Computer-based, online summative assessment in undergraduate pharmacy teaching: The Manchester experience

HARMESH AOJULA1, JILL BARBER1, ROD CULLEN2, & JULIE ANDREWS1

1School of Pharmacy and Pharmaceutical Sciences, University of Manchester, Manchester M13 9PL, UK, and 2Distributed Learning Group, University of Manchester, Manchester M13 9PL, UK

(Received 3 February 2006; revised 19 June 2006; accepted 27 June 2006)

Abstract

Computer-based assessment (CBA) has many advantages for staff and students in higher education. Although we have successfully used diagnostic and formative CBA for a number of years, the introduction of summative assessments raised additional concerns about development time, academic rigour, security and organisation. To explore these issues, a pilot study was set up to assess the running of online (summative) CBA for 240 students taking a first year course in Cell Biology and Biochemistry. We have compared the scores obtained from computer-based marking, with those obtained from: (a) optical mark readers; and (b) traditional hand-marking of scripts. Computer-based marking of text matching questions was initially compromised by spelling errors but these were reduced to an acceptable (<1%) level by introducing an online subject-specific list of correctly-spelled options. The development and evaluation of secure online examination procedures is also discussed. We conclude that, with these improvements, online CBA can be used successfully for a range of summative assessments in undergraduate Pharmacy courses.

Keywords: Assessment, CBA, online, pharmacy, summative, WebCT

Introduction

Computer-based assessment (CBA) is defined as “an assessment in which the questions or tasks are delivered to a student via a computer terminal”. In most cases, the student’s answers are typed in at the computer keyboard and recorded and marked electronically. CBA can be delivered: (a) using stand alone machines; (b) using a local area network of machines; or (c) using an intranet or internet. For students in higher education, CBA has many advantages (Pitt & Gunn, 2004). The most obvious of these is that a computer can achieve very rapid, objective, anonymous and accurate marking. However, CBA provides a number of other benefits as well. High quality colour figures, photographs and even video clips, animations and sound can be easily incorporated into online assessments. Furthermore, many conventional modes of cheating can be eliminated, firstly because each student can receive a unique examination paper of randomly shuffled questions (Ward, Gordon, Field, & Lehmann, 2001; Thelwall, 2000) and secondly because students do not need to take pencil cases or calculators into the examination room (the on-screen calculator is sufficiently powerful for most applications in pharmacy). Computer-based examinations can also, in some situations, address the needs of students with certain disabilities. For example, the use of cascading style sheets (.css files) in online (web-based) CBA allows students with special needs to personalise (using web browser Accessibility options) the display of pages on their computer. A student with impaired vision, for example, may choose a personal style that applies a large, clear font and a contrasting background colour, while students with controlled epilepsy or severe dyslexia may choose quite different settings appropriate to their own disabilities. Despite these advantages, there are few reports of investigations of the reliability of online examinations in the literature (Howarth, Messing, & Altas, 2004).
Here in Manchester, we have had several years experience of both diagnostic and formative online CBA and have found that this mode of assessment is generally well-received by our students. In 2001, students in the School of Pharmacy expressed a slight preference for online testing (55%) versus conventional (paper-based) testing (45%), but by 2006, this had risen to 75% of students expressing a preference for online testing. Furthermore, we consider it likely that the extent of this preference will increase even further as student intakes become more familiar with online testing and with computers in general. Indeed, the more widespread promotion of e-learning together with e-assessment and online CBA is now one of the key pedagogic objectives in this University, as it has been shown that high quality formative assessment has a strong positive effect on student learning both in general (Olson & McDonald, 2004) and specifically in pharmacy (Andrew, 2000). Until recently however, all the summative assessments within the School have been traditional paper-based examinations.

The case for using online CBA (summatively) for end-of-semester examinations is most compelling for modules with large numbers of students, especially when there is an established database of suitable questions. However, summative online CBA has not been widely adopted for a number of important reasons (Pitt & Gunn, 2004; Ward, Gordon, Field, & Lehmann, 2001; Thelwall, 2000). One factor is that the overall preparation and delivery of web-based assessments may be unfamiliar to academic staff and another is that the examination itself may require different administrative arrangements from an equivalent paper-based examination. Many of the problems are however logistical issues, due to lack of experience and/or unavailability of fully-evaluated protocols for successful implementation of summative examinations online. There has also been some concern as to whether online assessments unacceptably restrict the choice of question types, thereby devaluing the assessment (Ward et al., 2001) and a seemingly exhaustive list of potential problems has been addressed by Zakrzewski and Steven (2000, 2003). However, in our experience, most of the issues have been technical rather than academic. For example:

- Production of online assessment materials—unto what extent are web skills required?
- Candidate registration and authentication—how do we know who is sitting the paper?
- Security—how do we ensure that the paper is not accessed before the examination period, or answered from outside the examination room?
- Contingency plans—what arrangements are required in the event of computer failure?

In order to determine whether these issues could be resolved satisfactorily, a 6-month pilot study was set up, to assess whether it would be possible to run online examinations in Cell Biology and Biochemistry for a group of 180 first year Pharmacy and 60 first year Medicinal Chemistry students. The main aim of the project was to work within: (i) the existing University Examinations framework; and (ii) the “SQA Guidelines on Online Assessment for Further Education” (ISBN: 1 85969 457 8) published in 2003. The specific objectives were as follows:

1. To satisfy the University Examinations Office that online examinations would be administered in accordance with the normal University Examination Regulations.
2. To satisfy the Disability Support Unit that any students registered as having special needs would not be disadvantaged by this move towards online CBA.
3. To use existing site-licensed software (WebCT, Dreamweaver and Respondus) to create online versions of examination papers.
4. To use existing School/ Faculty PC clusters to deliver online CBA.
5. To ensure that contingency plans were in place, should the computers fail, for any reason, e.g. virus attacks, server crashes etc.
6. To analyse the exam marks generated in this way and to compare them with the marks generated by more traditional marking methods.
7. To report back and to make recommendations as to the feasibility of a more widespread implementation of this technology.

Materials and methods

The project involved preparation of documentation and assessment materials for undergraduate students, plus documentation for the Examinations Office and their invigilators. This required evaluation of both the procedures involved and the materials produced. In line with SQA recommendations, this was achieved by giving students prior information about the mode of assessment and allowing the students to practise using the system for formative assessment, well in advance of the planned summative assessment. Specifically, a series of four formative assessments were scheduled during the pilot study (February–May) before conducting the final online examination in June 2005. Details of the pilot study and the various tasks are given in Table I.

Choice of software/VLE

In the past we have successfully delivered diagnostic, formative and summative assessments using a range of different software packages including Diploma by Brownstone (http://www.brownstone.net/) and an in-house version of Questionmark Perception (QMP)
These presented us with several problems. In brief, we were unable to link Diploma to our student records system, which resulted in a time-consuming administrative burden. With regard to the QMP software, the version available to us involved a complex development process (tests produced within the school had to be uploaded to the QMP server by centralised IT staff), which proved unwieldy and difficult to manage. Furthermore, WebCT Campus edition (and shortly WebCT Vista) is The University of Manchester Virtual Learning Environment (VLE) of choice. The assessment tools that it provides are more than adequate for our purpose, in terms of the question types available, and the provision of centralised support was also an important consideration. Thus, students can be automatically entered into the WebCT system from the central student records system (administered by the University Learning Technology Unit) and as course designers we have complete control over the creation, editing and administration of assessments. For these reasons we chose to deliver the summative assessments using WebCT.

Development process for WebCT assessments

The stages involved in this process have been described in detail by Pain and Le Heron (2003). However, assuming that a database of suitable questions is already available, then the four main tasks can be summarised as follows:

1. Constructing a test. Exam questions, written in MS Word, are initially saved in rich text format and imported into Respondus 3.0 (Respondus 3.0 software is a Windows application that makes it easy to create and manage assessments offline and can be used in a range of VLEs including Blackboard and WebCT). This is because questions written in MS Word cannot be imported directly into the assessment tools of WebCT and Respondus 3.0 provides a convenient interface. Tests are then uploaded into a test area in WebCT (a secure environment available only to staff with designer access) where the presentation of individual assessments was completed by adding appropriate instructions and rubrics, in line with institutional examination procedures. WebCT allows both text and HTML content to be added to various parts of an assessment to achieve this. We created this content with Macromedia Dreamweaver™ software and cut and pasted the HTML instructions and rubrics into the assessments in WebCT. Scoring for individual questions was also entered at this stage.

2. Setting test specifications and security. Once the assessments were complete in the WebCT test environment, they were then uploaded to the WebCT system. The system allows both text and HTML content to be added to various parts of an assessment to achieve this. We created this content with Macromedia Dreamweaver™ software and cut and pasted the HTML instructions and rubrics into the assessments in WebCT. Scoring for individual questions was also entered at this stage.

3. Enhance WebCT capability for functions such as automatic submission, pop up warnings, prohibiting printing, etc.; Task 10: Prepare and run Tests 3 and 4 and evaluate the online Spellcheck Tool; Task 11: Prepare final summative assessment and launch exams online end of May 05; Task 12: Review outcomes and disseminate results end of June 05
area, they were thoroughly proofread and edited as required, before being transferred to a live WebCT server (by the Learning Technology Service) ready for delivery to the students. As with all formal examinations, steps needed to be taken to ensure that assessments were only available to authenticated candidates at the permitted times. This was achieved by using: (i) the LDAP Client Login authentication on the cluster computers; (ii) timed release (tests only available on the date/time of the exam); and (iii) a test-specific password (only given out at the beginning of the exam). The authentication is the standard for the University of Manchester PC clusters and the latter two are both features of WebCT assessment tools.

3. **Running an invigilated assessment.** The project involved close liaison with staff in the Examinations Office, who timetabled the examination and arranged for invigilation in the normal way. Invigilators were given a small amount of additional training in the issues specific to computer-based examinations. In particular they were alerted to the need to check task-bars for unauthorised windows and connections. The existing computer clusters were not originally designed with examinations in mind, so although some clusters had opaque glass screens between the computers, in others it was possible for students to have a limited view of neighbouring screens. Additional measures, such as randomisation of questions, would have compromised the contingency plans, so we relied on invigilation and the normal randomisation of seat numbers carried out by the Examinations Office. As with the term-time formative assessments, the final summative assessment (for 240 students) was run simultaneously in three computer clusters in 3 different buildings. A member of staff experienced with WebCT was assigned to each cluster as a “technical invigilator” in addition to the normal invigilation staff.

4. **Contingency plans: Specification of technical failure procedures.** One of the main concerns in this study was that there should be a contingency plan in case of computer failure on the day of the examination. Although not ideal, we agreed with the Examinations Office staff that there would be a paper version of the examination paper available “just in case”. Invigilators were given a sealed envelope containing a full set of printed examination papers. In the event of a global failure on the day of the exam, such as a failure on the server, the assessment could quickly revert to a paper-based examination. In addition, students were asked to double-enter all responses: MCQ (Section A) answers were entered via mouse click into the computer and were also marked in pencil on an answer sheet designed to be scanned by an optical mark-reader (OMR sheet). Section B responses (mostly single word answers) were also entered into the computer and written onto an answer sheet. This provided additional insurance against problems arising from partial computer failure, affecting only some computers or only a part of the examination. Although the double-entry system meant we were not taking full advantage of the computer technology (no randomisation of questions), it was useful for two reasons. Firstly, it gave peace of mind for the staff in the unlikely event of a computer/power failure. Secondly, it also allowed, in this pilot study, a comparison of the different marking methodologies to be carried out.

Finally, one of the key features of this paper is the development and implementation of a Spellcheck Tool (not native to WebCT) and this is described below.

**Development of the spellcheck tool**

Automatic marking of text-matching questions (where students type one or two words into an answer box) has, in the past, proved complicated. This is because it is difficult, if not impossible, for the WebCT Quiz marking Tool to distinguish between genuinely incorrect answers and simple spelling mistakes. This seemed harsh, as we are not specifically testing spelling ability at this level. To address this problem, we therefore decided to provide access to an extensive subject-specific list of terms from which students could select (correctly-spelled) answers and cut and paste them into the appropriate box. There were approximately 1500 words in the list, which were derived from lecture notes, past exam papers and the index pages of text books recommended for this particular module. Care was taken to keep the content of the word list specific to the module and yet broad enough to reduce the chance of guesswork when selecting an answer. The words were listed alphabetically and the A to Z letters hyperlinked for ease of use. Once the model answers were written and the content of the assessments finalised, the word list was fine-tuned to contain additional distractors (plausible but incorrect answers) and also to remove any ambiguous answers. It was sometimes necessary to present the same answer in alternative forms. For instance “Citric acid cycle” could be called “Krebs cycle” or “Tricarboxylic acid cycle” in which case all the possible answers were included in the listing. The grading option was set to mark any one of these as a correct answer. The word list was published as an HTML file within WebCT and a hypertext link was included in each short answer question.
Results

Identification of a pharmacy module suitable for conversion to online assessment

The first year module “Cell Biology and Biochemistry” was chosen for the pilot study. This was because there was already an extensive database of questions for this module and also because it is taken by more students than any other module in the MPharm degree. The numbers are higher because about 60 chemistry students join the 180 pharmacy students giving a total of 240 students. The marking load is therefore especially heavy and the combined time taken for three members of staff to mark and moderate the papers, together with the time required for spreadsheet entry and checking has been estimated at almost 60 h. The basic format of the 2 h examination has changed little over the past 10 years, and because the module is relatively elementary, there have been no wholesale revisions of the material. The examination paper has two sections: Section A, consisting of 50 multiple choice questions (MCQs) and Section B containing four longer (15 min) questions, usually involving labelling a diagram, supplying missing words/descriptions (text-match) or carrying out a calculation. We currently hold an extensive database of Section A MCQs (ca. 800 questions) and a smaller database (ca. 60) of Section B questions. Examples of questions from each section are shown in Figures 1(A) and 1(B). In previous years, four to six formative assessments, in a similar format to the examination have traditionally been held during the semester. Since 2001, some of these have been computer-based. These assessments are made available to students during the revision period, with answers and feedback.

Development and evaluation of online test procedures

The pilot study began in January 2005. Although online CBA had previously been used for diagnostic and formative tests, students had no experience of online CBA under formal examination conditions. In order to familiarise students with the technology and format of the online examination, four formative tests (Proteins, Membranes, Structure and Function of Subcellular Organelles and Metabolic Processes) were made available to give the students the opportunity to practise taking online examinations. Students took the...
tests at set times, in invigilated computer clusters. “Guidelines for Invigilators” and “Login Instructions for Students” were prepared to facilitate this. In order to access the tests, students first had to log into the normal University Computer System via a Novell client and to be authenticated using their University username and personal password. Once they had access to the standard University cluster image they were able to use the Internet Explorer browser to log into the WebCT server. Students were instructed to use their same username and password in order to gain access to the WebCT server. Prior to the first test taking place, students were given information about the mode of assessment and asked (in lectures and via email) to check that they could access WebCT before taking the test. The email message gave information to help any students with login problems. On the day of the first test, after logging into WebCT, access to the test was via a final (test-specific) password issued when all students had successfully logged in. The entire login procedure took approximately 2–3 min. The online assessments proceeded smoothly and the majority of the students had no difficulties at all with using the computers. The existing computer provision was generally satisfactory. The only major issue was whether it was possible for some students to overlook one another’s screens. Some clusters were better designed than others in this respect.

Several potential problems were anticipated and were dealt with successfully. Despite the precautions described above, a few students forgot their own login username or password details. Guest logins were therefore available to the invigilators and students could be logged in by them if necessary. Computer clusters were booked at only 90% capacity in order to allow for equipment failure (the most common problem was a malfunctioning mouse).

Comparison of online CBA and Optical Mark Reader (OMR) marking of MCQs

Comparison of OMR and WebCT–generated marks for the MCQ Section A showed that these were virtually identical. All observed discrepancies were due to students having (inadvertently?) marked the OMR sheet differently from the response they typed into the computer. Generally speaking, it is probably much easier for students to type the correct response into the computer than to enter an answer onto a specially designed OMR sheet. For this reason, we believe that computer marking will eliminate most of the accidental errors that students make when recording responses in paper-based MCQ-based examinations.

Text-matching marking accuracy: Development of an online spellcheck tool

Comparison of the students’ test results (during formative testing) for Section B (text-matching) questions showed that the WebCT computer program marked more “strictly” than staff, especially with respect to spelling mistakes. The average difference in marks on this section was about 4% (range 0–10%) and this was felt to be unacceptable. We therefore developed and made available an online list of approximately 1500 (correctly spelled) words containing all the words required for the text-match and diagram labelling questions. The discrepancy between hand marking and computer-based marking in the subsequent formative tests was immediately reduced to less than 1% overall.

Enhancement of WebCT assessment functions

In the early tests, a small number of students lost track of the time, even though a clock was prominently visible on the screen throughout the test. This is probably due to lack of familiarity with a digital clock on a computer screen. A short piece of Javascript code was therefore inserted so that pop-up reminders of the time appeared half way through and 5 min before the end of the exam. It was also noted that a very few students failed to submit their responses for grading at the end of the assessment and so we modified the test settings so that this was done automatically for them, 1 min after the test or examination finished.

Running the invigilated summative assessment

In order to minimise the risk associated with computer failure, the following precautions were taken: (a) a back-up server operated throughout the examination in case of server failure; (b) WebCT was set to back up answers as the examination proceeded, so that answers already entered would not be lost in the case of a partial/subsequent failure; (c) a full set of printed examination papers was available (in a sealed envelope) in each cluster, in case of a serious University-wide computer failure. The examination was made available on the Live Server only on the day of the examination approximately 1 h before the examination start time. It could only be accessed using a test-specific password issued by the Chief Invigilator. Each student was permitted only a single attempt at accessing the paper.

Three versions of the examination were released simultaneously. The Pharmacy students took a 2 h paper in Cellular Biology and Biochemistry, the Chemistry students took a similar 1 h paper in Biochemistry and a small number of Chemistry students, registered with the Disability Support Unit, were provided with a separate version of the Chemistry paper which allowed 25% extra time before automatic termination of the examination and submission of the responses. The students in all three clusters began the examination within 3 min of one another and were individually timed, by the
computer, from the moment at which they accessed the examination paper. Candidate identity was confirmed in the usual way by asking each student to place their University ID card (with colour photograph) on the corner of the desk during the examination, for inspection by the invigilators.

Disability support

The University’s Disability Support Unit advised throughout the pilot study and during the development of the online assessments. It is worth noting that students who normally have to take examinations in the Disability Support Unit were, in this instance, able to join their colleagues in the computer clusters. Those requiring extra time were allocated seats at one end of the room, in order to minimise any disturbance when other students left the room.

Post-assessment marking, moderation and archiving of exam papers

By the time the module co-ordinator had made the short walk from the examination room to her office, the results were available. It took approximately 5 min to grade the papers and to download the results into an Excel spreadsheet. Thirty-two papers (submissions) showing failing or borderline marks were then identified, the individual papers opened on-screen and the short text answers given in Section B compared with the answers in the marking scheme.

Scrutiny of the students’ responses revealed that on just a few occasions, a student had entered a correct or partially correct response which had not been anticipated in the marking scheme. However once this had been identified, it was only a matter of moments to modify the marking scheme and to re-grade the entire set of papers. In all the later assessments, students had been strongly advised to use the word list with the copy/paste facility, in order to minimise such occurrences.

Questions which had not been answered on-line were checked against the paper-based answer sheets and it was found in every case that no answer appeared on the answer sheet either. Ten scripts out of 240 were moderated as a result of this scrutiny. In all cases the students had failed to use the word list provided correctly and their responses were either unexpected alternative answers (not in the word list) or incorrectly spelled (hand typed not copied and pasted). The moderation resulted in the addition of a small number of marks, the range being 0.5–5.0%. Only one student moved from borderline to pass as a result of this moderation. Interventions were all logged online for Quality Audit. The human intervention in the marking process took two academic staff 30 min. Complete exam papers for all students were subsequently archived onto CD and were left on the server until after the Final Examiners’ meeting in case any of the external examiners wished to see individual student papers. There should be no unauthorised access to assessment data and the Data Protection Act must be followed with regard to personal data. In addition to moderation purposes, all such data should be retained in case of queries and appeals. Backups should be made and kept in secure locations.

Discussion

In order for an institute to consider this form of online summative assessment the following are required:

(a) An appropriate VLE with built in assessment tools, e.g. Blackboard, WebCT or Moodle.
(b) The investment of time to familiarise staff with: writing questions for CBA, VLE assessment tools and other software such as Respondus and Macromedia Dreamweaver.
(c) Identification of a suitable location, equipment and appropriate staffing for online invigilated assessments.
(d) Training of students with formative tests (with feedback) prior to summative assessment.
(e) Training of staff in the Exams office (and the associated invigilators) to deal with online assessments and to handle the validation of protocols relevant to CBAs.
(f) A contingency plan in case of workstation, intranet or internet failure.

Blackboard and WebCT are the most popular commercial e-learning platforms. Since the Respondus software that we used to import assessments into WebCT is compatible with both of these VLEs, it is likely that the processes described here should be easily transferable between these managed learning environments. In principle, the extrapolation of this practice to other open source platforms such as Moodle, which offers similar features with regards to assessment tools, should also be possible.

The traditional use of CBA is in the marking of multiple choice questions and our results suggest that computer marking is preferable to the use of OMR sheets. Students made fewer mistakes in entering their responses via the keyboard or mouse than they did entering their answers onto an OMR sheet. However, CBA is sometimes said to restrict the testing to factual knowledge by limiting the question types to multiple choice, multiple response, text match or true/false formats. To test a higher level of understanding, it is important that tests can be extended to allow students to input text themselves. The Respondus software that underpins WebCT allows for text-matching and calculation questions but this can lead to marking discrepancies, since the text matching is sensitive to spelling errors. Where the spelling mistake is not
critical to understanding the subject material, this methodology could potentially disadvantage students from overseas (where English is not the first language) or students with spelling difficulties, especially dyslexia. The on-line Spellcheck Tool, developed with such students in mind, satisfactorily addressed this problem. Manual checking of all the failed and borderline students (University Examinations Policy) revealed that some of the weaker students did not always use the word list or used it incorrectly. However, since the spellcheck facility is a new tool, we anticipate improved accuracy as the tool becomes more familiar to the students. With the Glossary in place, we feel that the range of question types available to us is satisfactory for the level of skills which need to be tested in a first year module.

The use of computer clusters as examination rooms under the control of the Examinations Office worked well. Invigilators were content with the arrangements once they had been thoroughly briefed. Because of the novel nature of the examination, the invigilator: student ratio was relatively high (with 1 invigilator per 20 students) but in future years this should not be necessary. Booking multiple PC clusters did cause some extra work for the Examinations Office staff. Ideally we would have had access to a purpose built examination cluster such as a large hall where each numbered seat would be provided with a laptop computer and with wireless intranet access, configured so that external sites could not be accessed. This would resolve most of the management issues associated with summative CBA and many of the security issues as well. We encountered no evidence of cheating other than the time-honoured “notes written on hand” by one student during one of the formative tests.

The online mode of assessment was particularly successful for students with disabilities. It was especially advantageous for students who wanted to use their own personalised computer settings. Once the online spellcheck tool was in place, it was not necessary to make any special arrangements other than allowing extra time for a small group of students.

In terms of automated marking, the examination marks were available almost immediately to the academic staff. Within half an hour of checking and reviewing, compared with 60 h previously, results were finalised. However in moving towards this mode of assessment, we found that the preparation of web-based examination papers, briefing of Examinations Office and Disability Support Unit staff, preparation of the online word list, confirmation of student registration on WebCT and many other facets of the project were (for this first year only) at least as time-consuming as marking, although much more interesting. With the prior establishment of successful pilot studies and appropriate procedures, we hope that the design and administration time will in future be significantly reduced, or (at least) redistributed to administrative and technical support staff. Following on from the pilot study, we are currently extending our use of online summative CBA to several more modules in Pharmacy including Foundation Physics, Pharmaceutical Analysis and Physical Pharmacy and the use of Spellcheck Tool has been extended to include other subject-specific word lists for Physics and a numbers “listing” for calculation type questions.

Further developments in online summative assessment should aim at harmonising and securing the assessment environment. Many computer clusters in universities are not designed with assessment in mind and are too small with little or no provision for privacy. These issues are easy to address but require developments in infrastructure. Whilst the security, as far as the integrity of exam is concerned, seems adequate; there is still a requirement to combat the possibility of cheating during online examinations. In relation to this, Respondus have recently produced a trial version of a “lockdown browser” which prevents students from accessing any websites, other than WebCT, until the examination has ended.

Finally, we believe that the methodology described here provides a means of carrying out valid, reliable, secure and rapid summative assessments for first and second year Pharmacy modules. The process could, of course, easily be extended to other related science and engineering courses.

References


