Extending the pedagogic role of online interactive assessment: providing feedback on short free-text responses

Sally Jordan, Barbara Brockbank and **Philip Butcher**
The Centre for the Open Learning of Mathematics, Computing, Science and Technology (COLMSCT)
The Open University
s.e.jordan@open.ac.uk

OVERVIEW

This case study describes an extension to the range of interactive computer based assessment question types already in use at the UK Open University (OU). The question type under development marks free-text answers of around a sentence in length, and gives tailored feedback to students on incorrect and incomplete answers. The questions are authored using linguistically based software provided by Intelligent Assessment Technologies Ltd. (IAT) and early student responses are used to refine the answer matching. A small number of these questions will be incorporated into regular summative iCMAs (interactive computer marked assignments) for a new interdisciplinary science course to be presented to students for the first time in 2008. Formal evaluation has not yet taken place, but informal evaluation indicates that most students enjoy responding to questions of this type and find the feedback useful. Conference theme : great designs for assessment.

INFORMATION ABOUT THE CLASS, MODULE OR PROGRAMME

The current project is designing assessment tasks for the level 1 (introductory) Open University Science Faculty courses *S103 : Discovering Science* (current course) and *S104 :Exploring Science* (which will replace S103 from February 2008). Each of these is a 60 CATS point course, presented twice a year and with about 1900 students per presentation. Open University students are primarily adult distance-learners, studying at home from printed text and various computer-based resources. However students are supported by tutors who give feedback on their tutor-marked assignments (TMAs) and offer face-to-face and electronic tutorials, as well as one-to-one support by telephone or email when required.

Open University undergraduate courses are open access, so students studying S103 and S104 come with a wide range of previous educational qualifications, and none. They may be returning to study after a long gap. Most students study part-time alongside employment, and although most students are UK-based, increasing numbers are at other locations in the EU and elsewhere. A significant percentage of OU students have disabilities, and students include those in prison and at remote locations such as oil-rigs and ships. All of these factors must be considered when assessment tasks are being designed.

S103 and S104 are interdisciplinary courses, embracing the scientific disciplines of physics, chemistry, biology and Earth Science. It should be noted that, although the current work is being done in the context of these courses, one of the project's aims is to evaluate the wider usefulness to the Open University of assessment tasks of this type.

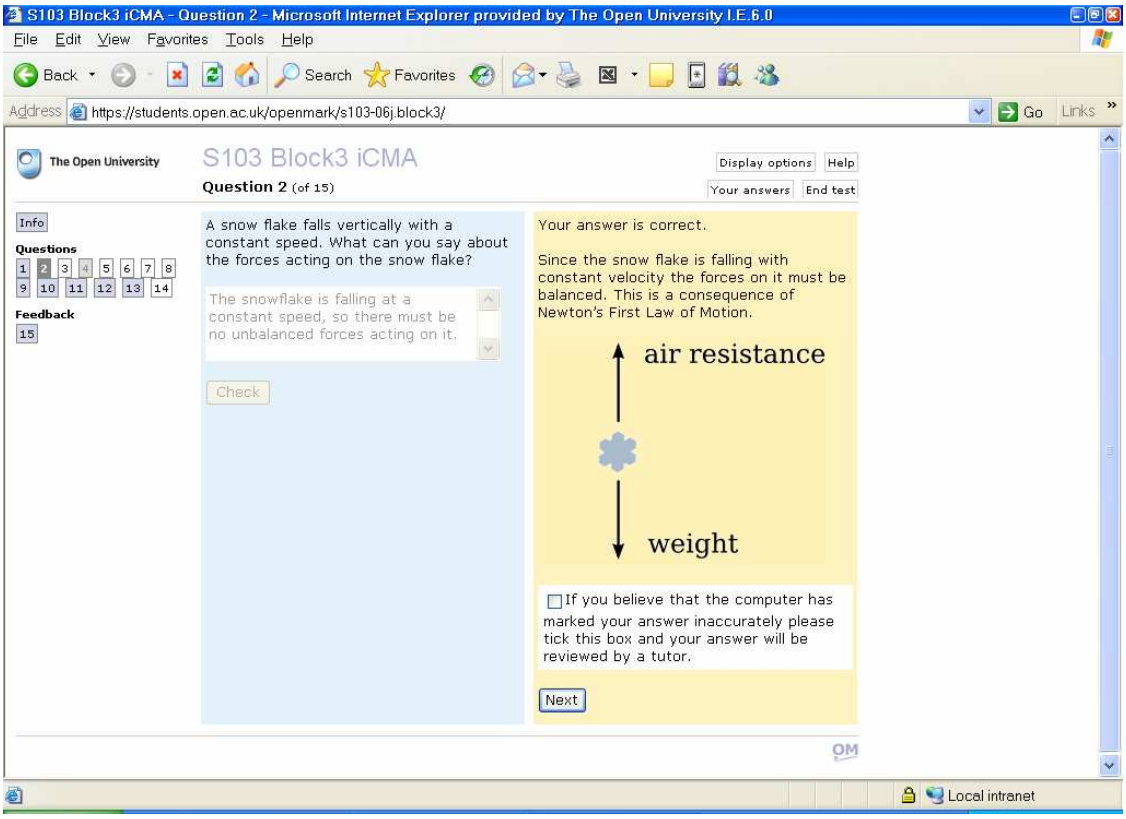
DESCRIPTION OF THE CASE

The Open University will present two new large population introductory science courses (*SDK125 : Introducing Health Sciences* and *S104 : Exploring Science*) to students for the first time in 2007/2008, and both of these will use regular summative (but low stakes) iCMAs (interactive computer marked assignments). iCMAs will only form one part of the courses' assessment strategy (alongside conventional tutor marked assignments and, in one case, an end of course examination). The iCMAs will use the OU's 'OpenMark' system, which enables

students to be provided with targeted feedback on their responses to questions (not just multiple choice questions). The OpenMark system allows students three attempts at each question, with increasing amounts of teaching feedback provided after each attempt.

The OpenMark question types currently available mean that the tasks that can be assessed are limited. In particular questions requiring free-text answers are limited to those requiring numbers, symbols or answers of no more than a word or two. In an attempt to explore other possible question types, a pilot study, jointly funded by the Centre for Open Learning of Mathematics, Science, Computing and Technology (COLMSCT) and the OU's VLE Project, is using an authoring tool supplied by Intelligent Assessment Technologies Ltd. (IAT) to write questions requiring free-text answers of up to around 20 words in length – typically a single sentence. The IAT authoring tool (Mitchell *et al*, 2003) is linguistically based, which means that an answer such as 'the dog bites the man' is recognised as being different to one such as 'the man bites the dog'. In the example shown in Figure 1, answers such as 'there are no unbalanced forces'. 'the resultant force is zero', 'the weight is equal and opposite to the air resistance' and 'the forces are balanced' should all be marked as correct, but 'the forces are unbalanced', 'there are no forces' and the 'forces act in opposite directions' are incorrect or incomplete. The sophistication of the authoring tool means that these responses are indeed correctly matched, as are the vast majority of incorrectly spelt responses, responses in poorly written English, and many of those which are expressed in unexpected ways.

Figure 1 A typical (correct) student answer to a question requiring a short free-text answer.

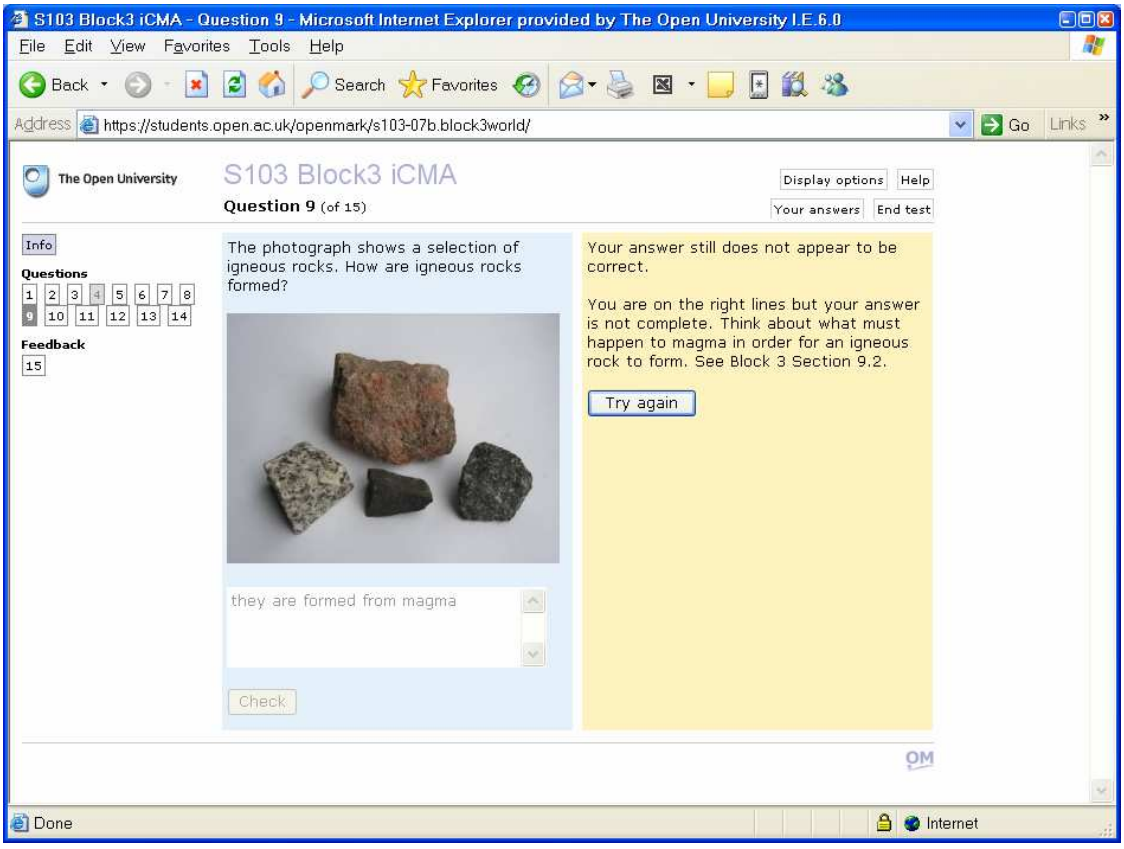


The Open University's pilot use of the IAT software has two major differences from that of most users of this software and other linguistically based short answer free-text systems (e.g. Sukkarieh and Pulman, 2005). Previous users have used student responses to paper-based questions in order to provide appropriate answer matching for the computer based version. The OU pilot is gathering responses for use in improving the answer matching by offering the questions online to students and staff on one presentation of the existing

Science Faculty course *S103: Discovering Science*. Responses to the developmental versions of the questions are monitored regularly and the answer matching is amended whenever necessary (this can be done very quickly, whilst the system is live). The questions are then released to students on the next presentation of the same course (4 months later) for more extensive evaluation, though still in a purely formative setting.

The second novel focus of the Open University's use of short answer free-text questions is the emphasis being placed on the provision of instantaneous teaching feedback. The IAT questions are currently being presented online to students via the OpenMark system, so a decision was made to replicate OpenMark's three stage feedback. At present the system attempts to recognise answers as correct or specifically or generally incorrect, and then to provide pre-written feedback including (as for conventional OpenMark questions) a reference to the relevant course material. Generic feedback is provided from OpenMark, but feedback targeted to specific student misunderstandings is generated within the IAT system. So, in Figure 2, the feedback in response to a student answer of 'from magma' (which is not considered a sufficiently complete response) has been generated from within the IAT system.

Figure 2 Targeted feedback on an incomplete student response to a question



The Open University's trial of online assessment questions requiring short free-text answers is in its very early stages, with four batches of questions released to students in their developmental phase, but only two of these batches released to students on the subsequent presentation of the course. An investigation has just begun into student perceptions of the questions and their use of the feedback provided.

RATIONALE IN TERMS OF EDUCATIONAL IDEAS

It is widely recognised that feedback on assessment tasks has an important part to play in underpinning student learning, encouraging engagement and promoting retention (see for

example Yorke, 2001). Gibbs and Simpson (2004) articulated eleven conditions 'under which assessment supports students' learning', and seven of these conditions relate to the provision of feedback. Online assessment provides an opportunity to give virtually instantaneous feedback, thus 'feedback is provided quickly enough to be useful to students' (Gibbs and Simpson condition 6). However, providing feedback which is targeted to an individual student's specific misunderstandings is more of a challenge for computer based assessment. Multiple choice questions enable specific feedback to be provided in response to predefined responses, but these questions narrowly constrain how students may respond (Sukkarieh and Pulman, 2005) and concern has been expressed over their reliability, especially when used for summative purposes (Lawson, 2001).

Throughout the OU's thirty-five year history, assessment has been seen as an integral part of the learning process. Tutor marked assignments (TMAs), although usually graded and thus having a summative function, also provide an opportunity for a student's tutor to give detailed and personalised feedback and a summary of points to assist with subsequent assignments. Feedback and 'feed forward' at this level is considered particularly important for students who are studying at a distance and have limited opportunities to participate in face-to-face or electronic tutorials. The OU also has a longstanding commitment to the use of appropriate media, including computers, to support student learning. Computer based assessment, initially in the form of batch-marked multiple choice computer marked assignments (CMAs), has been used since the early days, and since the 1990s, interactive formative questions of several types (including drag and drop and numerical input) have been provided to Science Faculty students on CD-ROM and DVD-ROM.

In September 2002, a 10 CATS point course *S151: Maths for Science* was presented to students for the first time. This course is studied over a relatively short time period (10 weeks – 3 months) and does not allow the opportunity for tutors to provide feedback to students via TMAs. However, the subject matter, and the fact that many of the students studying the course are lacking in mathematical confidence, meant that it was considered particularly important to provide rapid and meaningful feedback. It was therefore decided to use online interactive assessment for both formative and summative purposes. This enables students to be provided with instantaneous and targeted feedback on their responses to the questions, only about 25% of which are multiple choice (Jordan and Swithenby, 2005). The aim is to simulate 'a tutor at the student's elbow' (Ross *et al*, 2006), pointing out the student's error as specifically as possible, and providing a suggestion for how it might be corrected. The student is allowed three attempts at each question, with increasing amounts of teaching feedback provided after each attempt. The student is thus able to 'act upon the feedback to improve their work or their learning' (Gibbs and Simpson condition 6).

The regular summative use of interactive computer marked assignments in *S104: Exploring Science* will enable students to be provided with instantaneous and targeted feedback on their answers, but the regularity of these iCMAs and the fact that they carry some credit will also encourage students to keep up to date in their studies. Thus the 'assessed tasks distribute student effort across topics and weeks' and 'sufficient feedback is provided, both often enough and in enough detail' (Gibbs and Simpson conditions 2 and 5). It is hoped that this will encourage student engagement with the course and so increase student retention. However, it is easier to write meaningful iCMA questions for some parts of S104 (typically the more numerate parts) than others. The current project is exploring the feasibility of extending the available question types to include ones requiring short free-text answers.

Learning from students

Monitoring student responses to interactive online assessment questions, essential in the development of sophisticated questions of the type described in this case study, can also be used to provide valuable feedback to academics about student misunderstandings. Analysis of student responses to the conventional OpenMark questions in summative assignments for *S151 : Maths for Science* has provided valuable insight into the mathematical misunderstandings of adult distance-learning science students (Jordan, in press). Even in

the developmental phase, student responses to the short answer free-text questions has provided similar insight. For example, in answer to the question 'What does an object's velocity tell you that its speed does not?', a small but significant number of students gave the answer 'It tells you about a change in direction'. This answer was initially marked as correct (it is similar to the correct answer 'It tells you the direction in which the object is moving') but it is *not* a correct answer, and the source of the misunderstanding was tracked to an unfortunate section heading in the printed text. In both this case and the more systematic analysis of 'Maths for Science' questions, increasing understanding of student misconceptions has enabled changes to be made both to the assessment questions and to subsequent Open University teaching.

EVALUATION

The current project is in its early stages. Formal evaluation will include:

- a human/computer marking comparison;
- an investigation into student perceptions of the questions, and their use of the feedback provided;

A later stage of the project will compare the current system, based on computational linguistics, with a simple 'bag of words' system. This is likely to be less accurate in its marking, but may lend itself to different question types (including, paradoxically, those requiring longer answers) and the provision of different types of feedback.

For the present, evaluation is limited to analysis of student usage of the developmental system and their responses to an informal 'feedback' question at the end of each iCMA. Of the 53 students who responded to the feedback question at the end of the first iCMA, 41 reported that they had enjoyed the iCMA, with just one saying 'no, it irritated me'. 40 reported that they had found the feedback useful or very useful, with just one student saying 'not really'.

Each time a student completes a question, they are provided with an opportunity to report if they think that the computer has marked their answer inaccurately. Most of the times this box is checked the student response has actually been marked correctly (and the student answer is usually wrong!) but questions which are relatively 'open' tend to generate the most uses of the 'inaccurately marked' box, perhaps because students fail to appreciate that they need to provide a precise answer (for example, if they are asked to compare intrusive and extrusive igneous rocks, they need to say *which* has bigger crystals).

The authors acknowledge that the students currently completing the iCMA are a self-selecting group, doing so as an 'added extra' to their studies. More reliable data will be obtained from the more formal evaluation, especially when the free-text questions are embedded within summative assessments.

Institutional implications: authoring questions

At present just two OU employees (the first two authors of this case study) are writing questions using the IAT authoring tool, but their experience of this is being monitored carefully, since it is hoped that other academics will author questions in the future. Framing appropriate questions, especially questions which assess understanding not simply recall, is perhaps a greater challenge for the question authors than using the authoring tool to generate appropriate answer matching. The time taken to write suitable questions and answer matching, and to modify this in the light of student responses, should not be underestimated. Using the IAT authoring tool to write questions of this type does not require any special linguistic abilities but it does require a logical approach to the authoring process.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of the Centre for the Open Learning of Mathematics, Computing, Science and Technology (COLMSCT) and the Open University VLE Project, and the assistance of many individuals, especially Stephen Swithenby (COLMSCT Director) and Tom Mitchell (Intelligent Assessment Technologies Ltd.).

REFERENCES

Gibbs, G. and Simpson, C. (2004) Conditions under which assessment supports students' learning.*Learning and Teaching in Higher Education*, **1**, 3-31.

Jordan, S.E. and Swithenby, S.J. (2005) Online summative assessment with feedback as an aid to effective learning at a distance. In *Proceedings of the 2004 12th International Improving Student learning Symposium, Birmingham,* ed Rust, C., pp 480-485. Oxford, UK: Oxford Centre for Staff and Learning Development.

Jordan, S.E. (in press) The mathematical misconceptions of adult distance-learning science students. *CETL-MSOR Conference, Loughborough, 11-12 September 2006.*

Lawson, D. (2001) Computer assisted assessment (CAA) in relation to learning outcomes. http://ltsn.mathstore.ac.uk/articles/maths-caa-series/oct2001/index.shtml (accessed 26th February 2007)

Mitchell, T, Aldridge, N, Williamson, W. and Broomhead, P. (2003) Computer based testing of medical knowledge. *10th Computer Assisted Assessment Conference, Loughborough, July 2003.*

Ross, S.M, Jordan, S.E and Butcher, P.G. (2006) Online instantaneous and targeted feedback for remote learners. In *Innovative assessment in Higher Education*, ed Bryan, C. and Clegg, K.V., pp 123-131. London, UK: Routledge.

Sukkarieh, J.Z. and Pulman, S.G. (2005) Information extraction and machine learning: auto-marking short free text responses to science questions. *Proceedings of the 12th International Conference on Artifical Intelligence in Education.*

Yorke, M. (2001) Formative assessment and its relevance to retention. *Higher Education Research and Development*, **20**(2), 115-126.