Educational research methods article

Evaluation tools for investigating the impact of assessment regimes on student learning

EVELYN BROWN[1], GRAHAM GIBBS AND CHRIS GLOVER[2]

1) *Faculty of Science, Open University, Walton Hall, Milton Keynes, MK7 6AA*
2) *Sheffield Hallam University*

Date received: 03/10/2003 *Date accepted: 05/11/03*

Abstract

The FAST project (Formative Assessment in Science Teaching http://www.open.ac.uk/science/fdtl/) is developing evaluation tools for science teachers to evaluate the impact of assessment regimes on their students' learning. This article describes the development and characteristics of these tools. The project will demonstrate how these tools can be used to diagnose where improvements could be made to assessment so as to support student learning better, and to measure improvements in student learning resulting from principled innovations in assessment. The tools currently include:

- *An 'Assessment Review Checklist'*
- *The AEQ (Assessment Experience Questionnaire) for diagnosis of the effects of assessment regimes*
- *A simple 'graph sketching' tool for students to report the pattern of their study effort across each week of their course, in relation to the schedule of assessment demands they experience.*
- *A coding system for categorising feedback that teachers write on assignments*
- *A questionnaire used to identify what feedback students receive, use and value*

This paper explains the use of these tools and provides a copy of each for use by lecturers.

Keywords: assessment, evaluation, questionnaires

The impact of assessment on students' learning

The starting point for the FAST project is the assumption that assessment has a profound impact on how much effort students put into learning, which topics they learn and which they overlook, and the quality of their engagement with the learning tasks they are set. The project also assumes that feedback to students on their understanding is a vitally important, though often under emphasised, component of teaching. A literature review on theories of assessment and of case studies of effective innovations in assessment, at school level and in higher education, led to the articulation of a set of eleven 'conditions under which assessment supports student learning' (Gibbs and

Simpson, in press) (see Table 1). These conditions have provided a conceptual framework for the review of science courses to see how well their assessment supports students in their studying. All the evaluation tools being developed relate to these conditions.

Table 1 *Assessment Conditions*

Student learning is best supported when the following conditions are met:
Assessed tasks capture sufficient student time and effort
These tasks distribute student effort evenly across topics & weeks
These tasks engage students in productive learning activity
Assessment communicates clear and high expectations to students
Sufficient feedback is provided, often enough & in enough detail
The feedback is provided quickly enough to be useful to students
Feedback focuses on learning rather than on marks or students
Feedback is linked to the purpose of the assignment and to criteria
Feedback is understandable to students, given their sophistication
Feedback is received by students and attended to
Feedback is acted upon by students to improve their work or their learning

This article is concerned with the range of evaluation tools being developed with which teachers can review their own assessment regimes and their impact (Gibbs, 2002, Gibbs et al 2003). The tools are designed for easy use by science teachers, rather than for educational researchers, and are provided in full in the Appendices for free use by teachers. Further evaluation tools will be added as they are developed.

Assessment Review Checklist

This checklist (see Appendix 1) is for teachers to review the assessment arrangements of a single course unit or module. It lists the 'eleven conditions' and simply asks about the extent to which the teacher believes that these conditions are met. The checklist also asks what evidence there is to back up these hunches. For example, little may be known about the distribution of student effort across weeks and across topics (re. Condition 2) which is why a simple tool has been developed to find out about students' distribution of effort (see Appendix 3). Evidence might take many forms. In relation to Condition 10 *('Feedback is received by students and attended to'*) you may already know, for example, that 25% of all assignments are not even picked up by your students from the department office. All such evidence will help to fill out an overall picture of how your students respond to your assessment. This assessment review checklist is a useful starting point for examining your own assessment and deciding where to look more closely and collect more evidence.

The 'Assessment Experience Questionnaire' (AEQ)

The 'Assessment Experience Questionnaire' was developed to provide quick and easy evidence from students about the extent to which students experience the 'eleven conditions' to be met. Questionnaire items were derived from unstructured interviews carried out at Open University science

summer schools in 2002, and from the literature. The AEQ contains six scales (each covering more than one 'condition'), and each scale contains six statements. Students are asked to indicate the extent to which they agree or disagree with each statement. Table 2 lists the six scales and illustrates one questionnaire item for each scale.

Table 2 *Assessment Experience Questionnaire scales and sample items*

Scale 1	Time demands and distribution of student effort
	e.g. *"In weeks when the assignments are due I put in many more hours"*
Scale 2	Assignments and learning
	e.g. *"Tackling the assignments you can get away with not understanding and still get high marks"*
Scale 3	Quantity and timing of feedback
	e.g. "There is hardly any feedback on my assignments when I get them back"
Scale 4	Quality of feedback
	e.g. "*The feedback shows me how to do better next time"*
Scale 5	Use of feedback
	e.g. "The feedback prompts me to go back over material covered earlier in the course"
Scale 6	The examination
	e.g. *"In the exam you can get away with not understanding and still get good marks"*

This questionnaire has been used with 1,050 students on seven science courses at the Open University and also on six science courses at Sheffield Hallam University. The scales are reasonably coherent, though they would benefit from some further development (Gibbs and Simpson, 2003).

The questionnaire has proved useful in distinguishing between courses (Gibbs et al, 2003). For example Table 3 highlights data demonstrating the differences between Open University courses in the way students use feedback. This has prompted follow-up studies to identify why students experience such wide differences in usefulness. This is very important to the Open University because a large proportion of course delivery costs are associated with providing feedback on assignments.

Table 3 *Comparison of Science courses within the Open University in terms of students' use of feedback*

AEQ items	**'Best' course**	**'Worst' course**
	% strongly agree	
The feedback helps me to understand things better	36%	6%
The feedback shows me how to do better next time	31%	4%
The feedback prompts me to go back over material covered earlier in the course	13%	1%
	% agree or strongly agree	
I do not use the feedback for revising	17%	44%

The AEQ is being used as a first stage diagnosis of what aspect of assessment to explore in more depth with the other more specific questionnaires and evaluation tools described below. A 'Special Interest Group (SIG) in Formative Assessment' (http://bio.ltsn.ac.uk/network/sigs/formassess/index.htm) has been established, supported financially by the FAST project and organisationally by both the Biosciences and Physical Sciences Subject Centres. This SIG is supporting science teachers to review assessment on their own courses and to evaluate the impact of innovations in assessment, using the AEQ.

Distribution of effort

You may not have much evidence to hand about the extent to which Condition 2 is met, concerning distribution of student effort across a course. Getting students to complete diaries in order to log how much time they spend on different parts of your course can be time consuming and also misleading. Usually as soon as students become aware of the way they use their time (by keeping such a log) they change their pattern of studying. However, it is possible to obtain a rough idea very quickly simply by asking students to draw a graph indicating the shape of their effort over successive weeks of the course. The horizontal axis of the graph should specify the weeks and also have marked on it the weeks in which any assignments, tests or exams are due. The vertical axis should indicate hours spent per week on the course (including all class contact time). You may even want to insert what topics were dealt with in each week, to help students to remember what they were doing that week. Student should be asked simply to draw the shape of the graph so as to give a rough idea of how (if at all) they change from week to week. It is also useful to ask them to write down why they spend their time in the way they do. Again these may not be the 'real' reasons but their responses will illuminate students' perceptions and the way they think about their effort in relation to assessment demands.

Such a 'distribution of effort' graph is unlikely to be a totally correct record – students may exaggerate how hard they work or may not be able to remember in detail. In fact the scale on the vertical axis may not be important. The actual number of hours worked per week may not be accurate but the relative effort in weeks with and without assignments is likely to reflect reality and students' explanations are likely to lend weight to this reality. Figures 1 and 2 below illustrate rather different patterns of student effort, in the form of histograms.

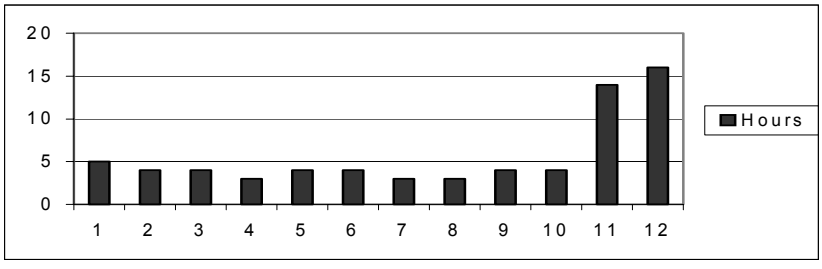

Figure 1 A student's estimate of his weekly study effort (in hours) on a course with no assignments and an exam in week 12. Total study effort = 68 hours

Instructions to students on how to draw up such a graph can be found in Appendix 3.

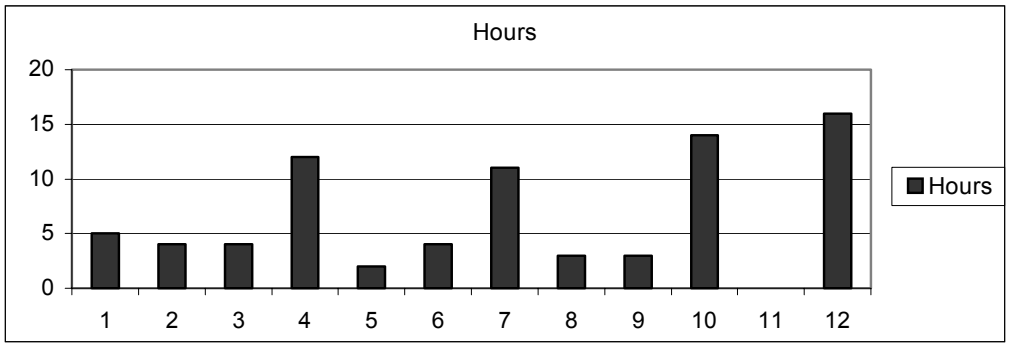

Figure 2*A student's estimate of her weekly study effort (in hours) on a course with three assignments due in weeks 4, 7 and 10 and an exam in week 12. Total study effort = 78 hours.*

Coding feedback to students

The feedback teachers write on students' work performs many functions, such as:

- identifying where errors have been made (or where work is correct);
- correcting these errors;
- explaining ideas or concepts the student has not fully understood;
- demonstrating techniques or procedures the student may not have used appropriately or correctly;
- engaging students in some thinking in relation to what they have written or presented;
- suggesting further study or reading
- explaining and justifying marks or grades
- suggesting how to approach subsequent assignments

… and so on.

Which of these functions is most important? Which types of feedback do students actually read carefully, think about and act upon? Would teachers be better concentrating on only some of these feedback functions? The starting point for answering such questions is to find out what types of feedback teachers actually give, and the second stage involves finding out from students which of these types they find most useful. At the Open University science tutors are coding each others' feedback. Appendix 4 contains categories of feedback and the codes used to record them, definitions and examples of each category, and a recording sheet to be completed for each assignment, to log the frequency of occurrence of each type of feedback.

Students will then be then interviewed, by a tutor other than their own, with their assignment feedback in front of them, to find out which feedback categories they are most likely to read, made sense of and use. The outcomes of this evaluation may include:

- briefing tutors about what kinds of feedback to concentrate on, given what students find useful;

- changing marking schemes so that feedback can concentrate on supporting student learning rather than on justifying grades;
- changing assignments so that feedback from one assignment is useful to students for the next assignment.

The value to students of different kinds of feedback

At Sheffield Hallam University (SHU) results from using the AEQ across a range of science courses suggested that students did not find written feedback to be as sufficient, prompt or useful as did students at the Open University. This is likely to be in large part because the OU allocates a much greater proportion of its course delivery costs to providing such feedback. Sheffield Hallam University science courses allocate far more resources than the Open University to face to face contact. The staff at SHU believed that students received a good deal of prompt and useful feedback orally, in a variety of contexts, but did not necessarily recognise or value this feedback (Gibbs et al, 2003). Sheffield Hallam University have developed a questionnaire to explore the way students perceive the existence and value of a wide range of types of feedback, much of it oral and some of it delivered to whole classes rather than to individuals. The questionnaire in Appendix 5 is being administered to students across a range of courses and asks them to reflect upon their overall experience of feedback of various types, and what kind of impact it might have. It does not distinguish between different courses but assumes that the entire department provides an assessment and feedback environment that students respond to in a global way. Certainly results from the use of the AEQ found that the institutional context contributed more to variance in students' perceptions than did the context provided by individual courses. It seems that institutions provide their own, unique, assessment environments.

Communicating author Evelyn Brown, Faculty of Science, Open university, Walton Hall, Milton Keynes, MK& 6AA. j.e.brown@open.ac.uk

Acknowledgements

The financial support of the Higher Education Funding Council for England, through the Fund for the Development of Teaching and Learning, is gratefully acknowledged.

References

Gibbs, G. (2002) *Evaluating the impact of formative assessment on student learning behaviour.* European Association for Research into Learning and Instruction Assessment Conference, Newcastle.

Gibbs, G., Simpson, C. & Macdonald, R. (2003) *Improving student learning through changing assessment – a conceptual and practical framework*. European Association for Research into Learning and Instruction, Padova.

Gibbs, G. & Simpson, C. (2003) *Measuring the response of students to assessment: the Assessment Experience Questionnaire*. 11th International Improving Student Learning Symposium, Hinkley.

Gibbs, G & Simpson, C. (in press) Does your assessment to support your students' learning? *Journal of Teaching and Learning in Higher Education*.

Links to these articles can also be found at http://www.open.ac.uk/science/fdtl/pub.htm