The Impact of Assessment on Student Learning: How Can the Research Literature Practically Help to Inform the Development of Departmental Assessment Strategies and Learner-Centred Assessment Practices?

Chris Rust

Active Learning in Higher Education 2002 3: 145
DOI: 10.1177/1469787402003002004

The online version of this article can be found at:
http://alh.sagepub.com/content/3/2/145
The impact of assessment on student learning

How can the research literature practically help to inform the development of departmental assessment strategies and learner-centred assessment practices?

CHRIS RUST

Oxford Brookes University, UK

ABSTRACT In the context of a worldwide paradigm shift towards student-centred outcomes-based approaches, and at a time when many UK departments are developing learning, teaching and assessment strategies, this article reviews what the research literature says about the impact of assessment on students’ learning. It then proceeds to translate that into practical suggestions for practice with the specific intention that this should help to inform departments in the development of appropriate assessment strategies and learner-centred assessment practices which meet the Quality Assurance Agency (QAA) general principles on assessment.

KEYWORDS: assessment, assessment strategies, learner-centred assessment, student learning

Introduction

Assessment is at the heart of the student experience. (Brown and Knight, 1994: 1)

From our students’ point of view, assessment always defines the actual curriculum. (Ramsden, 1992: 187)

Assessment defines what students regard as important, how they spend their time and how they come to see themselves as students and then as graduates. . . If you want to change student learning then change the methods of assessment. (Brown, 1997: 7)

Go to any textbook on teaching in higher education and you will find statements such as these, asserting that assessment methods and systems
influence student behaviour, and ultimately student learning. In my experience, most teaching staff, even though they may not have read any of these books, are nevertheless, at least partially aware of this relationship. They certainly know that if work does not have marks attached many students will either not do it at all or only do it in a perfunctory way. But to what extent have research literature findings about assessment actually informed and reformed our assessment practice? And conversely, to what extent are current assessment practices out of step with what the literature would suggest would be appropriate and effective in improving student learning?

In the UK, now is an ideal time to reflect on both questions. Institutions have been asked to produce learning and teaching strategies, and as a consequence many departments have also been charged with writing their own learning, teaching and assessment strategies. In the light of both, departments may be moved to reconsider and revise their assessment practices, and, in the process, perhaps ensure that they are soundly based on the wealth of research literature. Currently, departments are also having to consider how they will ensure that they meet the QAA codes of practice for assessment. The intention of this article is to review the research literature and its implications for practice to help inform the development of departmental assessment strategies.

Constructive alignment

Over the last decade in much of the English-speaking world (including the UK, the USA, Australia, New Zealand and South Africa) and in northern Europe there has been a paradigm shift, at least in the espoused rhetoric of higher education, from a focus on teaching to a focus on learning. Many mission statements can now be found which claim their institutions to be student-centred or learner-centred institutions. Beyond this rhetoric, possibly the most noticeable changes that can be seen, certainly in UK universities, are a greater emphasis on the development of skills, and in particular general, transferable, ‘life’ skills (and the notion of lifelong-learning), and the writing of course units and modules in terms of intended student-learning outcomes. (For the full implications of what a true paradigm shift might require see Barr and Tagg, 1995.) There does, however, appear to be a significant lag in the connection between changes in teaching methods and changes in assessment. Tom Angelo, describing the situation in the USA, comments:

We continue to assess student learning - and to graduate and certify students - much as we did in 1986, 1966, or 1946, without meaningful reference to what students should demonstrably know and be able to do. (Angelo, 1996: 3)
And in the UK, it is worth noting that in Quality Assurance Agency (QAA) subject reviews the aspect that most frequently loses a point is ‘Teaching, Learning and Assessment’, and the reason is almost always something to do with inconsistent assessment practices. Presumably because of these subject review findings, the Higher Education Funding Council for England’s (HEFCE) current invitation to bid for funds under phase four of the Fund for the Development of Teaching and Learning (FDTL) identifies assessment as the first of its generic priority areas, including matching assessment to learning outcomes (HEFCE, 2001). It would, therefore, seem fair to deduce that appropriate changes in assessment practice have been patchy at best. In all too many cases, although the course unit may now be written with four or five learning outcomes, the assessment task or tasks have remained the same and the linkage between the outcomes and the coursework essay, exam or whatever is tenuous at best, and almost always implicit. And may be further confused by the use of assessment criteria which may be additional to, or not directly relate to, the learning outcomes. As a result, students pursuing the courses often do not see any linkage between outcomes and assessment.

In an exam, where choices of which question to answer may be made by the student, it may well be possible to avoid one or more of the outcomes completely. In fact, exams offering choices of sections, or of questions within sections, may have exactly the opposite effect on the students’ approach to their learning than originally intended. A course might well include in its aims the encouragement of the students to read widely, but if the exam can be answered successfully by question spotting and selecting only a few topics to be revised thoroughly the result may be the exact opposite.

Of course, simply by having an exam question with the word ‘evaluate’ in the title, for example, does not necessarily mean that the students’ skills of evaluation are actually being assessed. They may be, but if the answer can be answered successfully simply by regurgitating the notes taken from the lecture you gave on the topic then all that may actually be being assessed are the students’ memorizing and essay-writing skills, along possibly with the skill of question spotting!

Assessment systems dominate what students are oriented towards in their learning. Even where lecturers say that they want students to be creative and thoughtful, students often recognise that what is really necessary, or at least what is sufficient, is to memorise. (Gibbs, 1992: 10)

When the final assessment decisions have to be aggregated into one mark or grade the linkage to the learning outcomes becomes even more untenable – an issue we will return to later.
So what does the literature suggest about how we might overcome these problems. Probably the most useful model can be found in the recent work of Biggs, and what he has termed ‘constructive alignment’.

The fundamental principle of constructive alignment is that a good teaching system aligns teaching method and assessment to the learning activities stated in the objectives so that all aspects of this system are in accord in supporting appropriate student learning. (Biggs, 1999: 11)

Essentially, in summary, Biggs’ model of ‘constructive alignment’ requires a shift in thinking about the process of course design, to the following three stage model:

1. Identify clear learning outcomes.
2. Design appropriate assessment tasks that will directly assess whether each of the learning outcomes has been met.
3. Design appropriate learning opportunities for the students to get them to a point where they can successfully undertake the assessment tasks.

Although the term ‘constructive alignment’ is not used, this kind of systematic thinking is exactly what the QAA are looking for when they refer to:

effective and appropriate measurement of the achievement by students of the intended learning outcomes. (QAA, General principle 6)

Departments mindful of the QAA requirements, and seeking to follow Biggs’ principles, would therefore be well advised to do two things:

1. To require all course modules or units to follow this design model, and to ensure that all assessment tasks, and assessment criteria, clearly and directly relate to the learning outcomes.
2. To audit all their modules’ or units’ learning outcomes and map them against the subject’s programme specifications, to ensure that all the programme specifications will have been assessed for any student successfully completing the course programme.

**Deep and surface approaches to learning**

Another theory which ought to influence our approach to course design and the assessment strategies we use, and that has been in the literature for considerably longer than Biggs’, is that concerning students’ approaches to learning (Dahlgren, 1984; Gibbs, 1992; Marton and Saljo, 1976, 1984; Ramsden, 1992). Essentially, this research says that students are capable of taking different approaches to their learning. The surface approach, defined as the student reduces what is to be learnt to the status of unconnected facts to be memorised...
[means that] the learning task [becomes] to reproduce the subject matter at a later date. Alternatively, a deep approach to their learning, defined as the student attempts to make sense of what is to be learnt, which consists of ideas and concepts (and) involves (the student in) thinking, seeking integration between components and between tasks, and 'playing' with ideas (Gibbs, 1992: 2). A key finding in this research, regarding the importance of assessment practice, is that most students can adopt either surface or deep approaches to their learning and one of the most important influences on which approach they take is the design of the course and the assessment strategies used.

Course characteristics associated with a surface approach are:

- A heavy workload
- Relatively high class contact hours
- An excessive amount of course material
- A lack of opportunity to pursue subjects in depth
- A lack of choice over subjects and a lack of choice over the method of study
- A threatening and anxiety provoking assessment system.

(Gibbs, 1992: 9)

Course characteristics which can foster a deep approach are:

- The engendering of intrinsic motivation in the students; students wanting and needing to know
- Learner activity
- Interaction with others
- A well structured knowledge base – i.e. where content is taught in integrated wholes and where knowledge is required to be related to other knowledge.

(Gibbs, 1992: 10-11)

So what are the practical implications of this research for assessment strategies? Well obviously to start with we need to ensure, as much as possible, that the workload expected of students is realistic, and that the assessment system is non-threatening and non-anxiety provoking. Some form of continuous assessment is almost certainly more likely to achieve the latter rather than a system in which the assessment ‘that counts’ all comes at the end. Within a continuous assessment system, we need to ensure that there is plenty of formative feedback at regular intervals, and all assessments need to have clear, assessment criteria known by the students before they undertake the work.

In the words of the QAA:

The principles, procedures and processes of all assessment should be explicit, valid and reliable. (QAA, General principle 2)
As stated above, students are also more likely to take a deep approach if they are intrinsically motivated. To be intrinsically motivated, they need to see the relevance and importance of what they are being required to do. Is it not interesting that one of the meanings of academic, as in the term ‘academic exercise’, is ‘of little or no real importance’. If students see an assessment task as essentially just a hoop to be jumped through, with no relevance or importance to them beyond passing the assessment, the research suggests that they are likely to take a surface approach. There is also evidence that they are more likely to justify plagiarism or other forms of cheating in achieving the task (Bannister and Ashworth, 1998; Carroll and Appleton, 2001; Cole and Kiss, 2000). (Students are also more likely to cheat if they believe the process is unreliable.) Assessment tasks are far more likely to appear relevant if they are ‘real-world’ tasks, i.e. the student can see why in the real-world someone might undertake such a task, and if there is the notion of a ‘real’ audience who might want to receive the resulting product (even if they know in reality that person will be their tutor). A report to a government select committee, a specialist magazine article or an information leaflet would all potentially meet these requirements, rather than a conventional essay question in which students may feel they have nothing new to contribute, and the task seen to have no purpose other than assessment. (For more detailed suggestions see Brown et al., 1994.)

Students are also more likely to be interested, and therefore motivated, if they have choice in their assessment task. If students are to be given choice in their assessment tasks without falling in to the problems of outcomes being missed or avoided (already discussed above) it is vital that the major assessment criterion must be that the learning outcomes are met. The use of learning contracts (see Knowles, 1986) can be a useful device to ensure this, sometimes coupled with the compilation of portfolios of evidence (see Redman, 1994).

Choice, coupled with the desire to create learner activity and interaction, would also suggest the use of assessment strategies such as independent project work, group-work and problem-based learning. Now although it is almost certainly true that more of these techniques are being used in universities currently, compared with 10 or 15 years ago, this is probably more so in new universities than old, and their usage still varies incredibly between different subject disciplines and between different institutions. If we are to take notice of the research evidence, surely we should all be asking whether we are using the whole range of these techniques as much as we might, and if not, why not?

Departments would therefore be well advised to audit the range of assessment strategies that they are using, their appropriateness, and to
examine critically the reasons why certain techniques may not be being used, and others perhaps over-used.

Widening participation and increasing retention

A further consideration for institutions and departments in the UK currently, as well as internationally, is the government’s concern to widen participation to more non-traditional groups of students, and to increase retention as part of its ‘social inclusivity’ agenda. This is clearly another area with implications for assessment policies and, here again, there is research data that can help. Although the importance of reducing the threat of assessment for all students has already been mentioned, it is possibly even more important for non-traditional students that the assessment system should be non-threatening.

Based on his research into student retention, Mantz Yorke proposes that to increase first-year retention the assessment at the end of the first semester should be primarily formative (part of an address to the ILT Symposium on Widening Participation and promoting student retention, in London on 27 September 2001). He further suggests that non-traditional students may be disproportionately lacking in self-belief, which obviously has implications for the need for good and sensitive feedback (see later). Other implications from his research are that drop-out can be reduced if longer assignments (such as dissertations) set sub-goals, and that programmes’ assessment strategies should allow for ‘slow learning’.

Any department that is serious about widening participation and/or increasing retention should surely give these ideas serious consideration in the development of an assessment strategy.

Preparing students for assessment

As has already been discussed above, if the assessment system is to be as unthreatening as possible, not to mention fair, the assessment process and criteria should be explicit and transparent to the students. In addition, one might be tempted to believe that having explicit criteria might also help the students to produce better work, but sadly research evidence does not bear this out (O’Donovan et al., 2001; Price and Rust, 1999). Giving students explicit assessment criteria alone is unlikely to result in them producing better work and, actually, we should not be surprised by this. If we acknowledge the research evidence that to learn and understand students need to actively engage in some way with what is being taught, why should this be any different if we want them to understand the assessment requirements?
So what might this active engagement with assessment criteria look like? Well, as a bare minimum it might involve time spent discussing the criteria with the students, but much more effective would be to involve the students in a marking exercise where they are actually required to apply the criteria to a number of pieces of work. A research study has shown (Price et al., 2001, submitted) that this can significantly improve the students’ performance when they subsequently undertake the task, and that this improvement continues in other similar pieces of work in which similar criteria are used at least a year later. And although this significant improvement can be achieved for as little as a 90-minute or 2-hour intervention, to my knowledge very few courses currently invest even this much time in such marking exercises.

Similar improvement in student performance has been demonstrated in courses in which it is possible to involve the students in peer-assessment, and peer marking using model answers (Forbes and Spence, 1991; Hughes, 1995).

Even more student engagement can be achieved through:

Involving students in the design of assessment, choice of assessment task and negotiation of criteria, for example through the use of contracts and self and peer assessment. (Gibbs, 1992: 17)

Any consideration of a department’s assessment strategies should surely therefore consider at least the inclusion of marking exercises in year one modules or units, as an important part of the students’ skills development, and then possibly moving on to involving the students in the design and choice of assessment tasks, and the negotiation of the assessment criteria.

**Marking and feedback**

Institutions should ensure that appropriate feedback is provided to students on assessed work in a way that promotes learning and facilitates improvement. (QAA, General principle 12)

As well as learning through undertaking the assessment task itself, assessment should also aid student learning through the subsequent feedback, provided by the tutor, on the strengths and weaknesses of how the task was done. If feedback is to be effective a review of the relevant literature would suggest the following, which readers might like to use as a checklist of their own practice (adapted from an unpublished workshop handout of David Jaques):

- Feedback should:
  - be prompt;
  - start off with a positive, encouraging comment;
RUST: THE IMPACT OF ASSESSMENT ON STUDENT LEARNING

• include a brief summary of your view of the assignment;
• relate specifically to the learning outcomes and assessment criteria;
• balance negative with positive comments;
• turn all criticism into positive suggestions;
• make general suggestions on how to go about the next assignment;
• ask questions which encourage reflection about the work;
• use informal, conversational language;
• explain all your comments;
• suggest follow-up work and references;
• suggest specific ways to improve the assignment;
• explain the mark or grade, and why it is not better (or worse);
• offer help with specific problems;
• offer the opportunity to discuss the assignment, and comments.

Although feedback which follows these guidelines is likely to be better than nothing, sadly research evidence would also suggest that just giving feedback to students without requiring them to actively engage with it is likely to have only limited effect. A study in the 1950s, replicated and developed further only last year (Fritz et al., 2000), seems to show that the emotional and psychological investment in producing a piece of work for assessment has a much stronger effect on the student than the relatively passive receipt of subsequent feedback. Consequently, if asked to repeat the task sometime later the student is likely to replicate what they did the first time, including the mistakes, despite the fact that these were pointed out, and supposedly ‘corrected’ in the feedback. The implication of this would seem to be that, in the same way as we need to use things such as marking exercises in order to be effective in getting the students to actively engage with marking criteria, we probably need to create feedback exercises that force students to actively engage with that feedback. It may be that only if we actually get the students to redraft the work in the light of the feedback or take it on to some further stage in some way will we have a significant effect on future performance.

Within any assessment strategy, as a minimum, a department should surely include a specific marking and feedback policy.

Pacing learning

If we know from the literature that students are likely to take a strategic approach to their studies, and in general only seriously engage with learning tasks if they are going to be assessed, with marks attached, then we should use this knowledge strategically ourselves in the way we design our courses. If the only assessment on a course unit or module is the final exam,
many students are likely to leave doing any serious work until a week or so before that exam. So, if we want to help students to pace their learning, and to engage seriously with the material from week one, we need to build in regular assessment tasks. In some accumulative, linear subjects, such as accountancy, this can be especially important. Intermediate tasks either need to carry marks, or be made course requirements (i.e. unless you carry out a minimum number of these tasks during the course you can’t sit the final exam). And if neither of these approaches seems appropriate, we need to be imaginative in finding other incentives. For example, an introductory chemistry course puts multiple-choice questions (MCQs) on the course’s website each week for the students to do in their own time. There are no marks for these tests, and no record is kept of who has taken them, but the students do the questions because they know from the beginning of the course that the end of module exam will include a section of MCQs and that half of these questions will have been selected from those questions used in the weekly tests (Rust, 2001). An additional benefit of an approach like this is that it does not involve any additional work for the tutor as there is no increased marking for such tasks. Peer assessment, and peer-marking using model answers are other techniques with this advantage (for examples see Rust, 2001).

Any departmental audit of assessment tasks and techniques should, therefore, not limit itself to a consideration of summative assessment, but should also consider the amount and nature of formative assessment techniques used, especially with a view to subsequently disseminating and spreading good practice, and creative ideas across the department.

**Marks and grades**

Ultimately, at the end of the whole assessment process on a programme, we give the students a degree classification; before that, on each module or unit, we gave them a numerical mark, or a letter grade (and in some cases both), which was probably derived from combining other numbers and/or grades together. I happen to think that there are a lot of good arguments against doing any of these things (Rust, 2000; Winter, 1983) but for this article I will confine my criticisms to the impact on student learning. So what’s wrong with classifications, marks and/or grades?

My first criticism would be that they don’t mean very much, and are not helpful in providing feedback. This is especially the case for numbers. What does 52 per cent, for example, actually mean? A number of students could all get 52 per cent but for very different reasons, and have considerably different strengths and weaknesses. Although this may be less true for a B grade, for example, it is still true to an extent. But despite their
problematic significance, numbers and grades can be invested with meaning that they do not have, especially by students who may virtually ignore the accompanying feedback. When this happens, students focus on the accumulated average mark or grade point average, or on getting a better mark or grade than others in the class, rather than what has been learnt, or what strengths and weaknesses have been demonstrated in the work.

Fundamentally, however, the main objection must be that with the move to a student-centred outcomes-based model of course design, aggregating students’ achievement makes little sense. If the student’s score is derived from assessing four or five outcomes (some of which may have been met well, some only partially, and maybe one not at all) to arrive at a number or grade, this does not help the student identify their strengths and weaknesses; neither does it help the course programme ensure that by the end of the programme the student will have met all the degree’s programme specifications.

So what is the alternative? Although it may sound radical, I believe the only logical conclusion of an outcomes-based model of course design is an assessment system that assesses explicitly against each individual learning outcome, and whether they have been not met/partially met/met/met well, rather than assessing the module as a whole. Such a system would mean that students received clear direct feedback on their achievement of individual outcomes, and it could explicitly ensure that all programme specifications are eventually met. The overarching checking that all outcomes are eventually met could be linked to progress files and help to promote serious engagement with them by the students and it could greatly reduce the need for resits (providing outcomes not fully met can be achieved elsewhere in subsequent course units or modules).

How this might work has been described at greater length in an earlier article (Rust, 2000) and I would urge any department that is seriously reappraising its assessment practices to at least consider whether it could make this more radical step.

Summary and conclusion

There is almost certainly a greater variety of assessment practices being used in our universities today, compared with 10 or 15 years ago, many of which could be justified with reference to what the research literature tells us about the impact of assessment on student learning. However, this variation is patchy, across disciplines and institutions, and not systematically implemented. There is also some research that has yet to make any real impression on practice at all. In considering the QAA general principles on assessment, now is the time for departments to develop clearly
articulated assessment strategies that are explicitly founded on the findings in the research literature. In these strategies, departments are recommended to:

- require all course modules or units to follow a constructively aligned design model, and to ensure that all assessment tasks, and assessment criteria, clearly and directly relate to the learning outcomes;
- audit all their modules’ or units’ learning outcomes and map them against the subject’s programme specifications, to ensure that all the programme specifications will have been assessed for any student successfully completing the course programme;
- ensure, as much as possible, that the workload expected of students is realistic, and that the assessment system is non-threatening and non-anxiety provoking;
- audit the range of assessment strategies that they are using, their appropriateness, and examine critically the reasons why certain techniques may not be being used, and others perhaps over-used;
- engender intrinsic motivation (and reduce plagiarism) by encouraging assessment tasks which resemble ‘real-world’ tasks and involve active engagement by the student, and by providing choice of tasks;
- allow for ‘slow’ learning;
- ensure there are sufficient formative assessment tasks, providing more feedback and helping to pace student learning;
- include explicit guidelines on giving effective feedback;
- prepare students for assessment through the use of marking exercises and self and peer-assessment;
- consider alternatives to the use of marks and grades.

References
RUST: THE IMPACT OF ASSESSMENT ON STUDENT LEARNING


HIGHER EDUCATION FUNDING COUNCIL FOR ENGLAND (2001) Invitation: FDTL Phase Four October 01/60, Bristol.


PRICE, M., RUST, C. & O'DONOVAN, B. (submitted) 'Improving Students' Learning by Developing Their Understanding of Assessment Criteria and Processes'.


Biographical note

CHRIS RUST is Head of the Oxford Centre for Staff and Learning Development and Deputy Director of the Human Resource Directorate at Oxford Brookes University. He was Course Leader for the university's initial training course for new teaching staff for six years, and with four colleagues he helps to provide both staff and educational
development support to the university's academic schools and support directorates. He has researched and published on a range of issues including: the experiences of new teachers in higher education, the positive effects of supplemental instruction, ways of diversifying assessment, and the effectiveness of workshops as a method of staff development. He is a Fellowship holder of the Staff and Educational Development Association (SEDA) and a member and accreditor of the Institute for Learning and Teaching in Higher Education (ILT).

Address: Oxford Centre for Staff and Learning Development, Oxford Brookes University, Gipsy Lane, Oxford OX3 0BP, UK. [email: cirust@brookes.ac.uk]