

## **Self-evaluation of assessment performance can enhance students' perception of feedback on computer-generated tests.**

Georgina Fyfe, School of Biomedical Sciences, Curtin University,  
Jan Meyer, School of Anatomy and Human Biology, UWA  
Sue Fyfe, School of Public Health, Curtin University,  
Mel Ziman, School of Exercise, Biomedical & Health Sciences, ECU,  
Kathy Sanders, School of Anatomy and Human Biology, UWA,  
Julie Hill, School of Anatomy and Human Biology, UWA

Funded by the Carrick Institute\*, researchers from three Western Australian universities explored students' perception and use of feedback to on-line assessment tasks. Informed by surveys administered to several hundred students, the team developed and implemented feedback comments for all the possible responses to multichoice questions that tested various syllabus modules in Human Biology units. Data on students' use of the assessment tasks and associated feedback comments were compiled and analysed, and revealed differences between particular demographic cohorts of students that has implications for universities regarding the way in which on-line tests are accessed and administered.

The Carrick project team also developed an on-line survey that asked students to reflect upon their assessment performance in light of the automated feedback they had received. Although not all students engaged with this activity to the same degree, many perceived that an opportunity to evaluate themselves in a concise and convenient format was useful. This paper argues that providing students with an on-line format for self-evaluation, available on completion of a test once feedback on their performance has been obtained, prompts students to formulate their own strategies for improvement and thus gain more from the feedback and enhance their reflective practice skills.

### **Introduction**

#### *Importance of feedback to learning*

There is much evidence to support the belief that feedback is an important aspect of any cycle of learning. Within higher education in Australia, systematic reviews of students' course experience across the sector and at many levels flag feedback as an area where students perceive inconsistency or inadequacy in the feedback they receive. In spite of a perception from some academic staff that students do not read the feedback they are provided with and are only interested in their grade or mark, students consistently report in surveys that they do not receive sufficient feedback on their assessment tasks to help them understand their strengths and weaknesses in order to develop strategies for improvement (Krause, Hartley, James, & McInnis, 2005; Poulos & Mahony, 2007). Research on how students perceive feedback, and what aspects of feedback are most valued by students, is providing insight into how best to provide feedback to maximise its usefulness in evaluation and in transforming learning (Juwah et al., 2004; Lea & Street, 2000; Orsmond, Merry, & Reiling, 2005; Peat & Franklin, 2002).

#### *What constitutes effective feedback?*

For feedback to be most effective it should be appropriate and timely (Ramsden, 1992) which in the context of feedback on assessment tasks means within a timeframe that allows students to recall their responses and the understanding that informed their decisions. Shute defined feedback as "the review of information communicated to the learner to modify his or her thinking or behaviour for the purpose of improving learning" (Shute, 2008) and agreed that providing students with timely feedback is important. Feedback comes in many forms, from the very general to the very specific, but perhaps the most commonly used at an undergraduate level in the sciences is feedback that Shute describes as task-level feedback, which is less more specific to the particular assessment task and is delivered soon afterwards (Shute, 2008). In addition to timely feedback, students want a personalised experience from their feedback and they are not always satisfied with what they perceive as generic comments (Gleaves, Walker, & Grey, 2008;

Meyer et al., 2007). Feedback is most effective when students perceived the feedback comments show evidence of deep engagement of the assessor with students' work (Lizzio & Wilson, 2008). Whether or not students perceive the feedback they receive as being relevant and helpful will influence possible improvement on similar future tasks. Studies have shown that "elaborated feedback" that is specific to a student's particular response will enhance learning compared with other forms of feedback merely showing correct or incorrect response or allowing multiple tries until correct (Shute, 2008).

#### *How can reflection contribute to effective feedback?*

Reflection on past experiences is an important aspect to learning (Boud & Walker, 1998) and is being incorporated into more university courses in response to the importance of reflective practice as a generic graduate outcome to support lifelong learning. Reflective writing contributes to the assessment tasks of courses of study across many disciplines but can be difficult to assess objectively (Boyd, 2002; Davis, 2003; Donaghy & Morss, 2000). There is a tension between an assessment task requiring the learner to present themselves as competent and sure of their understanding whereas reflection by definition encourages articulation and analysis of self-doubt and the questioning of accepted theories and paradigms (Elliott, 2008). However, reflection can resolve the conflict necessary for learning and create new understanding and frameworks for thinking, and although this usually takes much time and effort but can happen very quickly under the right conditions (Launso, Rieper, & Rieper, 2007). Reflection by the student on their assessment performance following feedback on that assessment task can be quick and relatively easy, and students will comply when they perceive that such reflections provide the basis for dialogue with a tutor to help them improve for subsequent assessments (Fyfe, 2002). The use of computer technology to facilitate quick and timely reflection may increase compliance further, although there are reported differences in how students approach reflection on computer as opposed to handwriting in a journal (Gleaves et al., 2008).

#### *Providing effective feedback en masse*

Although the provision of effective feedback on assessment is generally agreed to be an important component of learning, it is a challenge for those responsible for teaching large classes (Orsmond, Merry & Reiling, 2005). Various strategies, including peer marking and the use of computer technology to help provide timely feedback to students following assessment tasks, have been tried to overcome impossible workload impositions on academic staff. However, this feedback will necessarily be somewhat generic regardless of the skill in constructing feedback comments that meet the criteria that constitute good feedback, but providing students with an opportunity to reflect on their performance can add a more personalised component. Post-assessment reflection has been perceived as useful by students when using a paper-based system (Fyfe, 2002) but when the assessment task is delivered by computer, this post-test reflection process is more readily delivered online.

## **Background**

This paper draws on work done during a project that explored issues relating to feedback in first-year undergraduate students studying units in human biology within three universities in Western Australia (a unit is defined as a portion of the semester load contributing to the whole course). In each of the three Human Biology units there was some aspect of formative or summative assessment mediated via computer-assisted testing. Our cross-institutional project team sought to understand what students perceived as feedback, what aspects of feedback they considered useful and helpful, and whether or not they were able to use the feedback from one context to improve their performance in similar or different subsequent assessments.

Funded by the Carrick Institute for Learning and Teaching in Higher Education (now the Australian Learning and Teaching Council), we approached our task in three stages. Firstly we surveyed the students to gain insight into their perceptions and their expectations of feedback in their course. The survey results provided a demographic snapshot of the students and guided our development of resources to provide online feedback to the computer-assisted assessment tasks in each course. Secondly we developed and tested feedback responses to multiple-choice and true/false questions within the online test banks before launching the feedback-enhanced online tests for each of the three institutions. Thirdly we added an online reflective survey for

students to complete following their online tests, and we gathered and analysed data on usage patterns, achievement and perceived usefulness and matched it against demographic data.

### What students (n=1099) told us in the initial survey

Firstly, demographic information was collected from the students including :

- a. age,
- b. gender,
- c. experience of study at university,
- d. whether studying full time or part time,
- e. language spoken at home,
- f. their anticipated mark in the unit.

Secondly, we asked them about their level of experience of online feedback and the extent to which they believed feedback was useful in different learning contexts (e.g. did it tell them how to improve, where they went wrong, how to answer similar questions and so on). We included some open-ended questions asking them to indicate what types of feedback they had previously found most and least helpful.

Thirdly we asked students to recall their past experience of their satisfaction with the amount and timeliness of the feedback they had received, and how they had used it.

Finally we asked students to tell us what type and amount of feedback they anticipated receiving for the human biology unit in which they were currently enrolled.

**Table 1. Demographic and other information gained from the initial student survey**

	University 1	University 2	University 3	Overall
n	421	564	114	1099
Age mean yrs SD (n)	18.25 SD 1.94 (418)	18.88 SD 2.87 (563)	23.97 SD 7.25 (108)	19.85 SD 3.68 (1089)
Age group expressed as a percentage of sample in each institution				
16-18	310 (74.2% )	356 (63.2%)	30 (27.8%)	696
19-21	90 (21.5%	154 (27.4%)	29 (26.8%)	273
>21	18 (4.3% )	53 (9.4%)	49 (45.4%	120
Gender expressed as a percentage of sample in each institution				
Male	167 (40%)	158 (28%)	48 (42%)	373 (34%)
Female	254 ( 60%)	406 (72%)	66 (58%)	726 (66%)

Full or Part-time enrolment broken down by age group			
Age group	Full time enrolment (%)	Part time enrolment (%)	
16-18	695 (65.2%)	1 (4.8%)	P<0.001
19-21	268 (25.1%)	4 (19%)	
>21	103 (9.7%)	16 (76.2%)	
Previous post-secondary experience broken down by age group			
Age group	Yes	No	P <0.001
16-18	42 (6.1%)	647 (93.9%)	
19-21	90 (34%)	175 (66%)	
>21	63 (55.3%)	51 (44.7%)	

The results from the survey were illuminating in many ways. They showed that there were differences in the way particular students viewed feedback. Not surprisingly, older students (>21) and students with prior experience of learning perceived feedback more positively and they said that they followed it up more effectively. They reported that they were able to use it in other

contexts to improve their performance in subsequent assessment tasks. The students also told us that they valued personalised feedback and wanted to know why they were right as well as why they were wrong in their answers. Armed with this information, the project team developed feedback comments for each possible response for multiple-choice and true/false questions in selected syllabus modules of the online tests at each of the universities in the study. Guided by the survey responses, the feedback comments were of medium length (1-2 sentences), provided meanings of keywords used in the question stem, explained why right answers were right as well as why wrong ones were wrong, affirmed correct responses and tried to keep a positive and personal voice. We aimed to provide an explanation of the underpinnings of the question rather than just providing content knowledge, while also supporting students to deal with criticism. Following trials the online feedback was implemented into the test banks for the three universities, mediated through either Web CT or Blackboard platforms. Both these platforms have tracking capabilities so we were able to follow patterns of use, such as whether or not the tests were completed in one sitting. These tests were dissimilar in some ways, including the number of tries a student may have at a particular test, security of test location and whether or not any marks were allocated to the outcome of the test. In the case where the tests did contribute to the semester mark for the unit these were of low risk, e.g. 6 or 7% of the total semester mark.

### **The optional online post-test survey**

We sought information from the students about how they perceived the usefulness of the feedback in the online tests with questions incorporated into an online post-test survey. This survey was optional for students and was done immediately, or some time following, the completion of their online test and after they had received their score and the associated feedback on their right and wrong answers. We asked them how they had looked at the feedback comments; either just scrolling through them, or only reading particular ones, or carefully reading all feedback comments. We asked them if they believed that the feedback comments helped them to understand their right answers, their wrong answers, or answers to other questions on the same topic. In a separate section we asked students whether the mark they received for their test was higher, lower or the same as they had expected, and whether or not the test was an accurate reflection of their knowledge. Another separate section asked students to report particular problem areas in either test technique or parts of the syllabus content, and to suggest strategies for improvement. These separate sections provided radio buttons with some options, gathered from previous paper-based reflective practice tools, and an open-ended comment section.

The responses from the post-test survey were gathered and analysed. Some of these results are summarised in the following sections.

### **Response to the post-test survey**

There was some difference in the pattern of response to the post-test survey between the different universities. At University 1, where the computer-based testing was used by students for review and could be tried a number of times, 283 students used the tests and of those 86 completed the post-test survey. At University 2, where the tests constituted part of the students' summative assessment, 279 students used the tests and of those 123 completed the post-test survey. At University 3, where the computer-based tests were used by students as formative assessment, 243 students used the tests and of those 106 completed the post-test survey. Therefore 315 students (39%) of a possible 815 students using the tests completed the post-test reflection survey. Generally, older students ( $\geq 21$  yrs) were more likely to complete the post-test survey and this was partly the reason for the higher percentage of respondent at University 3 where the cohort was generally older than in the other two universities. Students who had scored well in their computer test, students with previous study experience, and females were also more likely to complete the post-test survey.

#### *How did they report their use of the feedback comments?*

Almost half the students responding to the post-test survey stated that they read all the feedback comments provided for each of the responses in the questions but almost as many read only the answers to the ones they got wrong. Those reading all the comments were more likely to be from

University 1 where the assessment was formative while those reading only the comments on the ones they got wrong were more likely to be from University 2 where the assessment was summative. Eighty-nine percent of respondents agreed that the feedback helped them understand why they got answers wrong and over 90% believed they could apply what they had learned from the feedback to other Human Biology modules or units.

*Was their result higher or lower than they expected?*

A small majority of students (45%) answered in the post-test survey that the mark they received for their test was the same as they expected. Thirteen percent said that their mark was higher than expected while 42% said it was lower than they expected. A significantly higher than expected proportion of students speaking a language other than English at home reported scoring a lower mark than they expected, and was also seen in the total cohort of students at University 2 where the testing was summative.

*Did they believe that the test was an accurate assessment of their understanding?*

Predictably, a larger percentage of students scoring lower scores in their test thought that the test was not an accurate assessment of their understanding.

*What were some problems areas indicated by students?*

Ninety-one students indicated that they could not understand what some of the questions in the test were asking. Interestingly, the majority of these were younger students (<19yr) but there was no difference between native English-speakers and those who spoke a language other than English at home (LOTE). Ninety-two students indicated that they doubted their own knowledge and may have changed their test question answer or left it blank, and older students and native English-speakers were less likely to select this radio button. Thirty-seven students indicated that poor test technique contributed to their wrong answers, and the majority of these students were LOTE speakers. Seven students said they ran out of time to complete the test.

## **Discussion**

Although only 39% of students who had the opportunity to do the post-test survey responded, the level of engagement of those students was relatively high, judged for example by those who, when asked if they could think of any other examples or strategies to include in an open-ended comment section, politely answered “no”. Those students who reported that they received the marks that they expected from their assessment task were more likely to have scored well and to be older, whereas most students reporting that their marks were lower than expected were likely to be younger and LOTE speakers and their assessment was more likely to have been summative. Subsequently, LOTE students were also more likely to think that the assessment was not an accurate measurement of their understanding, and may suggest unrealistic expectations within this group of students. Students who reported that they had received the mark they had expected were more likely to deny any need for action, or to offer strategies for improvement. Perhaps these are strategic learners who aim for a particular grade and know what they need to do to achieve it and have no time, energy or motivation to aim for improvement.

There was an inversely-proportional pattern between test scores achieved and issues identified with test technique but not with syllabus knowledge problems, which were similar regardless of test score. This may indicate differences in identified locus of control, that is, if they blame external rather than personal factors for their results.

Those students who reported their test score as not an accurate assessment of their knowledge also engaged more with the survey, providing fewer negations to the radio buttons and offering more open-ended comments. This perhaps shows that these students welcomed an opportunity to vent in some form which in itself may be of benefit.

There seemed to be particular subgroups that did not take the opportunity to reflect after their test, including younger males and those who have scored less well in the test. Perhaps the post-test survey needs refining to either appeal to these subgroups more, or to select specifically for low-scoring students and approach their reflection in a different way.

## **Future directions**

There is an area of higher education teaching and learning where feedback is usually absent; feedback on students' final examination performance. Students often receive no information about their examination performance unless they specifically request an interview with the unit coordinator to review their paper. This is time-consuming for large classes and cannot be done until after the Boards of Examiners have ratified the examination grades, by which time students are focussing on the next semester's units. An online feedback site and associated reflection instrument available more immediately after the final examination may help students to think through their performance and learn from their experience. More work is needed to determine whether or not students benefit from opportunities for them to be reflective, and whether or not it enhances the gains seen from the provision of elaborate feedback. Longitudinal studies could also show whether or not students' reporting in CEQ and similar post-completion surveys of the adequacy and usefulness of the feedback they receive has changed.

## Conclusion

The results of our initial survey, our tracking of patterns of use of the computer-mediated tests across the three universities and analysis of the post-test survey show that students will engage with computer-mediated surveys that ask them to reflect on performance and to suggest reasons and strategies which will improve their performance on subsequent tests. Students want immediate, personalised, relevant feedback before and after assessments and they may need to receive it in different ways, driven by demographic and other factors. They would like to have a conversation about the feedback on their work prior to using it to effectively transform their learning. If feedback is delivered well it will enhance their learning, but online feedback related to computer-assisted assessment tasks cannot easily diagnose individual student problems, which may relate to understanding, question comprehension or test technique, leading to poorer outcomes. The project and study on which this paper is based showed that computer-mediated reflection can work and students will comply with a structure that is easy and quick to do, is non-judgemental and anonymous, but which provides a framework for them to think about at least some aspects of the feedback they receive after an assessment task, and to reflect on what it means in relation to their learning and study habits, and how they might make positive changes for future assessments.

## References:

- Boud, D., & Walker, D. (1998). Promoting reflection in professional courses: The challenge of context. *Studies in Higher Education*, 23(2), 191-206.
- Boyd, L. D. (2002). Reflections on Clinical Practice by First-Year Dental Students: A Qualitative Study. 66(6), 710-720.
- Davis, M. (2003). Barriers to reflective practice: The changing nature of higher education. *Active Learning in Higher Education*, 4(3), 243-255.
- Donaghy, M. E., & Morss, K. (2000). Guided reflection: A framework to facilitate and assess reflective practice within the discipline of physiotherapy. *Physiotherapy Theory and Practice; An International Journal of Physiotherapy*, 16, 3-14.
- Elliott, C. (2008). Emancipating Assessment: Assessment assumptions and critical alternatives in an experience-based programme. *Management Learning*, 39(3), 271-293.
- Fyfe, G. (2002). Building Reflective Practice into Student Learning. In A. Bunker & G. Swan (Eds.), *Focusing on the Student*. Western Australia: Edith Cowan University.
- Gleaves, A., Walker, C., & Grey, J. (2008). Using digital and paper diaries for assessment and learning purposes in higher education: a case of critical reflection or constrained compliance? *Assessment and Evaluation in Higher Education*, 33(3), 219-231.
- Juwah, C., Macfarlane-Dick, D., Matthew, B., Nicol, D., Ross, D., & Smith, B. (2004). *Enhancing Student Learning Through Effective Formative Feedback*. York: The Higher Education Academy.
- Krause, K.-L., Hartley, R., James, R., & McInnis, C. (2005). *The first year experience in Australian universities: findings from a decade of national studies* o. Document Number)
- Launso, L., Rieper, J., & Rieper, O. (2007). Evaluative feedback as a contribution to learning between groups of professionals. *Evaluation*, 13(3), 306-322.

- Lea, M., & Street, B. (2000). Student writing and staff feedback in higher education: an academic literacies approach. In M. Lea & B. Stierer (Eds.), *Student writing in higher education: new contexts*. Buckingham: Open University Press.
- Lizzio, A., & Wilson, K. (2008). Feedback on assessment; students' perceptions of quality and effectiveness. *Assessment and Evaluation in Higher Education*, 33(3), 263-275.
- Meyer, J., Fyfe, S., Georgina, F., Ziman, M., Plastow, K., Sanders, K., et al. (2007). *An articulated approach to the development and evaluation of automated feedback for online MCQ quizzes in Human Biology*. Paper presented at the National Uniserve Conference, Sydney.
- Orsmond, P., Merry, S., & Reiling, K. (2005). Biology Students' Utilization of Tutors' Formative Feedback: A Qualitative Interview Study. *Assessment and Evaluation in Higher Education*, 30(4), 369-386.
- Peat, M., & Franklin, S. (2002). Supporting Student Learning: The Use of Computer-Based Formative Assessment Modules. *British Journal of Educational Technology*, 33(5), 515-523.
- Poulos, A., & Mahony, M. J. (2007). Effectiveness of feedback: the students' perspective. *Assessment and Evaluation in Higher Education*, 33(2), 143-154.
- Ramsden, P. (1992). *Learning to teach in higher education*. London: Routledge.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153-189.

**Contact Details for Dr Georgina Fyfe:**

Senior Lecturer  
School of Biomedical Sciences,  
Curtin University,  
PO Box U1987  
Perth,  
Western Australia 6845  
Phone: +61-8-9266 7364  
Fax: +61-9266 2342  
Email: g.m.fyfe@curtin.edu.au

\*Support for this study was provided by The Carrick Institute for Learning and Teaching in Higher Education Ltd, an initiative of the Australian Government Department of Education, Science and Training. The views expressed in this paper do not necessarily reflect the views of The Carrick Institute for Learning and Teaching in Higher Education.