

Chapter for *'Innovation in Assessment'*, Bryan, C. and Clegg, K.V. (eds.), Routledge Falmer, London, 2006

Online instantaneous and targeted feedback for remote learners

Shelagh Ross, Sally Jordan and Philip Butcher (The Open University)

Providing feedback on assessment in distance education

The challenge of this project was to develop assessments for a distance learning course in 'maths for science', such that no hand-marking was required, students were given immediate feedback tailored to their own answers and multiple attempts at each question were permitted.

The importance of feedback for learning has been highlighted by a number of authors, emphasising its role in fostering meaningful interaction between student and instructional materials (Buchanan, 2000, p.199), its contribution to student development and retention (Yorke, 2001), but also its time consuming nature for many academic staff (Gibbs, Chapter 1). In distance education, where students work remotely from both peers and tutors, the practicalities of providing rapid, detailed and regular feedback on performance are vital issues. Four of Gibbs's 'eleven conditions' under which assessment supports student learning (Gibbs and Simpson, undated, and Gibbs, Chapter 2) are particularly worthy of examination in the distance education context.:

- a) the provision of sufficient feedback (in terms of both frequency and detail) [No 5]
- b) the provision of timely feedback [No 7]
- c) the delivery of feedback in such a way that students have to engage with it [No 10]
- d) the provision of feedback that can be acted upon by the student in future learning tasks [No 11]

The nature of assessment in relation to its 'formative' and/or 'summative' purpose is also important. As noted by Brookhart (2001) and others, these terms are not used consistently in the literature. Most authors subscribe to Sadler's (1999) definition of formative assessment as being explicitly 'intended to provide feedback on performance to improve and accelerate learning', with the further, often unstated, implication that such assessment is not used to allocate grades. Our usage here conforms to both the explicit and implicit parts of this description, with the proviso that formative assessment involves each student's answers being independently checked (whether by a person or a computer program), thus distinguishing it from *self*-assessment, in which students draw their own conclusions by comparing their own answer with some kind of specimen answer. Summative assessment, on the other hand, counts towards an award.

As noted by Gibbs in Chapter 1, at the UK Open University (OU) the tutor-marked continuous assessments have always been seen as a crucial element of learning and the principal means by which individual feedback is supplied to students during the course. The majority of courses in the OU's Science Faculty last for eight months with students typically submitting assignments every six to eight weeks. These are annotated by a tutor with corrections, comments and suggestions to help with future elements of the course, as well as a grade. Since these assignments fulfil both formative and summative purposes they conform to Boud's (2000) characterization of 'Double Duty' assessments. Such tutor-marked assessment can satisfy Gibbs's conditions a) and

d) above. Whether condition c) is met depends largely on the extent to which students are willing to work through feedback that necessarily relates mainly to a task that has already been completed. Condition b) is only partly satisfied by tutor-marked assessment, in that each assignment relates to study carried out over a period of many weeks and the marking process itself usually takes several weeks. This turnaround time makes feedback via hand-marked continuous assessment unsuitable for courses lasting less than about three months.

Partly for reasons of turnaround time, and partly due to their expense, tutor-marked assignments have not been considered suitable for the Science Faculty's programme of 'short courses'. Instead, credit for these 10 CATS point courses is awarded on the basis of a single open book, un-timed, end of course assessment (ECA) that students do at home. The development of a new Level 1 course entitled '*Maths for Science*' within this programme brought a number of the issues raised above into sharp focus. Presented four times a year, it has recruited large numbers of students: currently about 900 students complete the course in each calendar year. Students work on printed course material and a CD-ROM at home, without face-to-face tuition or meetings with fellow students. However, they do have telephone and asynchronous computer conference access to specialist advisers and other students.

The team developing the course was very aware of the fact that the majority of the students would be at or near the beginning of their OU studies, and that many such mature learners bring emotional 'baggage' arising from previous negative experiences of maths. It would therefore be particularly important to build students' confidence, despite the lack of continuous assessment to provide on-going support. Furthermore, in the case of *Maths for Science* the inevitable delays associated with hand-marking and administration of large numbers of ECA scripts were considered unacceptable: students need rapid feedback on their mathematical strengths and weaknesses before proceeding to other science courses. For the distance education market, it was felt that the best way of providing this kind of immediate and personalised feedback would be to use an interactive Web-mediated format.

A Web-based assessment strategy

The rationale for the project was to build a complete teaching strategy in which feedback on both formative and summative parts of the Web-based assessment would be a crucial part of the learning process. Figure 1 shows one example of the kind of teaching comment that may be provided.

Figure 1: *An example of targeted feedback, showing the combination of a response-specific explanation of the student's error with a pointer to the course material.*

The feedback aims to simulate for the student a 'tutor at their elbow'. To this end, an entirely formative 'practice' assessment (PA) can be accessed throughout the course, as many times as the student wishes, with the option of generating different numbers and variables for each question on each visit. As well as providing subject-specific practice, it was hoped that the PA would help to reduce anxiety by allowing the students to become familiar with the computer system (Sly, 1999). The summative ECA is available for the second half of each presentation of the course. Both the PA and the ECA have the following features in common:

- they provide individualised, targeted feedback, with the aim of helping students to get to the correct answer even if their first attempt is wrong. Thus in the example shown in Figure 1, the student is told that they have taken the wrong tack, and given a reference to the relevant section of the course material. This feedback appears immediately in

response to a submitted answer, such that the question and the student's original answer are still visible.

- they allow students to make up to three attempts at each question, with an increasing amount of teaching feedback being given after each attempt. Figure 1 shows feedback provided after a second attempt.

This instantaneous feedback, together with the opportunity to use it immediately (Boud, 2000, p.158) by amending answers if necessary (Clariana, 1993), is designed to encourage students to engage with the feedback and to learn from it actively (OLTC, 1996).

Formats for interactive feedback and multiple try answers had previously been developed in other OU science courses for formative assessments presented on CD-ROM. However, in pedagogic terms, this assessment strategy was different from anything that had preceded it, in that it:

- continues the *teaching* feedback through the ECA, helping students to amend their answers even *during the process of summative assessment*;
- allows *partial credit* with the mark awarded decreasing with the number of attempts and thus with the amount of feedback given. This is not a replication of the partial marks typically awarded during script-marking. Rather, the strategy is based on answer-matching and feedback that helps the student to make another attempt if necessary
- allows students to form a reasonably accurate picture of their strengths and weaknesses across the subject matter, both from on-going feedback and from a final summary screen. If at this point they are not satisfied with their own performance, they can choose not to submit their ECA for credit, but to defer and take a new ECA with the next cohort. If they choose to submit the ECA immediately, the feedback informs them of weaknesses, which they may then address before starting their next course.

As well as fulfilling the pedagogic aim for this particular course, the Open University also saw the introduction of a Web-based assessment system as institutionally desirable in other ways. The expense of hand-marking large numbers of scripts is removed. In principle, bank of templates and questions types can be built up. When appropriate, random-number generation can produce many similar questions at minimal cost. Barriers to study of time and geography are further reduced. However, many of these perceived advantages have to be balanced against important design issues and against the requirements for very robust systems.

Developing the assessments

The design of the assessments had to satisfy five main drivers:

1. *interactive feedback*: both the formative PA and the summative ECA had to provide students with response-specific comments that would help them correct their misunderstandings and learn while doing the assessment.
2. *easy to use and robust software*: given the nature of the course, many students are inexperienced internet users, with fairly low specification machines, modem rather than broadband connections, and a wide range of software and ISPs. The assessments must sit robustly on top of this variety of hardware and software. Students must be able to take a break (or suffer a dropped connection) and be returned on reconnection to the exact point reached on their last visit.
3. *simple input*: it is widely acknowledged (e.g. Beevers and Patterson, 2002) that finding an easy way for students to input mathematical expressions is a significant challenge for computer-aided assessment in mathematics. For example, superscripts were enabled by the provision of a button, as illustrated in Figure 2.

4. *minimal use of multiple choice formats*: Although inputting problems can be avoided by setting multiple choice questions, there are well recognised difficulties inherent in such formats (see, for example, Lawson, 2001). Especially important for the context described here is the fact that multiple choice questions do not easily assist in the diagnosis of many student misconceptions: the question setter can write distractors based on common errors, but cannot anticipate all possible answers, and of course a student may pick the correct answer simply by guesswork. Multiple choice formats were therefore used only when answer formats would have been too difficult for students to input.
5. *institutional record-keeping*: every attempt at a question had to be securely recorded, despite any breaks in communication, and marks transferred to the University's examination data-handling system.

Figure 2: *A partially correct answer requiring a superscript, with targeted feedback. In this case, feedback and pointers to course material both cover two independent mistakes in the student's second attempt at the question.*

The practice assessment is a cornerstone of the overall pedagogic strategy, and has to fulfil a number of purposes. Its main aims are:

- to engage students during their work on the course, by giving them additional questions for practice in a format in which they receive *specific* feedback (e.g. 'you appear to have calculated the first rather than the second derivative'), as opposed to the printed self-assessment questions which have full answers but may not help students to diagnose their own difficulties or misconceptions.
- to build learners' confidence in the course and themselves by receiving immediate positive feedback on questions they have tackled successfully.
- to make sure that students appreciate all the elements of a 'correct' answer (e.g. not just a value, but a value given to appropriate precision with units of measurement included, as illustrated in Figure 2).
- to provide students with an objective formative assessment of their performance, counteracting the tendency of many students to delude themselves in self-assessment. We have seen many examples in pre-course diagnostic maths tests and in tutorial sessions of students believing they have only made a slip when in fact it is a fundamental error, and not engaging with questions they think they can do. Colleagues in other universities report similar attitudes to paper-based solution sheets, whereby students don't fully analyse the answers, or think 'I could have done that' only to discover when it comes to an exam that actually they cannot.
- to provide a 'dry-run' for the ECA in terms of ensuring that students have the necessary software in place, understand the input mechanisms and have confidence in the technology.

Student behaviour and perceptions

In accordance with standard OU practice, the first intake of students was surveyed by a postal questionnaire, covering all aspects of the course. A very small number of students in a later intake completed the general Assessment Experience Questionnaire (AEQ) discussed by Gibbs in Chapter 2. However, the most detailed information came from a cohort numbering about 500 selected to receive a postal questionnaire focussing specifically on the assessment; 270 completed questionnaires were returned. As well as responding to multiple choice questions, these students

were encouraged to write freely about aspects of the assessment format that they particularly liked or disliked.

Despite many respondents' relative inexperience with technical Internet issues, the software proved suitably user-friendly. Responses from those who did not complete the ECA confirmed that while a very small number of students are prevented from submitting by technical difficulties (and alternative provision can be made available when technical problems prove insurmountable), the main causes for non-submission are of the kind revealed in every OU retention study, with lack of time the most reported.

The pedagogical drivers for this assessment format were the immediate teaching feedback and the summary screen that would enable students to make a realistic assessment of their mathematical competence before submission and ahead of entry to other science courses. It was slightly disappointing that the latter was mentioned far more often than the former in students' written questionnaire responses. It seems that for a significant number of students the importance of 'feedback' relates mainly to performance in terms of their idea of the pass/fail divide (what Higgins *et al* term 'strategic consumerism'), rather than to learning.

Nevertheless, it was a vindication of the strategy that 74% of respondents who had submitted the ECA agreed with the statement that 'the interactive feedback associated with the ECA questions helped me to learn' and only 5% disagreed. 79% agreed that the feedback had helped them to amend their initial inputs to obtain the correct answer at a subsequent attempt, with only 6% disagreeing. 88% agreed that the immediate acknowledgement of correct answers increased their confidence, again with just 6% disagreeing.

Tied up with students' comments about their overall performance are their perceptions of the marking process. Some students are convinced they would gain significantly better marks if they got credit for 'working'. Given that a proportion of the questions are essentially 'one liners', and that others are likely to be done wholly on a calculator, this assumption is not valid in all cases. It also ignores the fact that many students manage to correct their initially wrong answers and get credit for having done so. This issue probably becomes most contentious when answers have several elements (e.g. value, plus the precision to which it is quoted and/or units of measurement), though some students clearly have an erroneous belief that they would not be penalised by a human marker for what they regard as 'minor errors'. In fact, since every input made by a student is recorded, borderline ECAs and those with unusual distributions of responses are manually re-appraised.

Although it had been hoped that the PA would help students both to learn and to distribute their effort sensibly throughout the course (cf. Gibbs's condition 2 in Chapter 2), roughly three quarters of them accessed it for the first time just before starting the ECA. This is consistent with the fact that most students said they used the PA mainly as a 'mock' for the ECA, rather than as a bank of additional questions. It has become very clear that students are much more likely to submit the ECA if they have previously accessed the PA. This may be due to various factors operating in combination, as with the findings of Sly (1999) that students undertaking formative practice tests significantly improved their mean mark on summative tests compared to those who did not do the practice tests. Those submitting are a self-selected group, so may include all the 'better' or more motivated students who would naturally choose to do the PA. Students who have done the PA may approach the ECA with more confidence in their ability to handle both the subject matter and the computer interface.

The time spent by students on the PA and the ECA varies widely, in terms of both the total number of hours and the number of online sessions over which their effort is distributed. This time range is illustrated in Figure 3, which also shows that there is no clear correlation between time spent and success in the ECA.

Figure 3: *time spent on the summative ECA by one cohort of students.*

Conclusion

The effort required to devise and program these assessments was substantial, although this has been justified by the large student enrolment over a number of years and the longer-term saving in staff time. The project may be considered a success in terms of the student approval of both the interactive feedback and the multiple-try format. However, students' appreciation of the feedback within the summative component was focussed more on their attainment standard than on its learning function. Every input to the assessments made by every student is electronically stored and available for analyses that would not be possible for script-marked or multiple choice assignments. From a teaching point of view, this is contributing substantially to our understanding of common student misconceptions.

Acknowledgements

We thank Spencer Harben, Greg Black, Peter Mitton, Pat Murphy, Isla McTaggart and Stuart Freake, for their contributions to the project. A grant of £20 000 from the eOU Strategy Group towards development costs is gratefully acknowledged.

References

- Beevers, C.E. and Patterson, J.S. (2002) Assessment in mathematics. In Khan, P. and Kyle, J. (eds.) *Effective teaching and learning in mathematics and its applications*. Kogan Page, London.
- Boud, D. (2000) Sustainable assessment: rethinking assessment for the learning society, *Studies in Continuing Education*, Vol. 22, No. 2, 151-167.
- Brookhart, S.M. (2001) Successful students' formative and summative uses of assessment information, *Assessment in Education*, 8 (2), 153-169.
- Buchanan, T. (2000) The efficacy of a World-Wide Web mediated formative assessment, *Journal of Computer Assisted Learning*, 16, 193-200.
- Clariana, R.B. (1993) A Review of Multiple-Try Feedback in Traditional and Computer-Based Instruction, *Journal of Computer-Based Instruction*, Vol.20, No.3, 67-74.
- Gibbs, G. and Simpson, C. (undated) Does your assessment support your students' learning? <http://www.open.ac.uk/science/fdtl/documents/lit-review.pdf> (accessed January 2005)
- Higgins, R., Hartley, P. and Skelton, A. (2001) Getting the Message Across: the problem of communicating assessment feedback, *Teaching in Higher Education*, Vol.6, No.2, 269-274.

Lawson, D. (2001) Computer assisted assessment (CAA) in relation to learning outcomes
<http://ltsn.mathstore.ac.uk/articles/maths-cao-series/oct2001/index.shtml>
(accessed November 2004)

OLTC [Open Learning Technology Corporation] (1996) Learning with Software: Pedagogies and Practice: Operant Conditioning
<http://www.educationau.edu.au/archives/cp/04j.htm>
(accessed December 2004)

Sadler, D.R. (1998) Formative assessment: revisiting the territory, *Assessment in Education*, 5 (1), 77-85.

Sly, L. (1999) Practice tests as formative assessment improve student performance on computer-managed learning assessments, *Assessment and Evaluation in Higher Education*, Vol. 24, No. 3, 339-343.

Yorke, M (2001) Formative assessment and its relevance to retention, *Higher Education Research & Development*, Vol 20, No 2,

S151 Practice Assessment - Microsoft Internet Explorer

File Chapter Help

Chapter 8 Question 4

You have one more attempt at this question.

Ten Dunlin (small wading birds) are trapped and the length of their bills measured. The results are listed below. What is the median of these data?

Bill length/mm
34.0
38.0
32.5
37.5
32.0
33.5
31.0
36.5
34.5
35.5

median = mm

Your answer is still incorrect.
You have calculated the mean, rather than the median. The calculation of median values is discussed in Section 8.2.6.

ok

Click on 'ok' or press Enter

33 of 42

S151 Practice Assessment - Microsoft Internet Explorer

File Chapter Help

Chapter 3 Question 5

You have one more attempt at this question.

If $L = 6.1 \times 10^{30} \text{ W}$ and $F = 4.6 \times 10^{-10} \text{ W m}^{-2}$, find d in the equation

$$d = \sqrt{\frac{L}{4\pi F}}$$

You should give your answer in scientific notation, with the correct number of significant figures and the correct SI base units. *N.B. You do not need to understand the underlying science or the units used in order to answer this question.*

$d =$

Normal Superscript

Your answer is still incorrect.
You have given your answer to an incorrect number of significant figures.
In addition, your answer does not include units.
Significant figures and rounding in calculations are discussed in Section 3.1.2.
You can use the units given for F and L to work out the correct units for d (see Section 3.5.4).

ok

Click on 'ok' or press Enter

13 of 42

