

**FACTORS ASSOCIATED WITH THE
SUCCESSFUL INTRODUCTION OF
ON-LINE DIAGNOSTIC, FORMATIVE
AND SUMMATIVE ASSESSMENT IN
THE MEDICAL SCIENCES DIVISION
UNIVERSITY OF OXFORD**

Vivien Sieber and Damion Young

Factors Associated with the Successful Introduction of On-Line Diagnostic, Formative and Summative Assessment in the Medical Sciences Division University of Oxford

Vivien Sieber and Damion Young, Medical Sciences Division,
Learning Technologies (MSD-LT), Oxford, 01865 272175

Introduction and context

The University of Oxford is a large, traditional, research-led, collegiate institution. The Medical Sciences Division (University of Oxford) includes undergraduate degrees (Medical, Physiological Sciences, Experimental Psychology, Biochemistry), masters and doctoral students. Clinical students are based at hospitals and health centres throughout the region.

The Medical Sciences Teaching Centre (MSTC) was created to provide a central teaching “hotel” for three large research departments. It opened in 2002, contains a dedicated IT room with 48 workstations and one of the laboratories, normally used for laboratory classes, can be used as an IT room with a further 46 workstations for examinations. Three factors led to the identification of the need for a teaching centre: a 30% increase in the number of medical students per year. The need for additional space for research, which was gained by consolidating teaching in a single space, released underused teaching rooms within departments across the Science Area. The syllabus was revised for pre-clinical anatomy, reducing the amount of teaching in year two and introducing a three week intensive course “Principles of Clinical Anatomy” (PCA). This provides a bridge between year 1 - 3, taught in the Science Area, and years 4 – 6 taught in the geographically separate John Radcliffe Hospital.

Applications and technologies

The university VLE WebLearn (open-source, Bodington) is the primary source of administrative and teaching information which also acts as a common point of access for the range of e-learning resources. WebLearn tools are straightforward, simple to set-up and use but relatively unsophisticated. The multiple choice (MCQ) and questionnaire tools are generally used by academic or administrative staff locally within a course. WebLearn is primarily used as a spine connecting various resources to present students with a single coherent source of information and e-learning.

In 2003 Question Mark Perception v 3.0 was installed in the MSTC, access is restricted to the ox.ac.uk domain.

Range of Question Types

We use a wide variety of question types although those used in summative assessments tend to be restricted to MCQ, numerical, fill-in-the-blank/sentence completion, matrix (multiple true/false) and extended matching (EMQ). We encourage academic staff to include images, video and diagrams/graph and to think about how to encourage formative assessments, we encourage detailed feedback to reinforce the correct answers. We are also exploring further question types where the answer is not given in the question e.g. short text questions, graph drawing, etc.

EMQs

EMQs are commonly used in medical education (Case and Swanson, 2001). We have developed a semi-automated template that simplifies the process of creating an EMQ question (Figure 1). The format is characterised by a large list of equally likely answers (options) shown both in a table box at the top of the question and in a drop-down lists. The question items are presented as stems, which are often miniature case studies. Each answer option may be used more than once. Unlike the five-option MCQ, the large number of answer options in an EMQ makes eliminating incorrect answers difficult, making it more difficult to guess the correct answer.

Management of Skin Problems

A. Oral antihistamine	G. Oral methotrexate 10mg/week	M. Vitamin D analogue (Calcipotriol ointment) topically
B. Oral prednisolone 25 mg/day	H. Oral methotrexate 10mg/day	N. Cyclophosphamide
C. Topical antibiotic, fucidic acid ointment	I. Oral ciclosporin	O. Azathioprine
D. Oral tetracycline 500mg twice daily for 4 months	J. Topical antifungal such as terbinafine cream	P. A potent corticosteroid ointment
E. Oral tetracycline 500mg twice daily for 4 weeks	K. Bone protection with alendronate	Q. Ultraviolet light (UVB)
F. Oral aciclovir	L. Vitamin D orally	R. Topical permethrin (Lyclear)

For each of the following clinical presentations, select the most appropriate therapeutic intervention from the option list. Each option may be used once, more than once, or not at all.

(i) A 22-year old man has a one week history of urticaria. He has no lip or tongue swelling and the rash is worst at the end of the day but fades by the morning.

(ii) A 25-year old man has a long history of atopic eczema, asthma and hay fever. He has developed a very painful vesicular rash over the right scapula that is extending despite topical treatment with a very potent steroid (Dermovate) and oral antibiotics (flucloxacillin). He feels unwell.

(iii) A 25-year old patient has well-defined erythematous scaly plaques on the elbows and knees. He has nail pitting. He has applied a potent topical steroid ointment (Betnovate) for 4 months, but as soon as he stops treatment the plaques rapidly deteriorated and he and the GP are concerned about the long term use of steroids. He has not had any other treatments and is otherwise well.

(iv) A 25 year old man has a scaly rash with fissuring between the 3rd and 4th toes and the 4th and 5th toes on the left foot. The rest of the skin and the nails are normal. He is well.

(v) A 6-year old child with a one-week history of a slowly extending golden crusted lesion on the cheek. Her 8-year old sister has a similar lesion on the right shoulder. Their mother describes the lesion starting as thin-walled blisters that burst rapidly leaving the weeping crusted erosions. The children are well and are not bothered by the lesions.

Figure 1 Layout of an extended matching question (EMQ).

Application of CAA

Diagnostic

Across the Division all first year undergraduates take diagnostic IT (2004 -) and numeracy (2007-) assessments on-line during an IT induction workshop with the intention of helping students to identify their personal strengths/weakness and areas for potential development. The feedback and recommendations for training that are provided are dependent upon performance and tutors receive the overall marks and those of only their students. A secondary intention of the session is to introduce students to on-

line assessment itself, demonstrate the range of question formats and to check that they can access the software (QuestionMark Perception).

A second type of diagnostic test is also emerging. Third year pre-clinical students study a range of advanced topics. In pathology, students can test themselves before they begin a topic to make sure that they have the basic knowledge required to get the most out of the teaching. They are directed to appropriate resources if necessary. Similarly, first year students attending sessions in the dissection room are required to have sat a short quiz to ensure that they won't be holding up the whole group's teaching by not having covered preparatory material. Students who have not completed the test are expected to spend time with e-learning materials before admission.

Formative

The wet laboratories in the MSTC can only accommodate 70 (approximately half a year's intake) at any one time. The class is split, with half taking a practical class whilst the other half uses e-learning materials in the IT room. These classes are composed of web teaching materials provided by MedLearn (Magnolia CMS customized with Flash templates to allow simple authoring of MCQs, model answers), Case Scenarios and typically 10 Perception questions.

The introduction of PCA, which is timetabled from 9.00 – 17.00 daily for three weeks, provided opportunities to develop formative assessment further. As this course is solely assessed on-line; by an examination held at the end of each of the three weeks. Two e-learning sessions are timetabled per day. Daily teaching programmes were constructed in WebLearn with links to the appropriate part of the syllabus, MedLearn section, Case Scenarios and Perception questions for that session. Students are told that same images will be used in both formative teaching and summative questions.

Although the range of formats offered by the current VLE is limited to MCQs, short answer, or submission some groups run formative assessments on a regular basis. One masters incorporates weekly questions, based on the teaching that week, and is planning to move to summative examinations.

Summative

On-line assessments have replaced short answer papers in formal University examinations for First Bachelor of Medicine (BM) part I (year 1) and II (year 2), second BM (year 6), Medical Laboratory Sciences (year 4) and some clinical rotations.

Extent of use

Table 1 shows the number of student assessments offered (2006 – 7).

Examinations	1800 (28 papers)
Practice/Formative exams	2550
Self-test quizzes	8400
Student/assessments	12750

Table 1 The approximate number of student assessments run for students in the MSD (2006 - 7)

Process of introducing CAA

The drivers for successful introduction of CAA were primarily, gaining acceptance from academic, admin, and registry staff that the system was technically robust, that the organisation of the examinations was practically possible and secure. The barriers were attitude, rather than technology.

Procedures analogous to paper-based examinations

Figure 2 shows the processes involved in running an examination, from the appointment of external examiners to the distribution of marks for traditional paper-based examinations and on-line examinations. The administrative processes are similar until the examination is in electronic format, when administrative responsibility passes to the e-learning unit. We had to demonstrate that on-line assessments matched traditional processes in order to meet the requirements imposed by the university via departments (administrators, academics), institution: regulations, exam organisation (Schools, invigilators) and regulations (Proctors).

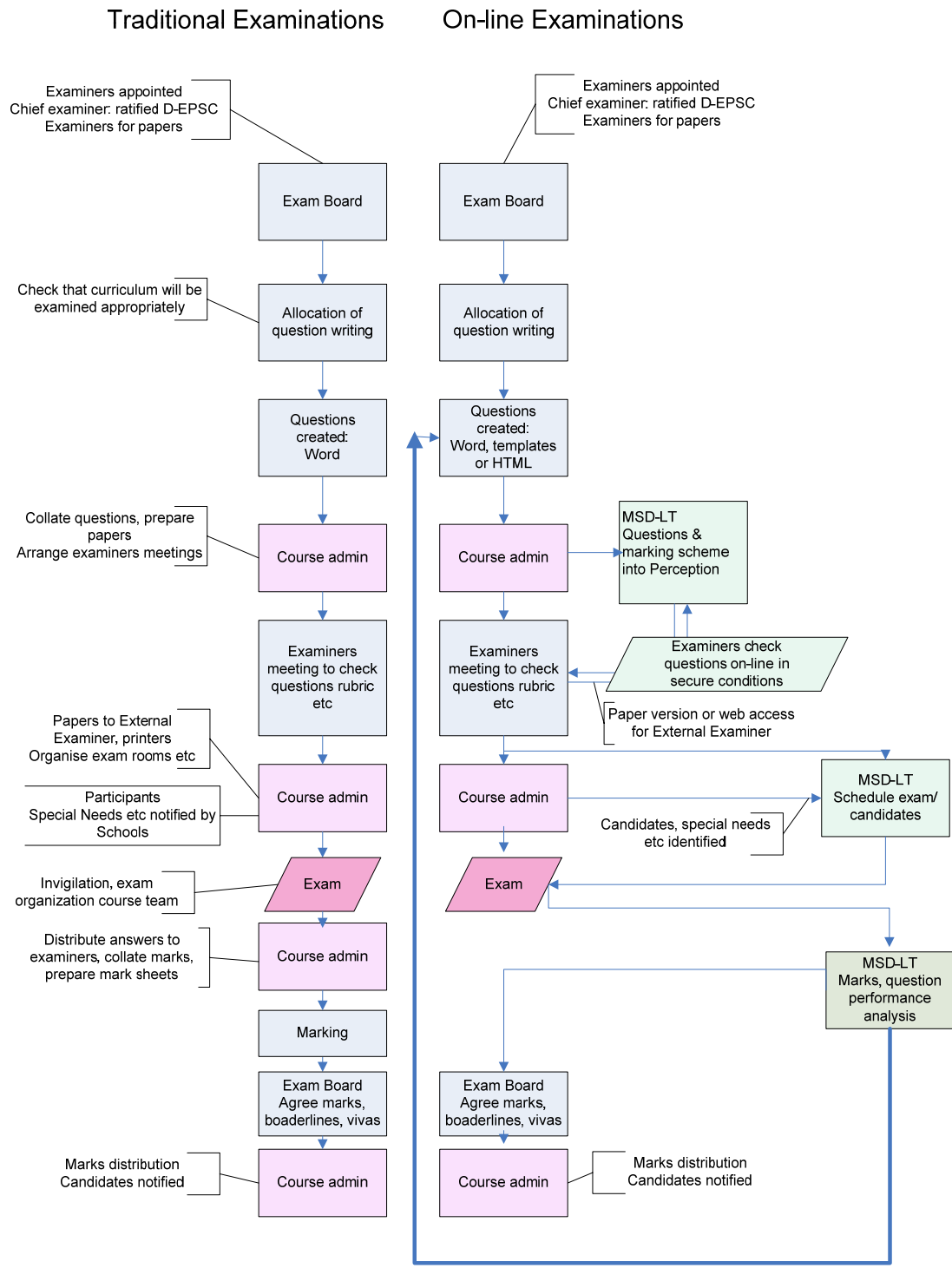


Figure 2 Processes required to deliver a conventional paper-based exam (left) and an on-line exam (right)

Regulations

Introducing CAA required a major change in examination regulations for existing courses whilst it was specified in the initial course descriptor for the new course, PCA which is a department based qualifying exam. Once PCA demonstrated that on-line assessment was technically and practically possible confidence was established and examination regulations were changed for First and Second BM examinations.

Security

The Proctors regulate and must approve security arrangements for examinations and had to approve the process for on-line examinations. An existing decree prohibits sending exam questions by unencrypted e-mail or storing them on shared servers. Questions are normally written in Word, transferred to MSD-LT on a flash stick. Once converted to Perception, academic staff physically visit MSD-LT to check their questions *in situ*. Numerous undertakings and various contingency arrangements were agreed in the event of power failure to the exam rooms, servers, and building; or a student confined to college through illness. Again, these processes mirror those of traditional paper examinations and a range of practical solutions were identified (for example, the ability to produce 160 colour photocopies quickly in the event of power failure)

Staffing

For traditional examinations, exams are organized and invigilation provided centrally, funded by top-slicing departments. At Oxford examinations are held in a dedicated building "Schools", examiners are required to be present for the first 30 minutes of any paper (in case of problems with a question), whilst senior administrators collect and deliver papers during their various stages of preparation. Any matter relating to security or permission to depart from normal regulations, for example granting extra time for a dyslexic candidate, is regulated by the Proctors office.

On-line examinations are held in the MSTC, invigilated by Schools personnel. A process, analogous to that of paper-based examinations, had to be agreed before permission to run official University examinations could be granted. This involved meeting the Proctors, and senior Schools staff, to show them round the building demonstrating how an exam would operate, the examination rooms, rooms where candidates could be held before or after an examination. Further issues arose in terms of responsibility for preparing rooms for examinations, a function normally carried out centrally by Schools. Questions of responsibility, for example should this be admin staff in the building or those responsible for the examination? Technical support during examinations is a further issue. The servers hosting Perception are maintained by the MSTC whilst assessments are created and scheduled by learning technologists working from a distant building.

Space

Approximately 150 students take formal examinations in the IT room and classroom which, between them, can accommodate about 80 students in one sitting. The student cohort is split into two groups and the examination repeated for the second cohort. This raised a number of logistical issues, primarily to remove the opportunity for collusion between students having taken the examination during the first sitting and those about to take the exam. Additional rooms are required to hold candidates prior to their exam, to hold students from the first sit until all candidates for the second sitting are accounted for and the examination started. Both groups are invigilated during

this time. Further rooms and invigilators are required for candidates with agreed special conditions, for example extra time for dyslexic students.

Positive factors leading to successful implementation

Planning

The range and format of short answer questions used in paper examinations was changed gradually to resemble more closely formats suitable for computer marking. This conscious shift, over a number of years, simplified the transition to CAA in that academic staff were not being asked to write questions in unfamiliar formats. Care is taken where there are multiple options that each option is independent of one another so that a response for the first option does not pre-determine marks for the remaining options, limiting the range of possible and plausible questions.

Space

Examinations involving 150 candidates need rather a lot of space which is, fortunately available in the MSTC. Whilst it is entirely obvious that IT rooms and workstations would be needed, the amount of additional space for students to congregate, remain segregated, special candidates and simply to leave their bags, coats and other materials prohibited from the examination room is significant. Years 4-6 has different term dates to the pre-clinical school, which means that the clinical school examinations can be organized out of normal undergraduate term when the rooms would normally be in use for teaching. Clearly the building specification providing flexibility in room size and the way they can be used is an important factor in the successful introduction of CAA and provides a further example of the importance of teaching spaces. (JISC, 2006)

Opportunities offered by CAA

Immediate feedback

Feedback is provided for questions answered following submission for diagnostic and formative assessments. With the advantage that it occurs at the time of learning, students clearly revisit assessments several times, sometimes after several months. All formative assessments, known locally as “quizzes”, are now available to all students on open access from a central location in the VLE.

Neither results nor feedback are provided for formal summative examinations. For open-access examinations, where an assessment is available for typically one week, marks and feedback are sometimes made available once the deadline has passed.

Popularity with students

Student feedback is consistently positive and they are requesting quizzes in a wider range of subjects. Medical students are highly strategic and that they are prepared to spend time using quizzes indicates their perceived value. The

learning by 'rote' that this might encourage is possibly appropriate for some parts of the medical curriculum.

Some changes have subtle effects, for example moving from a formative assessment during a timetabled e-learning class to completing the assessment in the week preceding the class means that students no longer come to the e-learning class, choosing to work on their own computers. The opportunity to work through these materials in class with an academic available to help is then lost possibly reducing the quality of the learning experience.

Saving in staff time

One of the arguments for CAA is that marking is instant and accurate, freeing academic staff to spend time on more valuable teaching opportunities. Whilst this is clearly true, significant amounts of staff time must be invested in the initial creation of questions. Question performance data is then used to review questions, identify poor discriminators and inform subsequent revisions.

There is also a saving in administrative staff time in putting exam papers together, organising marking and collating and possibly analysing marks although there are additional burdens in terms of making sure that students know where to go and when in the logistically complex double-sitting exams.

Question performance – statistical data

The value of question performance data has extended beyond identifying weak questions in that discussion includes the relative merits of different question types, distribution of questions across the curriculum. This information is informing teaching itself and further curricular development.

Consistency of marking

There is some evidence that the move from paper-based short answer questions to online assessments resulted in an overall lowering of student marks. The fact that question formats had already been gradually getting closer to computer-marked formats over previous years suggests that this was not just wholly caused by a change in format. Rather, it is thought that the computer-marking no longer provides an opportunity for a human marker to give a mark for an answer that is 'more or less right'.

Negative factors hindering implementation

Technology

We have a stable installation of Perception v 3.5 but have not yet upgraded to v 4.2 because of technical difficulties encountered setting-up and installing a clustered server.

There is normally a lag following log-in before an assessment appears on a candidate's screen. To overcome this problem small groups (typically 4) of students are asked to login at one time. Whilst assessments do eventually appear on screens, the lag is unsettling for students, academic staff and

invigilators. There have been instances where one candidate has requested the assessment 20 times during this period, further slowing the servers.

Technical breakdown and failure

Initial assessments were low stakes, so where technical failures occurred their impact was low. As CAA has been used more widely and for higher stakes formal University examinations, we developed processes that reduce the chance of breakdown. Of course there have been, fortunately, minor hitches along the way which provide an opportunity for detractors to criticise.

Technical problems caused by hardware failure have been significantly more frequent than failure of the assessment software. This is attributable to exams being held in teaching rooms that are available for student use on open-access and lack of technical support.

Changing the boundaries

Administrative units

Computer systems do not fit into neat departmental or faculty units. The distinction between managing a server, server software and assessment software is also not clear cut. This, natural, confusion can present a barrier, especially where staff are asked to explore new territory or work in areas beyond their technical experience. Whilst it may be logical and efficient for local staff to help to run examinations for the clinical school, half a mile away, it is difficult to change traditional working practice and boundaries.

Ignorance

Clearly fear of the unknown and ignorance in basic IT is a barrier for some staff, for example the invigilators were alarmed when a screen saver switched on during an examination.

Conclusion

Sharpe, Benfield, Roberts, and Francis, (2006) in their review the learner experience of blended learning, distinguish between e-learning as an add-on to traditional delivery and those where the introduction of on-line learning resulted in or was a consequence of course redesign. The development of PCA falls into the latter and on-line assessment continues to drive curriculum change as part of the cycle of course evaluation and development.

The effect of introducing CAA continues to change the way teaching is delivered in that, student pressure for more quizzes is increasing the range of subjects that provide formative assessments. The effect of some changes are subtle, moving assessment to before rather than after practical classes, appears to have changed the way students approach the e-learning element of that course.

References

Case, S.M. and Swanson, D.B., 2001. Constructing written test questions for the basic and clinical sciences. <http://www.nbme.org/publications/item-writing-manual.html> (18.02.08)

Sharpe, R. Benfield, G. Roberts, G. and Francis, R. 2006. The undergraduate experience of blended e-learning: a review of UK literature and practice. http://www.heacademy.ac.uk/assets/York/documents/ourwork/research/literature_reviews/blended_elearning_exec_summary_1.pdf

JISC 2006 Designing Spaces for Effective Learning: A guide to 21st century learning space design. <http://www.jisc.ac.uk/media/documents/publications/learningspaces.pdf> (21.02.08)