Practice

Computer aided self assessment – an effective tool

Roy Lowry

Centre for Chemical Sciences, University of Plymouth
Email: R.Lowry@plym.ac.uk

Received 11 July 2005, accepted 26 September 2005

Abstract: Computer aided assessment (CAA) has been used to provide students with a system for formative self-assessment. Students could access the material at any time and two levels of feedback sought to guide further learning. Comparison of two groups of students (those who used all of the material and those who used none) revealed that those students who used the system performed significantly better in the end of module summative assessment. The difference in performance was not observed between these two groups in a similar assessment on material that was not supported by CAA. This points towards the conclusion that the CAA system has made a positive impact upon the learning experience of the students. [*Chem. Educ. Res. Pract.*, 2005, **6** (4), 198-203]

Keywords: effective learning, formative assessment, feedback self-assessment, multiple-choice assessment.

Introduction

Computer Aided Assessment (CAA) is being used increasingly to provide a quick method of marking summative assessments for large groups of students. Whilst this can be very effective in saving time for staff, it does mean that the feedback present in the more traditional coursework assignments is lost. For effective learning, it is important that students can try out their understanding and obtain constructive criticism, so that the learning cycle is complete (Kolb, 1975). There is, therefore, a need to provide feedback to individual students via formative assessments. However, this is an exercise that is very time consuming, especially for large classes. Various authors have attempted to address this using CAA. Whilst this can be purpose-written software for a particular course or module (e.g. Hunt, 2002), software suites for assessment are now available either as part of a managed learning environment (e.g. WebCT, Blackboard) or for assessment alone (e.g. QuestionMark, WebMCQ).

The use of CAA for formative assessment affords considerable advantages. For the institution, this is mainly time saving (after the initial 'cost' of setting up the system), but the advantages for the students are more numerous:

- to give students feedback;
- to guide student effort;
- to diagnose problems in learning;
- to give students experience in assessment methods.

For an excellent discussion of the issues and impacts of using CAA for formative assessment see Charman (1999).

One possible problem with the implementation of a computer-based system (formative or summative) is that the results could be affected by students' prior computing experience or anxiety about using the technology. However, work comparing assessment performance using

both computer and paper based multiple-choice tests (Lee, 2001) has demonstrated that there is no measurable effect.

This project sought to provide a formative self-assessment mechanism via CAA for our Stage I BSc Environmental Science students. The syllabus covered included solution chemistry and thermodynamics, and was a 10-lecture 2-workshop course, which was 50% of a 20-credit module in the first semester. These students traditionally shy away from this material as it is seen as 'hard' science and contains more mathematics than other disciplines. Thus, any mechanism that increases their interaction with the subject could be beneficial. Also, since this module was at the very beginning of their course, the course team wanted to encourage students to become 'deep' rather than 'surface' learners (Marton, F., 1976). It was hoped that self-assessment would help to promote this. One hundred and four students completed the module.

Method

As the summative assessment for this module was a multiple-choice test, it was decided that the formative self-assessments should also be in the multiple-choice format. In this way, students could practise both their understanding of the subject and the mode of the summative assessment to come. Using a support mechanism that mimics that of the final assessment should also improve the validity of the assessment as 'false negatives' due to lack of familiarity with the method are minimised.

The CAA system used was 'Perception' from Question Mark Computing. The system comprised a series of programs for question creation, assessment compilation and delivery/monitoring of the assessments. This last program was mounted on a server (the others are local programs) and supplied the assessments to the students as web pages, allowed different levels of security and collected data as to which students had performed the assessments and their scores (both at assessment and question level). Whilst the suite of programs has been improved over recent years, the implementation of the system at the University of Plymouth is very similar to that reported by Zakrezewski (1999) at the University of Luton.

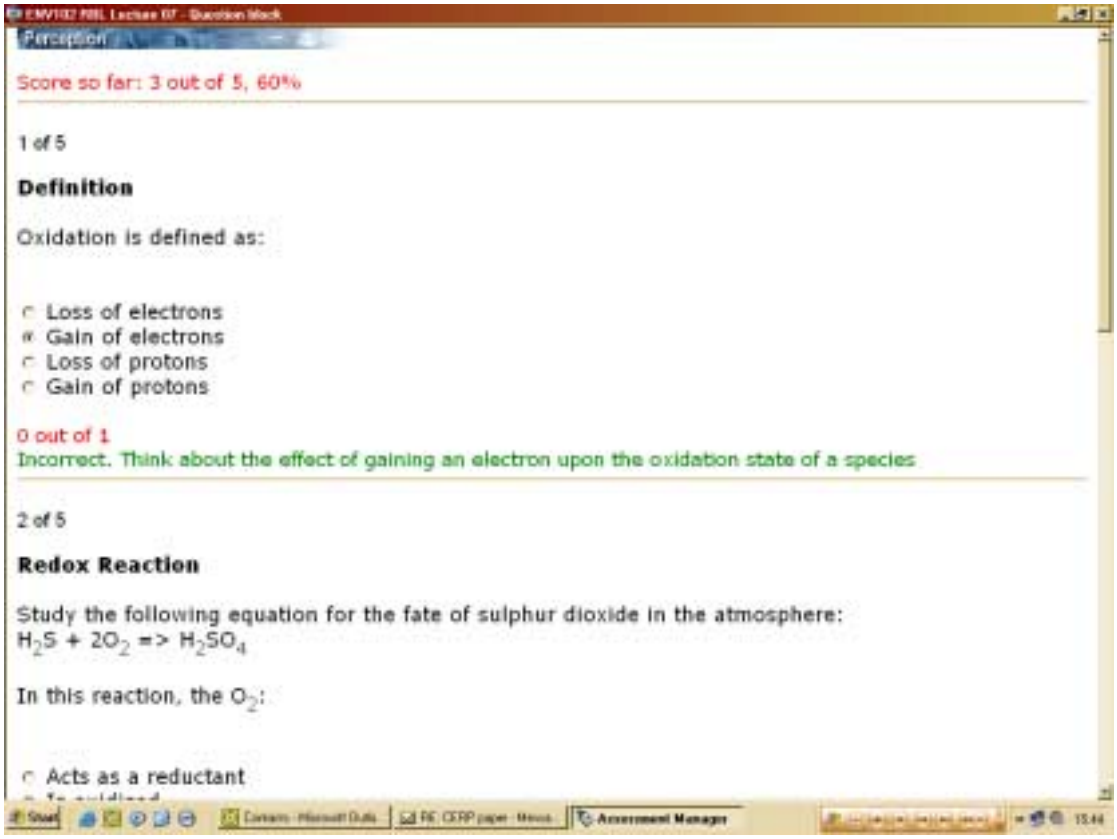
For each lecture, a short (5 question) self-assessment test was compiled. The test became available via the web at the end of the lecture and could be accessed whenever the student required and as many times as the student wanted. Access was not restricted to just the computers on campus and thus students could use the system from wherever they happened to be – especially during the vacation. At the end of the self-assessment test, the student was given feedback at two levels. The overall feedback gave the score gained and a message based upon the mark. These messages were as follows:

- 100% to 95% "*All correct. Well done! Celebrate your success and then use your valuable study time upon another area*."
- 94% to 50% "*There are some misunderstandings in your knowledge of this area. Use the individual feedback to identify these before further study. Come back to repeat the test if you wish.*"
- 49% to 0% "*Your understanding of this area requires further work. Go through your lecture notes with a text book and then re-try the test*."

After this page, students could access feedback relating to their answer to each question. This feedback was constructed so that it explained why an answer was incorrect, but not so that it gave the correct answer. The idea here was to get the students to consider their understanding and not just to memorise. Figure 1 shows a screen shot of one of the self-assessment tests. The end of year summative test consisted of 30 questions to be attempted

within 45 minutes and was held under normal examination conditions. It was marked by an optical mark reader.

Figure 1. Screen shot of one of the CAA formative assessments



Results

Although the self-assessment tests were completely formative and non-compulsory, 42% of the students used them all at least once and 65% used at least some of the tests. This high level of use may be due to the similarity between the formative and summative assessment methods (thereby increasing the value to the students). Another possibility is the relatively mature nature of this group of students; 32% of this cohort of students were over 21 years of age when they started the course. It is interesting to note that the majority of those students who accessed all of the tests started using the system within the first two weeks. It would seem that encouraging students to form the habit of testing their understanding early in the course is beneficial as it reduces the possibility of such activities being squeezed out by assessments in other modules. Another factor that may have increased the use of the system was that students were told that the system was not connected to the university record system and that it was 'safe' to try out their understanding multiple times without fear that this would affect future grades.

In order to gauge the success of the system, the summative results of two groups of students were compared. The first group (Group 1) had accessed all the self-assessment tests at least once, whilst the second group (Group 2) had not attempted any of the tests. This information was automatically collected by the system and was downloaded into an Excel spreadsheet for analysis. The average result for the end of module summative assessment for these two groups was found to be 52% for Group 1 and 36% for Group 2. Figure 2 shows the frequency histogram of these two groups and clearly there is a difference in performance. To

check the significance of this difference, a one-tailed t-test was applied to the data using an Excel spreadsheet, a summary of which is given in Table 1. This confirms the difference to be statistically significant at the 95% confidence level.

Figure 1: Summative assessment results

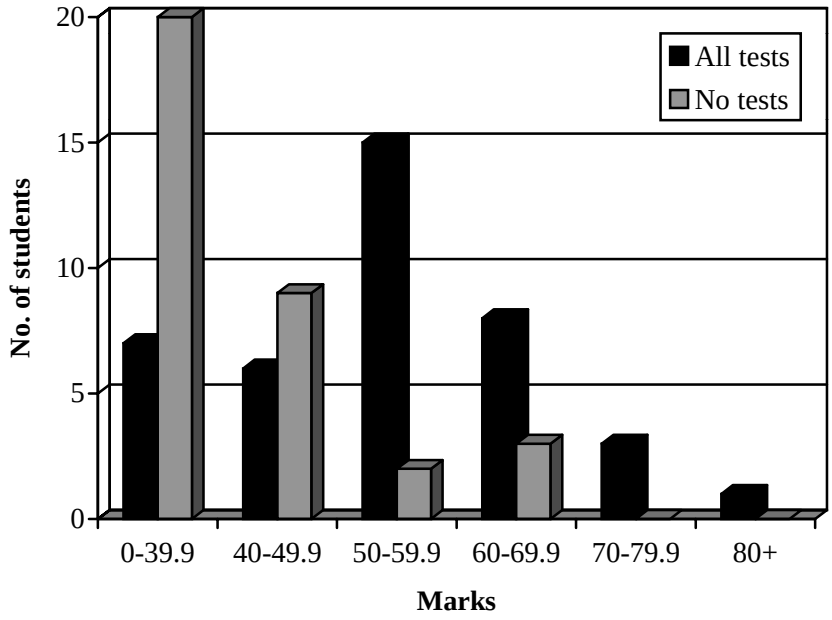


Table 1. Statistical evaluation of results

	CAA supported		Unsupported	
	Group 1	Group 2	Group 1	Group 2
No. of students	40	34	40	34
Average score (%)	52.3	36.1	66.8	63.3
Standard deviation	14.1	13.3	18.6	18.9
t_{stat}	5.001		0.804	
t_{crit} (one tail)	1.667			

Comparison with other material

It is possible that the difference shown between the two groups was due to the nature of the students themselves. Since the groups of students were self-selecting, it could be that the more able or dedicated students elected to use the self-assessment tests, whilst others did not. If this were true, then these students could have performed better than their colleagues without the self-assessment CAA system being in place. To check whether or not that was the case, a similar analysis was performed on the results of another summative assessment, which dealt with material that was similar (inorganic chemistry) but was not supported by the self-assessment CAA system. In common with the previous material, the optical mark reader was used to mark the summative assessment for this series of lectures. The summative marks for

this material were analysed using the same groups as previously, i.e. those students that had used all of the available self-assessment tests for the previous material (Group 1) and those that had not used any (Group 2). Using these same groups of students, the means of the two groups differed by only 3.6%. The subsequent t-test confirmed that the difference between these two groups under these circumstances was not significant.

Examining results from these different areas of material results in a comparison between material that has a support mechanism in place (self-assessment via CAA) and material that does not. It could be said that any support mechanism should result in an increase in performance of the student body. The data presented in this paper demonstrate that it is likely that self-assessment CAA is suitable as a support mechanism for material of this type and is an effective tool. However, it should be noted that there were inevitable differences in both the type of material covered and assessed (it is possible that the summative assessment for the supported material was less discriminatory) and also the presentation of the material (two different lecturers). These differences ensured that the unsupported material was not a totally effective 'control'.

A similar study (Peat, 2002) based upon a variety of delivery modes for computer-based self-assessment for biology students also found that formative material provided via computer had a positive influence upon learning. Student feedback over a number of years was consistently positive, but this work did not include a statistical analysis of summative results.

Student perception

Whilst no formal gathering of data on how the students felt about the system was performed, many students who had used the system offered their opinions via various methods, including feedback from the student representatives on the staff-student liaison committee, and the end of module student questionnaire. One of the most common comments was that the system increased confidence. It would seem from this that many students think that they do not understand a concept, when in fact they do. This means that they spend time studying the particular issue when their time could be better spent reading about another area (which they think that they do understand, but possibly do not). In addition, this increase in confidence can mean that students are more open when they arrive in a future lecture, as their perception of the subject as a whole influences their learning (Johnstone, 1997).

Other comments included the usefulness being able to access the material at any time and not having to come into the university. These factors may well have contributed to the considerable uptake of the self-assessments.

Conclusion

The data collected in this study indicate that providing a CAA system for self-assessment positively affects the learning of those students who choose to use it. Once the question database has been set up, no further intervention is required by staff, apart from scheduling the assessments. Given this, the time required to set up the system is more than justified by the improvement in the learning of the students. However, this study does not contain a true control group and hence it is possible that the effect seen upon summative assessment performance was due to other factors.

Further work

The system has been successful in providing formative assessment and therefore we will move to using it for the summative assessment in the coming academic year. This will allow

a much greater range of question types to be used in both the formative and summative modes by removing the limitations imposed by the use of the optical mark reader for the summative assessments. In particular, multiple response (more than one correct answer from a variety of choices), ranking (place items in the correct order) and numeric questions are suitable for material of this type. For a discussion of possible question types and their relative merits/demerits see Clarke, 2001. The system will also be extended to cover all of the material in the module and the results of the study reported here used to encourage more students to use the system.

References

Charman D., (1999), Issues and impacts of using computer-based assessments (CBAs) for formative assessment in Brown, S., Race, P. and Bull, J. (eds) *Computer-assisted assessment in higher education* London: Kogan Page.

Clarke A., (2001), *Designing computer-based learning materials*, Gower, Aldershot, 61-78.

Hunt N., Hughes J. and Rowe, G., (2002), Formative automated computer testing, *British Journal of Educational Technology*, **33**, 525-535.

Johnstone A.H., (1997), …And some fell on good ground, *University Chemistry Education*, **1**, 8-13.

Kolb D.A. and Fry R., (1975), Toward an applied theory of experiential learning, in C. Cooper (ed.) *Theories of group process*, John Wiley, London.

Lee G. and Weerakoon P., (2001), The role of computer-aided assessment in health professional education: a comparison of student performance in computer-based and paper-and-pen tests, *Medical Teacher*, **23**, 152-157.

Marton F. and Saljo R., (1976), On qualitative differences in learning I. Outcome and process, *British Journal of Educational Psychology*, **46**, 4-11.

Question Mark Corporation Home Page http://www.questionmark.com/uk/home.htm (accessed September 2005).

Zakrezewski S. and Bull J., (1999), The mass implementation and evaluation of computer-based assessments, *Assessment and Evaluation in Higher Education*, **23**, 141-152.