning. Interviewees were also sensitive to student needs and stressed the importance of having an understanding of the student perspective towards new technologies, as well as the need to ensure equality of access to technologies. In particular, one lecturer stressed that if students did not have access off site, then it was questionable how much technology should be used in the course.

The type and form of support varied significantly across the institutions. A national audit of learning technologists and their locations within institutions (Beetham, Jones, & Gornall, 2001) also reveals significant differences, which may be a consequence of this being a new and emergent role within institutions which is not yet fully embedded. Some had well established and branded central or faculty support mechanism and named individuals that staff were aware of, others relied on word of mouth or peer support at the local level. Many of the interviewees classified themselves as average in terms of their expertise in using learning technologies. For this group it was stressed that help is needed in terms of getting started in using CAA, in terms of support to introduce them to the possible ways in which it can be used and how it can support learning and teaching. Many commented that when this type of structured advice and support was not available, there tended to be low take-up.

There was a range of motivations cited for using CAA. For some it was the interest in exploring new innovations and new ways of enhancing learning for their students. Others were driven by the prospect of potential resource savings in using CAA, particularly in terms of marking large student cohort exam scripts or in re-use of materials for different purposes. Some cynically observe that there had been an institutional three-line whip to increase the amount of CAA within courses. Interviewees also recognised that the sector had changed significantly in the last five years: increased student numbers and associated workloads, and an increasing emphasis on quality audit and control were but two examples of changes which impacted directly on the increased interest in and uptake of learning technologies like CAA. A number of interviewees also commented on the changing technological profile and in particular the increase of the level of ICT literacy amongst the ‘computer games’ generation of children and the increased presence of technology in the home. Related research on the use of computers in the home supports this and importantly highlights the engaging, active and contextualised nature of the types of learning which take place with use of computers in the home (Sutherland, Facer, Furlong, & Furlong, 2000).

One interviewee expressed concern that CAA was not adequately integrated with learning and teaching and felt that until this was adequately addressed, it would be very much of an add-on to the learning experience. There was also a concern that standard good practice should not be lost. It was evident that those who had explored the use of CAA had gained personal satisfaction and felt that it had improved their own practice generally. In particular, as has been found more generally with learning technologies, consideration of the use of CAA made staff stop and think about their general approach to learning and teaching and the nature of their practice per se.

There were a number of barriers to wide-scale uptake of CAA within institutions. A number of other studies have also reported on the barriers to use of ICT more generally which reflect these findings (Grainne Conole, 2001; Squires, Conole, & Jacobs, 2000). Some of the barriers are related to technical difficulties. In particular, the functionality of current CAA software is still improving. Many were frustrated by the limited functionality available and some felt that therefore CAA was not appropriate for teaching their subject area. The cost of software and, more importantly, the hidden cost of staff time in developing and implementing CAA were cited as barriers to some extent. Others included the lack of appropriate equipment and resources, and institutions varied in the extent to which they provided additional staff support and training. Not surprisingly the tension between time spent on research versus teaching was a recurrent issue, particularly cited by those staff in the more research-led institutions. This demonstrate a crucial issue about the perception of pedagogical research as somewhat secondary and

inferior to ‘mainstream’ research and a perception – even by the innovators – that papers in this area do not count in terms of research standing.

One of the key benefits of using CAA cited was the speed of marking and speed of providing results and feedback. Some felt that CAA was 'more objective' than marking by hand, particularly with large student cohorts. Interestingly, one lecturer related that a student of his, whilst disappointed with a mark he had got, was more inclined to accept this final result because he felt the mark produced by CAA was more accurate than if it has been marked directly by the tutor.

Managers interviewed felt that there had been a shift in the sector in recent years and that staff in general were more susceptible to using learning technologies to support their teaching. They cited a range of tensions and, in particular, that the increased drive to do research alongside the increasing quality demands for teaching, was problematic. However, there appeared to be a move towards more cohesive strategic thinking and many cited the importance of institutional learning and teaching strategies as helping them to enable and increase innovations in learning and teaching. One manager cited the importance of champions to encourage innovation and in particular that the characteristics of these champions was an important factor, someone with a missionary zeal for technology for example, was not necessarily most appropriate, this mirrors findings from a recent national survey in this area (Beetham, 2001). Disappointingly, few felt that the project had had a major impact within their institutions. Senior managers felt instead that it was the cumulative impact of an aggregate of projects of this kind within institutions that was resulting in a shift in practice. Encouragingly, there is evidence that this is happening and there is also an increase in embedding learning technology implementation into main stream strategy. Overall it is clear that the CAA centre was a successful and valuable initiative in a number of ways: through the professionalism of the centre and associated staff, through the outputs and materials produced, with the development of a network of expertise at institutional and at a national level through alignment of the CAA initiative with relevant institutional strategies.

Mini Case Study 2 – Development of a Testbank of Questions

The Electrical and Electronic Engineering Assessment Network (e3an) is collating sets of peer-reviewed questions in electrical and electronic engineering which have been authored by academics in the UK (http://www.ecs.soton.ac.uk/e3an.). The questions are stored in a database and available for export in a variety of formats chosen to enable widespread use across a sector which does not have a single platform or engine with which it can present questions.

One of the key problems of developing a testbank of questions is the difficulties associated with recruiting adequate numbers of question-authoring consultants to produce well-formed questions to populate the question database. To counter this, a detailed specification of the contents of questions was made and authors were provided with word processor templates which could be automatically processed and converted into the data items in the question database. It was also possible for authors to provide clear hand written detailed feedback in the form of tutor’s notes to be scanned. Authors were left to their own devices to write the questions, and it was assumed that the peer review process would identify and remedy any gaps in the quality of the question. It was also assumed that this laissez faire approach would enable detail specification and explain requirements to future authors. However, when it came to transferring the electronic versions of the completed into the database, the number of issues and problems encountered were significantly larger than was originally envisaged, to an extent which would not be sustainable in the longer term.
Activities associated with the production of the database were divided into a number of distinct stages – the pre-operational stage, the collection of questions themselves, and then the post processing of the questions. The pre-operational stage involved the project team in answering the following design questions:

- Database contents and function and specification
- Target themes
- Metadata
- Question Templates

The next stages were to build theme teams (by recruiting and briefing academic consultants) and start the authoring and peer reviewing process of the initial set of questions. These activities were all carried out in parallel with the continuing specification and design of the question database. When the authoring task was complete, additional processing needed to be carried out by the core team to:

- process sample questions for inclusion in the e3an testbank
- partially automate the conversion process
- review the success of the authoring process
- identify good practice for future authoring cycles

The testbank was conceived as a resource to enable academics to quickly build tests and example sheets for the purpose of providing practice and feedback for students, and also to act as a resource for staff providing exemplars when creating new assessments or redesigning teaching. Because questions for the database were to be produced by a large number of academic authors, a simple text format was chosen as the data entry standard. In addition, in order to enable the widest possible use across the sector, it was important that questions from the bank could be exported in a range of formats.

The question database was designed to allow users to create a 'set' of questions which could then be exported to create either paper-based or electronic formative, summative or diagnostic tests. Output formats specified included RTF, for printing on paper (as the lowest common denominator), HTML for use with Web authoring, and an interoperable format for use with standard computer based test engines which is a subset of the IMS Question Test and Interoperability (QTI) specification (Smythe C, Shepherd E, Brewer L, & S, 2002).
There are many types of objective question; QuestionMark (http://www.questionmark.com), supports 19 different types, and the latest IMS QTI specification (v 1.2) lists 21 different types. However, the project team wanted to experiment with a small number of objective question types, and decided to confine initial collection to multiple choice, multiple response and numeric answer. In addition to objective questions, the project team wished to collect questions that required written answers, because some areas of engineering, design in particular, are difficult to test using objective questions. Furthermore, a collection of exam style questions, along with worked answers, was felt to be a useful resource for both teachers and students.
Four initial themes (Analogue Electronics, Circuit Theory, Digital Electronics and Microprocessors, and Signal Processing) were chosen because they were seen as being core to virtually every degree programme, and were chosen to reflect the breadth of the curriculum along with the teaching interests of the project team members. A theme leader was appointed for each subject area, drawn from each of the partner institutions. An additional concern was that the material produced should, as far as practicable, be relevant to courses in electrical engineering. An electrical engineering subject specialist was therefore appointed to work with the four theme teams and encourage consultants to reflect “heavy current” interests. Consultants were recruited from the partner institutions. Project team members were initially asked to identify and canvass potential contributors from their own institution for each of the four subject areas. Prospective consultants were then invited to attend a half-day training session. Individual briefings were organised for prospective consultants who were unable to attend on either one of the two dates offered. The briefing session was divided into two main sections; an introduction to the e^{3}an project and objective testing, and a meeting between the theme leader and members of the theme team to discuss and agree specific objectives for their subject theme area. The introductory component included:

- A description of the e^{3}an project, objectives, participants, timescales and deliverables;
- Overview of issues in student assessment, including the benefits of timely formative feedback and the outcomes-based approach to assessment advocated by the Quality Assurance Agency for Higher Education (QAA, 2000a, 2000b);
- Familiarisation with guidelines for writing effective objective test questions, question types: multiple choice, multiple response, numeric answer and text response.

The latter, practical components of the briefing, were considered essential since the design of effective objective test questions is an acquired skill (McKenna & Bull, 1999). Some general guidelines were presented (Nielsen, 1994), along with additional examples drawn from the electrical and electronic engineering curriculum. Specific examples demonstrated how indicative questions from a “traditional”

examination paper might be converted into objective test format, and how an author might design a question to test a specific learning outcome. The second part of the briefing session involved members of theme teams meeting with the theme leader to discuss and agree the key curriculum areas within their themes and the spread of question types that would most usefully support these sub-themes. This activity was conducted over the two separate events, with details finalised by email. Theme leaders initially proposed the main sub-themes or topics and their indicative level. Questions were classified as "Introductory", "Intermediate" and "Advanced" levels which broadly correspond to the three levels of a full-time undergraduate programme in electrical/electronic engineering. However the schema is flexible enough to cope with the specialist nature of some degree programmes and acknowledges that both timing and intensity of study may vary between institutions.
There was also some debate amongst members of the project team about the designation of materials as being relevant to the fourth year of an MEng programme. An additional metadata item of tutor information provides an opportunity for question writers to append explanatory notes if they felt this would be appropriate. Once the training and allocation of questions was completed authors took a few weeks to produce their questions and then met again to peer review each others questions, before the questions were accepted for publication in the database. Consultants were paid on successful peer review of their allocation of questions. Once the questions were delivered, there were two questions which the project team needed to resolve; what information to store in the database and what database architecture to use.

The following output formats were agreed:

- Rich Text Format – in order to allow the creation of paper-based tests, with two formats - the questions alone and the questions with worked answers.
- HTML – to enable questions to be put on the Web, again in question only and question with answer format.
- XML to the Question & Test Interoperability (QTI) specification which will allow the questions to be used by a range of proprietary test engines including Blackboard and version 3 of Question Mark Perception.

The team were aware that many users might wish to import the questions into earlier versions of Question Mark, which do not support QTI, and for this reason QML output (Question Mark's proprietary XML format) was also provided. A programme was written to read the Document Object Model (DOM) of the questions that had been provided in the Word templates, and from this to create the XML QTI version of the question. In theory it would have been adequate to store only the original Word documents as all the other formats could have been created dynamically from these documents. However because of speed limitations the input documents were batch processed and all the possible output formats were stored.

The question of what architecture to use was decided by the fact that there was a need to be able to distribute the database to individual academics, maybe working without system support; so that authors could be sent a CD enabling them to install and query the database on their own. Microsoft Access was used and the system was delivered packaged with the Access runtime for those that did not have Access on their machines. In parallel an SQL server version of the database was developed, so that it could be queried over the Internet. The user could therefore assemble a question set either by browsing the whole database and making individual selections, searching using individual or multiple keywords or by a combination of both these methods. At this stage the user could also view the descriptive metadata for each question and preview the tutor version of each question in Microsoft Word.

Given the ubiquity of online shopping it was decided to use a trolley metaphor to enable the building of question collections. If the user wishes to go on to make use of a question or set of questions this is achieved by adding items to the "question trolley". The "question trolley" can be viewed at any time during the process and questions can be added/deleted. The question set can then be sent to the "question checkout to be exported. The user is given an option of saving the question set so that it may be modified or used again at a later date.

At the "question checkout" the user may choose which combination of versions and formats the questions are to be exported in. Each question has a unique identifier and the suffix of each question identifies its version and format. Currently each question can be found as a separate file after being exported

individually. Future versions will also enable optional export of concatenated collections of questions in the various formats.
The evaluation process identified a number of issues when users accessed the database. Firstly, users were unsure what was meant generally by the term metadata and hence unable to make best use of this. In particular, they did not always understand the purpose of specific metadata items (e.g. question level). Secondly, inconsistent naming conventions within the retrieval process created confusion. Thirdly, users found different stages of the process indistinct and fourthly users were sometimes presented with more information than they needed.

Subsequent testing via the scenario and simplified thinking aloud suggests that the usability problems originally found have been solved without creating additional problems. Most problems identified in each prototype were found by the first five test users, testifying to the power of these methods in identifying usability problems easily overlooked by users involved in the database specification. It showed how a perceived "easy to use" database can still result in unexpected user problems.

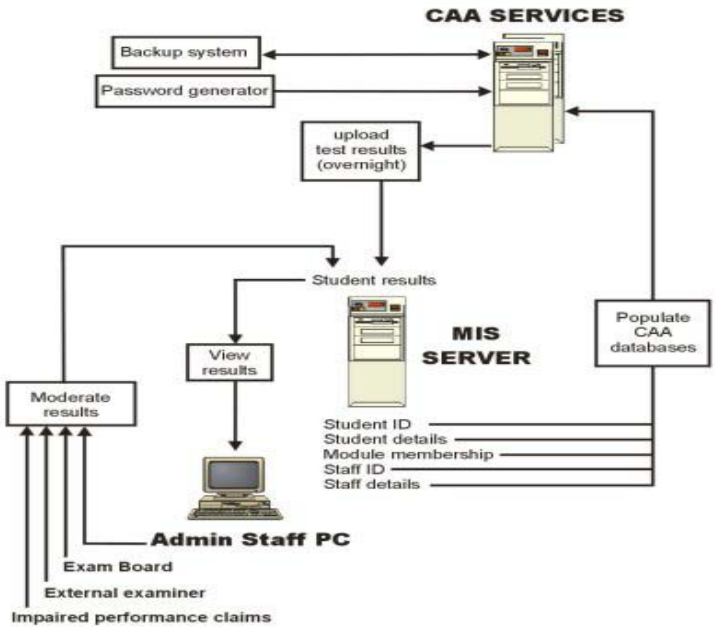
Unfortunately there is not yet any similar common methodology for the collaborative creation of database contents. For e^{3}an the objective achieving a high level of voluntary involvement among question authors was equally important to creating a useful testbank (White & Davis, 2001). The time needed to manually process the first batch of questions far exceeded original estimates, and authors variously showed imagination in the variety of problems they created for the design team. However these lessons have were incorporated into the author briefing sessions, as well as in a detailed guidance document and a validation tool.

Mini Case Study 3 – Institutional implementation

Over the past two years, Loughborough has been integrating a Web-based CAA system into their local Managament Information Systems (MIS). They have rolled out a commercial product, Question Mark Perception as a central online assessment system. In addition they received funding to integrate this system with a VLE (virtual learning environment) and corporate MIS. This goes some way towards an emerging Managed Learning Environment (MLE) for the university. The deliverable is documentation outlining the approach taken, problems encountered and solutions implemented. A central tenet of the implementation was to adopt as generic an approach as possible in order that the document is of value to other HEIs considering similar work.

There are a number of key features to this online assessment system and issues which have emerged during the project. An important aspect was carried out prior to any development work. This involved mapping the traditional assessment processes, to identify key stakeholders and opportunities to improve efficiency. This revealed a number of problems with the traditional assessment system. A clear lesson is not to assume that because a system is widely accepted it has been developed with any clear strategy, nor that it is necessarily an ideal solution. The Web CAA System has also been mapped to flow diagrams, one of which is shown below.

Fig 2: MIS interfaces



Assessment impacts on a number of traditional roles across universities and is clearly a high stakes core activity. To address this, the management structure for the project at Loughborough was set up to ensure maximum input from relevant stakeholders. Useability and accessibility issues were also built into the project, for example the screen was designed taking into account a number of features incorporated through negotiation with the Examinations Office, Disability and Additional Needs Service, screen design consultants, students and academic staff. Interestingly, the system was designed to include two areas of student support, i) online help and ii) a less obvious support requirement and benefit revealed by the authentication system. In addition, the candidate authentication system has parallels with traditional assessment procedures. The use of the Student ID number for authentication revealed several benefits and one draw back. The project showed that there were improvements over the traditional assessment procedure. The automated assessment system has improved features to assist in invigilation and these may be superior to the traditional assessment process. In total 10 different question types are supported by the system. It is available 24x7 and logs are generated to show that students take tests from 09.00 through to 03.00 given the opportunity. Loughborough have tied the digitally captured results into the university MIS offering an automated score upload and ratification feature. The draft of this feature has been well received by departmental administrators and the work was carried out by a university MIS officer. In terms of future development, the software house have released a major upgrade, and a project to deploy this has been scoped out and agreed. The next 12 months will see a new recruitment drive and the intention is to offer training through the system itself.

Scaleability was considered throughout the development phase. The finalised system is tied into the central Student Records system. This allows the scheduling of tests to groups (modules/courses) of students. Thus an individual menu of available tests is presented for every individual student. Tests are often time limited and this is a useful anti-cheating strategy. Further to this tests become available to take, and are withdrawn automatically and down to the minute. Through the provision of feedback to the software house a number of key features of benefit to the education sector have been developed in the more recent version of Perception. One continuing problem is the question and test publishing route. The Loughborough project team have developed a practical solution to this area, however this work has been taken forward further by K.C. Leuven University (Van Rentergem, 2001).

The project has required the involvement of key stakeholders across the institution and has also been firmly embedded into relevant institutional strategies. The latest Institutional Learning and Teaching

Strategy includes targets for CAA. In line with central Government policy the University is actively engaged in drawing up Departmental Learning and Teaching Strategies. It is envisaged that where appropriate these will include micro level details regarding CAA implementation and targets. Senior management buy in has enabled the continued support for the work and the service is now beginning to reap the benefits. The system has delivered over 15000 assessments during the 12 month pilot phase. Prior experience of screen delivered CAA suggests that the usage figures will increase exponentially. Spin off benefits for the University include a central question bank. Questions are authored and stored on a local client machine. To use with students they must be published to the main Perception Server via ODBC (Open Data Base Connectivity). The University has already developed a question bank comprising some 8000 items. The emerging benefits of interoperability have been exploited and the question bank exports to IMS QTI. This goes a long way towards the avoidance of software vendor lock in, as well as reassuring academic staff that the effort put into authoring questions need not be repeated should a new assessment engine be invested in the future. If time permits a live demonstration of the system is possible. This may include genuine tests in use with current students, as well as the Web based reporting tool showing live student data and revealing interesting trends in behaviour. The latter may be seen as evidence that under this testing regime students are taking responsibility for their own learning, prioritising their time, and revisiting tests in order to utilise the excellent feedback opportunities offered by this Web CAA System.

Mini case study 4 – Technologies for Online Interoperable Assessment (TOIA)

The creation of high quality assessment content is difficult, time consuming and expensive. Developing valid distractors for multiple choice questions is a highly skilled activity. The process is replicated on an institutional basis across FE/HE with few cross-institutional repositories offering high quality exemplars and content. It has long been recognised that a market for assessment content and national subject-based question banks from which an individual department or teacher could select questions and tests would be advantageous. A market has been slow to develop, partly because of the lack of an agreed format in which to exchange questions and tests. For the first time, however, an XML-based specification has emerged which provides an opportunity for assessment content creators and users to share data in a platform-independent format: the IMS Question and Test Interoperability (QTI) specification. With the planned incorporation of these specifications into the US Department of Defence's influential Shareable Content Object Reference Model (SCORM), QTI is likely to remain the only serious player in assessment content specifications. IMS's recent partnership agreement with the Open Knowledge Initiative further increases its status.

Members of the UK Centre for Educational Technology Interoperability Standards (CETIS) Assessment Special Interest Group (www.cetis.ac.uk/assessment/) have been following closely the development of the QTI specification and implementing it in a variety of ways at their own institutions. While a number of virtual learning environment and commercial online assessment system vendors claim to be making their systems QTI "compliant", only QuestionMark, which itself had a major input to the development of the specification, currently allows the import and export of questions in QTI format. While undoubtedly the leading product in this area, the cost of QuestionMark's Perception is prohibitive for many potential FE/HE users. There is an increasing need for freely-available tools to allow colleges and universities to create, store, exchange and deliver assessment content. A number of prototype systems and databases have been developed by members of the Special Interest Group. They are now planning on pooling their expertise within the *Technologies for Online Interoperable Assessment* (TOIA) consortum to develop administrative procedures and QTI-based systems and databases to facilitate and promote the uptake of online assessment across the sector.

In setting up the organisational and technical infrastructure for question and test banks TOIA intends to develop procedures and technologies for multi-institutional, cross-sectoral storage, retrieval and exchange of assessments and student performance data. All tools developed will be provided free to UK higher and further education communities. They will be fully accessible for learners with special needs, many of whom will benefit from the ability to engage in formative assessment at their own pace, rendered in ways which meets their requirements.

TOIA has three main areas of activity:

- Population of question and test banks
- Development bay for assessment content
- Interoperability between TOIA and other systems

TOIA will facilitate the creation of high quality online assessments through the provision of reliable administrative and technical infrastructures and tools. Initially the consortium intend to concentrate primarily on the development of tools and the "service" architecture for assessment content banks, it will also look at the population of the question and test banks with peer-reviewed questions, informed by the work carried out under the e^{3}an initiative. As described above, this network, has already assembled a large collection of questions in the area of electronics and electrical engineering in close collaboration with practitioners to ensure an appropriate pedagogical approach. The incorporation of question metadata for browsing, searching and selection reinforces the question author's understanding of the underlying educational objectives of the test items. Test banks are particularly useful in preparing students for examinations. Access to large numbers of appropriate questions as part of the revision process is likely to enhance educational achievement. Core response types in IMS QTI are multiple choice, true/false, multiple response, image hot spot, fill in the blank, select text, slide, drag object, drag target, order objects, match item and connect the points. TOIA's systems will implement these and also seek to integrate other innovative response types, develop appropriate XML schemas, and feed these back to IMS for incorporation in later versions of the QTI specification.

The separation of content from presentation inherent in QTI means that accessible questions and tests can be created, stored and exchanged. At run-time the questions can be rendered using alternative mechanisms such as screen readers for the visually-impaired. The IMS QTI specification now incorporates a schema for the exchange of results, another element of key interest to this project. The automatic capture of student input during formative assessment means that tutors can identify problem learning areas for individuals and across classes, take early remedial action and reduce the potential for drop-out. When used summatively, the administrative benefits of having student results collated and available instantly are considerable. The ability to combine the question and test part of the specification with the results interoperability section means that sophisticated analysis of the reliability and validity of items can also be carried out, allowing question developers to eradicate unsuitable items from future tests. Questions and tests in QTI format can be combined with other learning objects using the IMS Content Packaging Specification to form complete learning packages. This is currently handled poorly by systems such as Blackboard which exports questions within an IMS content package in its own proprietary format (Sclater, Low, & Barr, 2002). Pioneering work is currently taking place at Edinburgh university where interactive online courses in meteorology are being created using a combination of QTI and the Open University of the Netherlands' *Educational Modelling Language* (EML), currently being considered for adoption by IMS. The TOIA project is particularly interested in combining questions and tests with other online content using EML or IMS's Learning Design Specification, which address different pedagogical approaches to learning.

Conclusion

The four mini-case studies described in this paper provide a snapshot of current CAA activities and developments and give an indication of the diversity of work under development. The evaluation of the CAA centre clearly illustrates the way in which implementation of CAA is promoting a wider debate about learning and teaching and in particular highlighting issues of quality assurance as well as the role of assessment and feedback as part of the learning experience. The study was also encouraging in that it demonstrated that there has clearly been a shift in the last five years from the use of CAA being a peripheral activity, the province of the enthusiasts, to becoming part of central strategic thinking and part of everyday learning and teaching practices. This is further illustrated in the mini-case study describing the institutional implementation of a commercial CAA system, which demonstrates one institution's approach to maximising the benefits of CAA and illustrates the operational and technical issues which distinguish this form of assessment from traditional approaches. The mini-case study on the development of a cross-sector questionbank of resources for Engineering and the development of a suite of software and tools for creating, storing, exchange and deliver of assessment content illustrate the ways in which colleagues are seeking to work collaboratively to create shared resources, which can be adapted and re-

purposed to better meet the needs of students learning. Future CAA developments will need to address the wide range of issues that these case studies have highlighted, ensuring that pedagogic design is high on the agenda and that flexible, integrated systems and content are developed to enhance learning and assessment

It is encouraging that the mini-case studies show that staff are adopted a considered and reflective approach towards using technology. For example, staff recognised that students are becoming increasingly proficient in the use of technology to support learning, but that equally they were critical of the use of technology for its own sake and looked to see evidence of the enhanced benefits for learning. Projects such as the development of a shared questionbank and the institutional implementation of a commercial CAA system also provide an understanding of the cultural and human perspectives in the successful implementation and use of different forms of technology.

References

Beetham, H., Jones, S., & Gornall, L. (2001). *Career Development of Learning Technology Staff: Scoping Study, a final report for the JISC JCALT*. Bristol: University of Plymouth

University of Bristol.

Bull, J. (2002). *Implementation and Evaluation of Computer-assisted Assessment - final report.* Luton: University of Luton.

Bull, J., & McKenna, C. (2001). *Blueprint for Computer-assisted Assessment.* Luton: University of Luton.

Conole, G. (2001). *Drivers and barriers to utilising Information and Communication Technologies.* Paper presented at the IEEE conference on teaching and learning, Savoy Place, London.

Conole, G., & Bull, J. (2002). *Pebbles in the pond: evaluation of the CAA centre.* Paper presented at the Sixth international computer assisted assessment conference, Loughborough.

Davis, M., Oliver, M., & Smith, J. (2001). *ASTER Final Evaluation Report*. London: University of North London

University College London.

Harvey, J., & Oliver, M. (2001). *EFFECTS external evaluation: final report*.

Lepper, M. (1988). Motivational considerations in the study of instruction. *Cognition and Instruction, 5*, 289-309.

McKenna, C., & Bull, J. (1999). *Designing effective objective test questions: an introductory workshop.* Paper presented at the Proc. 3rd Annual CAA Conference, Loughborough.

Nielsen, J. (1994). *Guerrilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier - Designing Web Usability: The Practice of Simplicity*. New York: John Wiley & Sons.

QAA. (2000a). *QAA for Higher Education - Code of practice for the assurance of academic quality and standards in higher education. Section 6: Assessment of students*.

QAA. (2000b). *QAA for Higher Education - Handbook for Academic Review*.

Ramsden, P. (1992). *Learning to Teach in Higher Education*. London: Routledge.

Sclater, N., Low, B., & Barr, N. (2002, 9-10 July 2002). *Interoperability with CAA: does it work in practice?* Paper presented at the Proceedings of the Sixth International Computer Assisted Assessment Conference, Loughborough University.

Smythe C, Shepherd E, Brewer L, & S, L. (2002). *IMS Question & Test Interoperability: ASI Information Model Final Specification, Version 1.2, IMS*, February 2002, from http://www.imsproject.org/question/

Squires, D., Conole, G., & Jacobs, G. (Eds.). (2000). *The Changing Face of Learning Technology - Selected Papers from ALT-J* (1 ed.). Cardiff: University of Wales Press.

Sutherland, R., Facer, K., Furlong, R., & Furlong, J. (2000). A new environment for education? The computer in the home. *Computers and Education, 34*, 195-212.

Van Rentergem, L. (2001). *A Campus Wide Set up of Question Mark Perception V2.5 at the Katholieke Universiteit Leuven - Facing a Large Scale Implementation.* Paper presented at the 6th International CAA Conference Proceedings, Loughborough University.

White, S., & Davis, H. C. (2001). *Innovating Engineering Assessment: A strategic approach.* Paper presented at the Engineering Education: Innovation in Teaching Learning and Assessment Proceeding First Annual IEE Engineering Education Symposium, Savoy, London.