INSTRUCTIONAL DESIGN AND ASSESSMENT

Providing Formative Feedback From a Summative Computer-aided Assessment

David J. A. Lewis,^a and Robert D. E. Sewell^b

^aInformation Systems and e-Learning Services, University of Glamorgam, Wales, United Kingdom

^bWelsh School of Pharmacy, Cardiff University, Wales, United Kingdom

Objectives. To examine the effectiveness of providing formative feedback for summative computer-aided assessment.

Design. Two groups of first-year undergraduate life science students in pharmacy and neuroscience who were studying an e-learning package in a common pharmacology module were presented with a computer-based summative assessment. A sheet with individualized feedback derived from each of the 5 results sections of the assessment was provided to each student. Students were asked via a questionnaire to evaluate the form and method of feedback.

Assessment. The students were able to reflect on their performance and use the feedback provided to guide their future study or revision. There was no significant difference between the responses from pharmacy and neuroscience students. Students' responses on the questionnaire indicated a generally positive reaction to this form of feedback.

Conclusions. Findings suggest that additional formative assessment conveyed by this style and method would be appreciated and valued by students.

Keywords: formative feedback, evaluation, assessment, reflection, pharmacy

INTRODUCTION

There are many well-developed computer-aided assessment systems available that lend themselves to use for both formative and summative assessment in higher education. Most virtual learning environments (VLEs) such as *BlackBoard* and *WEBCT* include some form of inherent assessment tool, normally based around differing styles of multiple-choice questions. In addition to integrated systems, there are also standalone assessment products, probably the most popular and strongly developed of these being *Perception* from Question Mark Limited.¹ The Welsh School of Pharmacy has been using this computer-based assessment system in many evolving forms since 1992; first, as a standalone system, then as a LAN-based system,² and finally as a Web-based system.³

The use of computer-aided assessment has made many developments and innovations in teaching possible. As early as 1995, the Welsh School of Pharmacy utilized a LAN-based learning environment, with e-learning content, computer-aided assessment, and course information such as lecture notes and module information.³ These

facilities were included as part of the teaching program before the Internet was widely used as a medium for educational delivery. Since the early 1990s, computer-aided assessment has formed an integral part of the assessment methods, but was only used for summative assessments.

Since traditional, hands-on, laboratory-based practical experimental classes are progressively being replaced by computer-simulated experiments,⁴ the logical way of replacing the associated traditional written experimental report is to employ some form of computer-based assessment. This tandem approach has led to a considerable savings in staff time and improvements in student learning.⁵ The computer-aided assessment system was further developed to cover summative assessment in many areas of the pharmacy curriculum, initially pharmacology, then clinical pharmacy. This resulted not only in considerable savings in academic staff time, but also in a wealth of data regarding detailed student performance on individual test questions. These data were utilized to provide timely verbal feedback to groups of students on their group performance, but was never used on an individual basis. The form of the assessments dictated a strict question order together with shuffling of possible responses. Early in the development of computer-aided assessment, on-screen feedback during the test was discouraged because of the possibility of cheating

Corresponding Author: David J. A. Lewis. Address: Information Systems and e-Learning Services, University of Glamorgan, Wales, United Kingdom, CF37 1DL. E-mail: djalewis@glam.ac.uk

by individuals who could have copied the questions and the feedback provided for correct answers, thereby seriously disadvantaging honest students. A sad consequence was the provision of only a minimum level of individual on-screen feedback on the students performance in the computer-based assessment, which in most cases amounted to no more than their final percentage score. This had long been unsatisfactory, but methods of question shuffling, answer shuffling, or question banking to minimize any possibility of cheating could potentially have led to student complaints of the about the relative difficulty of their tests. Furthermore, the creation of extensive banks of new questions every year would negate any advantage for academic staff in time saving. So the possibility of creating individualized formative feedback from summative assessments was investigated in this study. The ideal situation would be to return a printed or electronic copy of their test to each student, with full feedback. This would be logistically difficult and require a completely new set of test questions each year. So a compromise between providing the students with formative feedback and preventing students from removing copies of the questions used in the summative assessment had to be reached.

It was in this context that the study—based on the processing of student results in a summative, computer-based assessment, and then returning interpreted feedback to the students in a lecture room situation—was examined. Student opinion on the form of the feedback, and how this would relate to their future educational practice was assessed using a questionnaire. The authors' objectives were to examine whether this was a positive development for students and whether students would use the feedback from the computer-based test to direct future study in this area.

DESIGN

Summative Computer-aided Assessment

In this pilot exercise, the test was based on a first-year pharmacy and neuroscience student computer-based experimental simulation entitled "Simulations of Pharmacological Experiments on the Guinea Pig Ileum."⁶ Students were expected to complete 5 sections of this large e-learning package, covering the agonists: (1) experimental set-up; (2) agonist theory; (3) constructing dose-response curves; (4) comparing dose-response curves; (5) agonists on the guinea pig ileum. The test, based on this learning material, consisted of 30 questions, divided into 6 per section. Each of the questions and their relationship to the section of the e-learning package to which they related was discussed and carefully revised by the authors until the format of the assessment was

finalized. Additionally, a few practice sample questions were available online to familiarize the students with the form of the assessment. The assessment was then attempted by 132 students, including 14 neuroscience students and 118 pharmacy students divided into groups of 30-40. Percentage scores for total correct answers were displayed on screen to the students at the end of the test. Once all the tests had been completed, the results were processed and fed back to the students during the same semester week at a scheduled session planned to provide timely, general feedback to the group and personalized written feedback to individual students.

Formative Feedback

A spreadsheet itemizing the student responses to each question on the computer-based guinea pig ileum test were retrieved from the computer-aided assessment system. These data were processed to provide personalized student results and feedback. The content of the student feedback is shown in Appendix 1. The papers were returned to the students individually to ensure fairness and privacy.

The students were shown a PowerPoint presentation on interpreting the test findings and given the remainder of the lecture time to review their results. The students were told that a result of "excellent" meant that they could move on to the next part of the self-directed learning. A result of "satisfactory" meant that the student had performed well enough to pass the assessment, but before their examination or moving on in the module, it would be advisable to review the section. A result of "unsatisfactory" meant that the student should completely review the section before moving on. This system of formative feedback was designed to inform the student and assist their process of reflection, while preserving the security of the electronic test.

Question processing was performed utilizing a "look-up" table, a cross-reference table, which in this case, matched each question to its corresponding topic. Each question had a particular topic associated with it, based on one of the 5 sections of the agonist component. Each student's result was processed and the question matched to its topic, and for each topic the individual scores were summed. All questions are weighted equally. The students were then given prescriptive feedback based on their score within that topic. An example of part of the set of source data for an individual student's results is shown in Table 1. This was cross-referenced with the reference table of question against topic, illustrated in Table 2 to generate the result in Appendix 1.

The data in Tables 1 and 2 demonstrate that within the first 5 questions, the specified student scored 50% for the

Table 1. A Truncated Example of the Source Data for Each Student's Responses and Performance on the Guinea Pig Ileum Computer-Based Test

Participant user name	sph4js
Participant group	03-04-MPharm 1 Group A1
Participant details	0412345 John Smith
Date/time started	November 24 2004 – 11:12
Date/time finished	November 24 2004 – 11:26
Status	Finished normally
Total score	27
Maximum score	53
Percentage score	51
Questions presented	30
Questions answered	29
Question description	Question 01
Actual score	2
Question description	Question 02
Actual score	2
Question description	Question 03
Actual score	1
Question description	Question 04
Actual score	0
Question description	Question 05
Actual score	3

topic “Experimental Setup” (2 out of 3 on question 2 and 0 out of 1 on question 4). Rather than giving an exact percentage score for each topic, it was decided that score bands would be translated into statement-based feedback. Consequently, a score below 40% would result in a grade of “unsatisfactory”; a score of 40%-70%, “satisfactory”; and over 70%, “excellent.” This type of formative feedback was intended to aid the students examination revision later in the year by indicating whether they needed to spend additional study time on any of the individual activities covered in the agonists e-learning package. It also provided actual feedback relating to students’ overall understanding of the topic.

This was the first time the students completing this assessment had received anything more than a percentage score, and the first time they were provided with individual formative feedback on this type of summative assessment. The problems associated with test security and preservation of the test questions made it impossible to provide detailed question feedback to the students, either on screen or on paper, applying the previous style of summative computer-aided assessment. Providing the feedback in printed form rather than on screen made it more convenient for students to reflect on their need for further study.

Table 2. Cross-reference Table Used for a Pharmacology Assessment

Question	Topic	Max. Score
Question 01	Comparing Dose-Response Curves	2
Question 02	Experimental Setup	3
Question 03	Agonists on the Guinea Pig Ileum	2
Question 04	Experimental Setup	1
Question 05	Constructing Dose-Response Curves	5

A questionnaire was employed to evaluate the form and method of feedback procedure and style. Thus, a simple questionnaire (Appendix 2) was included with the formative feedback sheet and the questions were worded simply, since complex phraseology and double negatives are invariably difficult for students to interpret, particularly if their first language is not English.

One hundred twenty-eight responses were returned from a cohort of 132 first-year pharmacy and neuroscience students. The undergraduate group of pharmacy students was made up of a majority of female students, approximately 70%-80%. Ten to fifteen percent of the students were from countries outside the European Union, with the majority coming from former UK colonies in East Africa and South-East Asia. Of the UK students, approximately 40%-50% were Asian or Afro-Caribbean, and approximately 20%-30% were from the South Wales area. The neuroscience students were mainly UK students (40% male), from a mixture of ethnic backgrounds and geographical locations. Ninety-five percent of both student groups were 18-20 years of age.

ASSESSMENT

Analyses were performed on scored responses (mean $\pm$ sem) recorded on a 0 - 4 Likert scale using the following statement scoring response scheme: strongly agree = 0, disagree = 1, no opinion = 2, agree = 3, strongly agree = 4.

The survey results were generally positive, with high levels of approval, especially for questions 1 and 2, for which approval approached an 80% “strongly agree” level. There was no significant overall difference between responses obtained from either pharmacy or neuroscience students. Both sets of students were studying a common pharmacology module (*Principles of Pharmacology*), but they were on divergent course programs, ie, a 3-year, science-based BSc in neuroscience versus a 4-year vocational masters degree in pharmacy. The examination entry scores of the 2 groups were comparable, with both having

a minimum of 3 “A” levels of grade B or above. However, pharmacy students had a higher grade average overall; thus, their expectations may well have been different by virtue of the particular requisites of the course they were pursuing. However, the matching survey responses of the 2 groups at least tended to refute this possibility with respect to formative feedback from the computer-aided assessment (Figure 1).

DISCUSSION

Based on our interpretation of both the qualitative and quantitative data returned in the survey, the students liked the extra feedback (Question 1) and wished to see it continued (Question 2). The authors believe there was a lack of understanding concerning the summative and formative terminology used in Question 5, leading some students to comment on the nature of the test rather than on the formative feedback. The students agreed that the nature of the feedback was a helpful aid to their revision (Question 3), but it would have been useful to ask whether they would use it to direct future study. They agreed with the form and amount of information provided (Question 4), though this question yielded the lowest agreement score in the survey. Students’ lack of understanding of University regulations prohibiting the return of test papers led to a poor score in this area with a significant number of students’ qualitative responses including a request to return that their test be returned to them.

Also in their qualitative survey responses, students requested that more information be included in the feedback, such as their exact grade, a percentage of correct answers per topic, and more bands on the grading/feed-

back scale. Many responses were positive in the area of electronic assessment, and additional assessment with feedback indicating that more formative assessment of this style would be welcome.

A proportion of responses were received regarding the design and format of the actual assessment, and modifications to the assessment were made by the authors as a consequence of these comments. However, these format changes were outside the scope of the present study.

Formative assessment is a relatively modern concept in education,⁷ and the contrasting terms *formative* and *summative* serve a functional purpose in describing the assessments.⁸ Tutors must change the way they teach in order to make best use of formative methods, which are generally more time consuming to perform. Therefore, the primary commodity that dictates the lack of formative assessment is teaching time, and specifically, academic staff time. As the number of pharmacy students taught by the Welsh School of Pharmacy as well as the number taught in other established and newly created schools of pharmacy in the UK continues to increase, the drive to maintain standards has never been greater.⁹ The time required for academic staff evaluation of assignments may be prohibitive, and the necessity to provide individualized feedback for every assignment strains even the most creative of minds. Undergraduate teaching has to compete against research for academic staff time in the majority of research-focused and Research Assessment Exercise-aware academic departments (a peer evaluation conducted every 5 years of the quality of research in the UK¹⁰). Consequently, time spent grading single undergraduate assignments is often limited by other work pressures. Students often fail to understand exactly how much time is required: completed examination from a typical cohort of 120 students could take 5 days of faculty time to grade. Distractions can cause grading time to be substantially increased, and a 2-3 week turnaround for a typical piece of coursework is not unreasonable. The grading process could be simplified by the utilization of other electronic methods of assessment, not just computer-aided assessment. However, other methods of grading, such as processing of spreadsheets or optical mark reading, must be weighed against established criteria.

In the future, rather than paper-based feedback, personalized feedback could be automatically e-mailed to students. Alternatively, the feedback might be uploaded to the VLE, whereby authenticated access would allow students to view only their individualized feedback. Both methods of delivering feedback would still be accompanied by the feedback lecture on test performance, the usual monitoring of academic progress by personal tutors, and individual student access to the module leader on a tutorial basis.

Figure 1. Returned survey responses of pharmacy students (n = 114) neuroscience students (14) to the questionnaire relating to provision of formative feedback from a summative computer aided pharmacology assessment. Analysis was performed on scored responses (mean ± sem) recorded on a 0-4 Likert scale using the statement scoring response scheme (9) shown on the ordinate.

The production of detailed feedback in this way would either reduce routine enquiries, or enable students to better target their questions during these feedback opportunities.

The authors foresee a growth in the number of methods available for providing formative assessment and thus formative feedback to students, improving the reflective and diagnostic aspects of any assessment process, whether through traditional or computer-based methods.

CONCLUSIONS

The effectiveness of providing formative feedback for summative computer-aided assessment was tested at the Welsh School of Pharmacy. The results indicated a positive response from students to this form of formative feedback, which directed their learning to areas in which they had performed well or badly, specifically related to the computer-based simulation that they had studied. Questionnaire results suggested that students would welcome more assessments with feedback conveyed using this style and method. The quantitative responses indicated that additional information within the feedback would be appreciated and valued by students.

REFERENCES

1. Perception from QuestionMark Computing Limited. Available at <http://www.questionmark.co.uk/> Accessed October 1, 2005.
2. Sewell RDE, Stevens RG, Lewis DJA. Multimedia computer technology as a tool for teaching and assessment of biological science. *J Biol Educ.* 1994;29:27-32.
3. Stevens RG, Lewis DJA, Sewell RDE. Controlling a multimedia CAL network in undergraduate pharmacy education using a courseware management system. *Am J Pharm Educ.* 1999;63:174-8.
4. Stevens RG, Sewell RDE. The replacement of pharmacology practicals by multimedia technology [Educational Career Supplement]. *Pharm J.* 1993;251:E11-13.
5. Sewell RDE, Stevens RG, Lewis DJA. Pharmacology experimental benefits from the use of computer-assisted learning. *Am J Pharm Educ.* 1996;60:303-7.
6. Lewis DJA, Stevens RG, Sewell RDE, Broadley KJ, Gilbert M. Simulations of Pharmacological Experiments on the Guinea Pig Ileum, COACS, University of Bath.
7. Scriven M. The methodology of evaluation. In: Tyler R, Gagne R, Scriven M, eds. *Perspectives of Curriculum Evaluation.* Chicago: Rand McNally; 1967;39-83.
8. Black P, William D. In praise of Educational Research: formative assessment. *Brit Educ Res J.* 2003;29: 623-38.
9. Taylor K, Bates I, Editorial: Pharmacy student numbers are bound to affect educational standards, *The Pharmaceutical Journal*, Vol 271 No 7271 p546, October 18, 2003.
10. RAE 2008. Research Assessment Exercise. Higher Education Funding Council for Wales, Higher Education Funding Council for England, Scottish Funding Council, Department for Employment and Learning. Available at: www.rae.ac.uk/. Accessed January 1, 2005.

Appendix 1. An example of a student feedback paper sheet.

TEST PH1105 Feedback

Details = 0412345 John Smith

Group = 0405-MPharm 1 A1

Login Name = sph4js

Time Started = "November 24 2004 – 11:12"

Time Finished = "November 24 2004 – 11:26"

Ended = Finished normally

Score = 27

Maximum Score = 53

Percent Score = 51%

Your result on topic Experimental Setup was Excellent

Your result on topic Agonist: Theory was Satisfactory

Your result on topic Constructing Dose-Response Curves was Satisfactory

Your result on topic Comparing Dose-Response Curves was Unsatisfactory

Your result on topic Agonists on the Guinea Pig Ileum was Unsatisfactory

A topic score below 40% is classified as unsatisfactory, between 40% and 70% is satisfactory, and above 70% is excellent

Appendix 2. A simple questionnaire survey format used for student feedback evaluation. The questionnaire was distributed to both pharmacy (118) and neuroscience (14) first year undergraduate students.

Please give your opinion on the PH1105 Ileum test Feedback Sheet					
Question	Strongly Agree	Agree	No Opinion	Disagree	Strongly Disagree
Q1. I think that giving feedback like this is a positive development					
Q2. I think that all the tests should give feedback like this					
Q3. I will use this to aid my revision					
Q4. I was happy with the form and amount of information					
Q5. I think this is a good way of obtaining formative feedback from summative assessments					

Please use the space below to write any comments you would like, related to the electronic test and subsequent feedback on PH1105: