leeds metropolitan university

Using assessment to support student learning

By Graham Gibbs

Contents

1. Introduction 1

2. How assessment influences student learning 2

3. Pedagogic principles underlying the use of assessment to support learning 8

4. Assessment tactics that support student learning 20

5. Case studies of the use of assessment to support student learning within Leeds Metropolitan University 29

6. Evaluating the impact of assessment on student learning 34

References 42

Acknowledgements

Section 2 of this publication draws on Gibbs and Simpson (2004). The kind permission of the Journal *Learning and Teaching in Higher Education* and of its Editor, Dr Phil Gravestock, are gratefully acknowledged.

Introduction

Assessment makes more difference to the way that students spend their time, focus their effort, and perform, than any other aspect of the courses they study, including the teaching. If teachers want to make their course work better, then there is more leverage through changing aspects of the assessment than anywhere else, and it is often easier and cheaper to change assessment than to change anything else.

This manual is designed to support Scheme, Course, Award and Programme level leaders to introduce changes to assessment with the aim of improving student learning. It is not meant to be a list of 'tips', although there are plenty of practical ideas here in Section 4, and case studies from within Leeds Metropolitan University in Section 5. Rather it is intended to provide a way of thinking about how assessment works, and how students respond to it, so that teachers can make sense of what is currently happening on their own courses, and make their own context-relevant decisions about what they might do to improve things. It reviews the available empirical evidence in some detail, so that, as far as possible, these decisions can be made with confidence that they will produce improvements in student performance. It also provides three evaluation tools to help diagnose potential problems and measure any improvements brought about by changes teachers might make.

Good luck!

Graham Gibbs
Honorary Professor, University of Winchester

2. How assessment influences student learning

In the early 1970s, researchers on both sides of the Atlantic (Snyder, 1971; Miller & Parlett, 1974) were engaged in studies of student learning at two universities. What they found was that, unexpectedly, what influenced students most was not the teaching but the assessment. Students described all aspects of their study – what they attended to, how much work they did and how they went about their studying – as being completely dominated by the way they perceived the demands of the assessment system. Derek Rowntree stated that "if we wish to discover the truth about an educational system, we must first look to its assessment procedures" (Rowntree, 1987, p.1). The Snyder and Miller & Parlett studies went further and highlighted the way students respond to these assessment procedures. More recently, qualitative studies have emphasised the importance of understanding the way students respond to innovations in assessment (Sambell & McDowell, 1998).

Snyder's work gave birth to the notion of the 'hidden curriculum': different from the formal curriculum written down in course documentation, but the one students had to discover and pay attention to if they wanted to succeed:

"From the beginning I found the whole thing to be a kind of exercise in time budgeting ... . You had to filter out what was really important in each course ... you couldn't physically do it all. I found out that if you did a good job of filtering out what was important you could do well enough to do well in every course."

(Snyder, 1971, pp. 62-63)

Once students had worked out what this hidden curriculum consisted of, they could allocate their effort with great efficiency:

"I just don't bother doing the homework now. I approach the courses so I can get an 'A' in the easiest manner, and it's amazing how little work you have to do if you really don't like the course."

(Snyder, ibid, p. 50)

Miller & Parlett focused on the extent to which students were oriented to cues about what was rewarded in the assessment system. They described different kinds of students: the 'cue seekers', who went out of their way to get out of the lecturer what was going to come up in the exam and what their personal preferences were; the 'cue conscious', who heard and paid attention to tips given out by their lecturers about what was important; and the 'cue deaf', for whom any such guidance passed straight over their heads. This 'cue seeking' student describes exam question spotting:

"I am positive there is an examination game. You don't learn certain facts, for instance, you don't take the whole course, you go and look at the examination papers and you say 'looks as though there have been four questions on a certain theme this year, last year the professor said that the examination would be much the same as before', so you excise a good bit of the course immediately ..."

(Miller & Parlett, 1974, p. 60)

In contrast, these students were described as 'cue deaf':

"I don't choose questions for revision – I don't feel confident if I only restrict myself to certain topics."

"I will try to revise everything ..."

(Miller & Parlett, 1974, p. 63)

Miller & Parlett were able to predict with great accuracy which students would get good degree results:

"... people who were cue conscious tended to get upper seconds and those who were cue deaf got lower seconds."

(Miller & Parlett, 1974, p. 55)

Many students are perfectly capable of distinguishing between what assessment requires them to pay attention to and what results in worthwhile learning, as this postgraduate Oceanography student explained:

"If you are under a lot of pressure then you will just concentrate on passing the course. I know that from bitter experience. One subject I wasn't very good at I tried to understand the subject and I failed the exam. When I re-took the exam I just concentrated on passing the exam. I got 96% and the guy couldn't understand why I failed the first time. I told him this time I just concentrated on passing the exam rather than understanding the subject. I still don't understand the subject so it defeated the object, in a way."

(Gibbs, 1992, p. 101)

Whether or not what assessment is trying to assess is clearly specified in documentation, students work out for themselves what counts, or at least what they think counts, and orient their effort accordingly. They are strategic in their use of time and 'selectively negligent' in avoiding content that they believe is not likely to be assessed. It has been claimed that students have become more strategic with their use of time and energies since the studies conducted in 1970s and that students today are even more strongly influenced by the perceived demands of the assessment system in the way they negotiate their way through their studies (MacFarlane, 1992).

The effectiveness of coursework assignments

Students tend to gain higher marks from coursework assignments than they do from examinations. Chansarkar & Raut-Roy (1987) studied the effects of combinations of various forms of coursework with examinations. They found that all combinations of coursework of varying types with examinations produced better average marks than did examinations alone: up to 12% higher average marks. Gibbs & Lucas (1987) reported an analysis of marks on more than 1,700 modules at Oxford Polytechnic. Modules with 100% coursework had an average mark 3.5% higher than modules with 100% examinations, and there were three times as many failed students on modules where there were only examinations. There was a significant positive correlation between the proportion of coursework on a module and the average marks students achieved. Bridges et al (2002) studied the differences in coursework and exam marks in six subjects at four universities. They found coursework marks to be higher by one-third of a degree classification in English and History (similar to the Oxford Polytechnic finding) and higher by two-thirds of a degree classification in Biology, Business Studies, Computer Studies and Law.

Students also prefer coursework. Starr (1970) reported that 90% of students from four departments preferred half or more of their marks to come from coursework and 56% preferred all their marks to come from coursework. Students consider coursework to be fairer than exams, to measure a greater range of abilities than exams and to allow students to organise their own work patterns to a greater extent (Kniveton, 1996).

Higher average marks and student preference would not count for much if coursework were inherently less valid as an assessment – but it is not. First, examinations are very poor predictors of any subsequent performance, such as success at work. A review of 150 studies of the relationship between exam results and a wide range of

adult achievement found the relationship to be, at best, slight (Baird, 1985). For example, first degree results explain less than 10% of the variance in performance in graduate studies (Warren, 1971). Second, coursework marks are a better predictor of long-term learning of course content than are exams. Conway et al (1992) reported a study of the performance of Psychology students on a range of tests of their understanding and recall of content of a cognitive psychology course taken some years before. They found that student marks on coursework assignments undertaken up to 13 years previously correlated with these test scores, while students' original exam marks did not. Presumably the kind of learning that coursework involves has long-term consequences, while the kind of learning involved in revision for exams does not. Studies of surface and deep approaches to learning have shown similar results: that any positive impact on test results of students taking a surface approach in preparation for the test are very short-lasting (Marton & Wenestam, 1978).

Third, in experimental studies in which students have studied either exam-based or assignment-based courses, the quality of their learning has been shown to be higher in the assignment-based courses. For example Tynjala (1998) compared two student groups: the first group studied via conventional lectures, a text-book and an exam. The second group studied via assignments based on the text-book, discussion with other students about these assignments, and a coursework essay marked by the teacher. This second group then also took the exam so as to enable a comparison with the first group, even though the students had not studied for the exam. The second group was found to place more emphasis on thinking and had developed more sophisticated conceptions of learning (see Säljö, 1982). In their exam answers they revealed more comparisons, more evaluations and more sophisticated structures to their writing (Biggs & Collis, 1982). These results (achieved with less teaching) were interpreted in terms of the assessment requirements for the second group being more 'constructivist' – they helped students to construct meaning from the material they were studying.

Assessment and student workload

It is a common observation of higher education teachers that if coursework is taken away from a module because of resource constraints, then students simply do not do the associated studying; for example students will rarely write unassessed essays. It is argued that you have to assess everything in order to capture students' time and energy. There are several problems with this rationale, both logistical and pedagogic.

It may not be possible to set enough assignments to actually capture much student time. A study of the nature of assessment systems at the level of whole degree programmes (Gibbs & Dunbar-Goddet, 2007) found, first, that programmes tended to have either high levels of assessed work for marks (and some degree programmes mark well over 100 separate assignments and examinations over three years) and very little assessed work for feedback only (as few as two assignments in three years), while other programmes had very low levels of assessed work for marks (as few as 11 in three years) and very high levels of assessed work for feedback only (up to 134 in three years) and also a great deal of oral feedback. A few programmes had neither high levels of summative assessment (for marks) nor formative-only assessment (for feedback), but no programmes had high levels of both. This seems to be simply economic reality – you can afford to mark often or give feedback often but you cannot afford to do both often. Second, it was found that programmes with low levels of marked work but high levels of feedback (with no marks attached) had students who worked harder and distributed their effort evenly across weeks and across topics on courses. In contrast, where there were only one or two marked assignments per course unit, these were all students spent their time on, largely ignoring all topics other than those addressed in the assignments and spending little time on any course that did not have an assignment due in that week (Gibbs & Dunbar-Goddet, 2007). The only way that it is possible to afford enough marking to capture students' time and distribute it evenly across weeks is through mechanised assessment such as computer-marked multiple choice question tests. The problem then is that mechanised assessment very often has substantial negative consequences for the quality of learning that students engage in (see below).

However, coursework does not have to be marked to generate the necessary learning. Forbes & Spence (1991) reported a study of assessment on an Engineering course at Strathclyde University. When lecturers stopped marking weekly problem sheets because they were simply too busy, students did indeed stop tackling the problems, and their exam marks went down as a consequence. But when lecturers introduced periodic peer-assessment of the problem sheets – as a course requirement but without the marks contributing – students' exam marks increased dramatically to a level well above that achieved previously when lecturers did the marking. What achieved the learning was the quality of student engagement in learning tasks, not teachers doing lots of marking. The trick when designing assessment régimes is to generate engagement with learning tasks without generating piles of marking.

The effectiveness of feedback

Summaries of what makes most difference to student achievement, involving reviews of many thousands of studies spanning decades of research, show clearly that the most powerful single influence is feedback (Hattie, 1987; Hattie & Timperley, 2007). Similarly, a comprehensive review of formative assessment (Black & Wiliam, 1998) emphasised the extraordinarily large and consistent positive effects that feedback has on learning, compared with other things teachers might try and do. There have been many attempts both to understand the nature of this impact and to harness its power through innovation, especially in schools, in response to this incontrovertible evidence.

In higher education, feedback to individual students in class must have declined significantly as class sizes have increased, though we have no evidence about this. Writing comments on assignments, however, remains a major component of teachers' workload in higher education. As resources per student have declined there have been some economies of scale in teaching (simply by packing more students into classrooms), but economies of scale are difficult to achieve for assessment: most assessment costs go up in direct proportion to the number of students. As a result, assessment costs can overtake teaching costs and teachers can find themselves spending much of their time marking. Is all this effort worthwhile?

In the Course Experience Questionnaire (Ramsden, 1991), used extensively in Australia and elsewhere to evaluate the quality of courses, the questionnaire item that most clearly distinguishes the best and worst courses is: "Teaching staff here normally give helpful feedback on how you are going" (Ramsden, 1992, p.107). Similarly, the variable on the National Student Survey that best distinguishes between universities and contributes most to determining their overall ranking concerns feedback.

This does not mean that higher education teachers in fact give helpful feedback. It means that the extent of helpfulness of the feedback they give makes more difference to students than anything else they do.

How well does feedback actually work? Maclellen (2001) surveyed 130 students and 80 lecturers at the University of Strathclyde about their perceptions concerning assessment. Among the 40 questions asked, four were about feedback and these revealed wide discrepancies between students and lecturers. While most teachers responded that feedback is frequently helpful in detail, helps students to understand and helps learning, most students responded that feedback was only sometimes helpful in these ways. 30% of students reported that feedback never helps them to understand. While 63% of lecturers responded that feedback frequently prompts discussion with a tutor, only 2% of students responded in the same way and 50% responded that feedback never prompted discussion. In another study, only 1% of students reported that reading feedback prompted them to go back to the subject matter and spend any more time on it (Gibbs et al, 2003).

There may be a problem here with the quantity and quality of feedback such that it is not actually helpful to students – after all, teachers are under enormous time pressure and it is difficult to provide comprehensive and useful feedback under such circumstances. But there are other problems. Studies of what students do with feedback makes for depressing reading. Feedback is often not read at all (Hounsell, 1987) or not understood (Lea & Street, 1998). Wojtas (1998) reported:

"Some students threw away the feedback if they disliked the grade, while others seemed concerned only with the final result and did not collect the marked work."

There is also a problem associated with both marks and feedback being provided. A grade is likely to be perceived by the student

as indicating their personal ability or worth as it is usually 'norm-referenced': it tells you, primarily, where you stand in relation to others. A poor grade may damage a student's 'self-efficacy' or sense of ability to be successful. Yorke (2001) elaborates on the ways in which formative assessment can affect student retention and emphasises its role in 'academic integration' (Tinto, 1993). In contrast, feedback on its own is more likely to be perceived as a comment on what has been learnt. In the absence of marks it has been reported that students read feedback much more carefully (Black & Wiliam, 1998) and use it to guide their learning. In the light of this (school-based) research evidence, some schools have adopted policies that all assignments should only have feedback and that no marks should be provided.

This is not a pretty picture. Assessment sometimes appears to be, at one and the same time, enormously expensive, disliked by both students and teachers, and largely ineffective in supporting learning, given the way it is often conducted. In the light of these problems, what follows is an attempt to justify a set of principles to guide assessment practice. The evidence is rarely conclusive enough to argue that if your assessment follows these principles then learning will inevitably be more effective, but they are a plausible set of guidelines given what we know.

This is not the first attempt to identify such principles but is the first attempt in the context of higher education. School-based research has identified lists of effects of formative assessment. Gagne (1977) argued on the basis of evidence that feedback:

1. Reactivates or consolidates prerequisite skills or knowledge prior to introducing the new material
2. Focuses attention on important aspects of the subject
3. Encourages active learning strategies
4. Gives students opportunities to practise skills and consolidate learning
5. Provides knowledge of results and corrective feedback
6. Helps students to monitor their own progress and develop skills of self-evaluation
7. Guides the choice of further instructional or learning activities to increase mastery
8. Helps students to feel a sense of accomplishment.

(derived from Crooks, 1988)

The principles outlined in this manual refer to two relatively distinct categories of influence:

- the influence of the design of assessment systems and assignments on how much students study, what they study and on the quality of their engagement
- the influence of feedback on learning.

3. Pedagogic principles underlying the use of assessment to support learning

3.1. There should be sufficient assessed tasks to capture sufficient student study time

This issue concerns how much time and effort students allocate – the 'time on task' principle (Chickering & Gamson, 1987) that if students don't spend enough time on something they won't learn it. Berliner (1984), summarising research in the 'time on task' principle, concluded that there was strong empirical evidence of a direct relationship between time allocation by courses, student time management and actual student time on task on the one hand, and student achievement on the other. The relationship between effort and marks is not always straightforward. Kember et al (1996) found that students' perceptions of their effort depended on their motivation more than on the number of hours they actually allocated, and that it was possible for students to put in many hours unproductively, especially if they adopted a surface approach to their studies. Some kinds of assessment can generate long hours of ineffective memorisation.

Courses in UK higher education are designed to involve a specified number of learning hours related to the number of credits for the course. Students are normally expected to spend between about one and four hours out of class for each hour in class (depending largely on the discipline involved). Innis (1996) found students at Leeds Metropolitan University spend between 1.4 and 3.0 hours out of class for each hour in class. How much of this 'out of class' time is actually allocated to studying may be determined largely by assessment demands. In the USA, higher education students on average spend less than half as many hours out of class for each hour in class as teachers expect: between 0.3 and 1.0 hours out of class when teachers, on average, expect 2.1 hours out of class for each hour in class (Moffat, 1989; Hutchings et al, 1991; Gardiner, 1997; Brittingham, 1998). The emphasis in the USA on attempts to improve student performance through assessment is on 'classroom assessment' – activities undertaken in class to test students – and then on using this assessment information to guide both students and teaching (Angelo & Cross, 1993). This focus on the classroom could be interpreted as a recognition of the failure to generate much out of class learning through the type of assessment teachers use. Diary studies (such as that by Innis, ibid) show how students in the UK allocate their time largely to assessed tasks and that this becomes a more narrow focus over time as they become more experienced, with students allocating as little as 5% of their time to unassessed study tasks by year 3. Subject areas with less frequent assessed tasks (e.g. text-based subjects) have students who study fewer hours (Vos, 1991). Science and technology subjects that generate greater total study effort tend to have more frequent (though smaller) assessed tasks, such as problem sheets and laboratory reports.

In mainland Europe only 5%–25% of full-time students (varying by country) study fewer than 20 hours per week, while in the UK 34% study 20 hours per week or less (Eurostudent, 2007). About 20 hours per week is the norm in three subject areas in post-1992 universities (Higher Education Policy Institute, 2007). 20 hours per week is part-time studying, so a significant minority of UK students are enrolled full-time but studying part-time, with their university receiving funding for full-time students. For these to be average hours, some institutions are likely to be doing worse than this. This pattern of low study hours has been found to be associated with modular courses with a large volume of summative assessment (Gibbs & Dunbar-Goddet, 2007). It seems clear that high volumes of marking may not achieve high volumes of student effort.

Students who put in fewer hours may be doing so because they are undertaking paid work in parallel with their studies. In the USA students normally 'work their way through college' by taking individual course

units as and when they have the time, given their need to work to support themselves. In the UK students seem to think it is acceptable to undertake paid work for many hours a week and still complete in three years. Studies of the impact of students undertaking paid employment in parallel with full-time study show that such students study fewer hours (Curtis & Shami, 2002) and perform significantly less well (Paton-Salzberg & Lindsay, 1993). Studies show that up to three-quarters of full-time students work during term-time and are likely to allocate their reduced study hours especially strategically in relation to assessment requirements. These studies show reduced reading and other out of class study activity.

Assignments are not the only way to capture student time and effort through assessment. The conventional way to do this is by having unpredictable sampling of course content in unseen examinations, so that for a student to ignore anything is a high risk activity. The quality, quantity and distribution of the study effort captured in this way is somewhat unpredictable and probably varies with student perceptions of the likely exam demands and the risks associated with choosing not to study some of the topics. However, the same rationale as is used in sampling the curriculum through exam questions can be applied to coursework. Students can be expected to undertake coursework on every topic and this coursework can be sampled for marking – perhaps two chosen randomly to be marked from eight that students have undertaken..

Time and effort can also be captured through social pressure, for example through the potential embarrassment of the poor quality of your work being seen by colleagues, as when a seminar presentation is assessed, or when a laboratory report is written and displayed publicly in the form of a poster. The potential censure from colleagues if a student were to fail to complete his or her component of an assessed group assignment can also generate effort. 'Bunking off' or 'social loafing' during group work (Latane et al, 1979) is less common when group size is small and students cannot hide or be anonymous (Kerr & Bruun, 1983; Williams et al, 1981).

3.2. Assessment demands should be designed so as to orient students to distribute appropriate amounts of time and effort across all the important aspects of the course

This principle concerns what student effort is spent on. Students usually distribute their time unevenly across courses, often focusing on topics associated with assessment and nothing else. If they drew a graph of weekly study effort for all the weeks of an individual course involving a sequence of assignments, it might look more like the Alps than Holland. Exams can have the effect of concentrating study into a short intense period at the end of the course with, for example, little study of lecture notes until many weeks after the lecture. Frequent assignments (such as short problem sheets) or tests (such as computer-based assessment) can distribute student effort across the course, often on a weekly basis, while infrequent assignments (such as extended essays) may result in intensive studying for a week or two immediately prior to the assignment deadline, while topics not covered by the assignment can be largely ignored. We know very little about the distribution of student effort, and higher education teachers tend to know little about what their students do with their time and when. Section 6 contains a simple prompt for students to tell you how they distribute their time.

3.3. Tackling the assessed task engages students in productive learning activity of an appropriate kind

This issue concerns the kinds of study and learning activity involved in tackling the assignment or in preparing for tests. Some assessment generates unhelpful and inappropriate learning activity, even if it produces reliable marks. Multiple choice question (MCQ) tests commonly mis-orient students to adopt a surface approach involving only memorising (Scouler & Prosser, 1994; Tang, 1994; Scouler, 1998), as can exams. Snyder (1971) described how students encouraged to be creative at Massachusetts Institute of Technology abandoned any such aspiration on discovering that most of the marks were derived from rote memorisation of material for multiple choice tests. It is important to recognise that it is students' perceptions of the demands of tests that determine how they go about their studying, rather than what the teacher who designed the test intended. I have myself set MCQ questions intended to test quite sophisticated analysis, and to encourage students to practise this kind of analysis in preparation, only for students to assume that as it is an MCQ test it is bound to require only memorisation.

It is not inevitable, however, that MCQ tests or examinations lead to worse learning. Macdonald (2002) has reported that at least some students adopt a deep approach to examination revision and learn effectively as a result of the integration of material that their revision involved, and others have reported a similar phenomenon (Entwistle & Entwistle, 2004). A recent study comparing institutional approaches to assessment (Gibbs & Dunbar-Goddet, 2007) suggested that it was the amount of time students had between teaching stopping and the exam that determined the quality of their engagement during revision. If this period is too short, students are very likely only to engage in memorisation. If it is much longer, students can experience revision as a highly engaging

integrative experience, even describing it as the most valuable part of the course. The following two quotations from students contrast these experiences of revision:

"You just go back to the lecture notes, and then just try and remember as much as I can ... just looking at lecture notes and trying to remember stuff."

"...you actually had time to – I think there was also an element of being able to go back and consolidate everything by revising for the exam."

(Gibbs et al, ibid)

Many assignments simply fail to engage students with appropriate types of learning activity. Submitting a laboratory report of a teacher-designed procedure is unlikely to help students to learn how to design experiments. Probably the only way to learn how to solve problems is to solve lots of problems. Probably the only way to gain facility with the discourse of a discipline is to undertake plenty of practice in using that discourse, for example through writing. Assignments are the main way in which such practice is generated.

Some assignments create appropriate learning activity as a by-product. For example, setting essays can generate 'reading around' and can support the working up of coherent arguments in a way that simply asking students to read what is on the reading list does not. If you were to take the essay away, the appropriate form of reading and thinking would not occur even in the unlikely event of a similar volume of reading of similar material taking place. The product, the essay, and the marks associated with it may be less important to the learning than the framework the assignment provides for the learning activities of 'reading around' and 'constructing arguments'. Similarly, with laboratory reports or design briefs, the product may be less important than details of the studying required to fulfill the assignment requirements. Group projects can engage students in much discussion and can confront individuals with alternative views and different standards of work. The quality of the group product (such as a report) that is marked may be less important than the qualities of the learning process that created it.

Students can tackle assignments that are intended as learning activities so as to maximise the marks they obtain rather than maximising the learning achieved from engaging with the assignment. This may involve 'faking good' and pretending to be competent or knowledgeable, deliberately covering up misunderstanding and ignorance, telling teachers what they think they want to hear rather than what they as students want to say, and so on. To some extent this is a consequence of the student's orientation (Beaty et al, 1997), but assessment tasks, marking régimes and the way in which feedback functions can override such individual orientations and even encourage student behaviour that reduces learning. In the example below an intrinsically-oriented student describes, in a learning log, the way he used to tackle assignments in Engineering so as to obtain marks at the expense of learning:

"The average lecturer likes to see the right result squared in red at the bottom of the test sheet, if possible with as few lines of calculation as possible – above all else don't put any comments. He hates that. He thinks that you are trying to fill the page with words to make the work look bigger. Don't leave your mistakes, either, even corrected. If you've done it wrong, bin the lot. He likes to believe that you've found the right solution at the first time. If you're still making mistakes, that means you didn't study enough. There's no way you can re-do an exercise a few months after because you've only got the plain results without comments. If you have a go, you may well make the same mistakes you've done before because you've got no record of your previous errors."

(Gibbs, 1992)

3.4 Assessment should communicate clear and high standards

This issue was highlighted in the 'seven principles of good practice in undergraduate education' (Chickering & Gamson, 1987; 1991): "good practice communicates clear and high expectations". Assignments need to appear challenging, but possible provided that you work hard enough, rather than easy. It also needs to be clear what kind of challenge has been set. The extent to which students experience "clear goals and standards" (as measured by the Course Experience Questionnaire, Ramsden, 1991) is closely associated with the extent to which students take a surface approach or a deep approach to their studies (Säljö, 1981). Students who don't understand what they are supposed to be doing tend to revert to a surface approach and simply reproduce material, in the absence of any clearer imperatives.

Much of the effort to communicate clear goals has been expended on specifying assessment criteria, and students do need to understand what counts as good or bad if they are to orient their effort appropriately and put in enough effort. However, the words used in articulating criteria are seldom meaningful to students and it is difficult for a student to tell what standard is expected or would be considered inadequate. For example Penny & Grover (1996) have reported the extent to which students misunderstood the criteria used to assess their final-year research project. The students expected criteria to be concerned with low-level goals such as style and presentation, while their teachers emphasised high-level goals such as theoretical and conceptual understanding. A comparison of degree programmes has shown that it is not when criteria are spelled out in detail for each assignment that students are clear about goals and standards, but when they get plenty of written and oral feedback (Gibbs & Dunbar-Goddet, 2007).

Much of the literature on the use of self- and peer-assessment is about the reliability of student marking, on the assumption that students are acting as cheap substitutes for teachers and that this is an acceptable practice provided that they can generate usable marks. But students do not need more marks: they need more feedback. The real value of self-assessment lies in students internalising the standards that are expected so that they can supervise themselves in relation to these standards and improve the quality of their own assignments prior to submitting them. This idea is at the heart of extensive work at the University of Strathclyde to support student learning through assessment (Nicol, 2006) and the focus of the Assessment Standards Knowledge Exchange, a Centre for Excellence in Teaching and Learning in Higher Education (**http://www.brookes.ac.uk/aske/!**). It seems clear from a range of studies (e.g. O'Donovan et al, 2008) that students do not come to understand much about standards by reading lists of criteria. Rather they need to see exemplars of work of different standards, to make their own judgments about the qualities of these exemplars (e.g. Orsmond et al, 2002), and to gradually calibrate their own judgments so that they are in line with the judgments their teacher would make. This is rather like inexperienced researchers learning about the standards they should be aspiring to in their research articles through acting as a reviewer of others' articles.

The influence of feedback on learning

"Knowing what you know and don't know focuses learning. Students need appropriate feedback on performance to benefit from courses. In getting started, students need help in assessing existing knowledge and competence. In classes, students need frequent opportunities to perform and receive suggestions for improvement. At various points during college, and at the end, students need chances to reflect on what they have learnt, what they still have to learn, and how to assess themselves."

(Chickering & Gamson, 1987)

Conventionally, feedback is conceptualised as an issue of 'correction of errors' (Bruner, 1974) or 'knowledge of results'. If a student is informed that she is accurate then she will learn. The following principles are concerned, instead, with how the provision of feedback affects student learning behaviour: with how feedback results in students taking action that involves, or does not involve, further learning.

3.5 Sufficient feedback needs to be provided, both often enough and in enough detail

This issue concerns what is conventionally defined as formative assessment: the impact on learning of feedback on progress, usually provided after a 'performance' on an assignment. The volume and thoroughness of feedback varies enormously between courses – far more than the variation in quantity or quality of teaching (Gibbs & Dunbar-Goddet, in press).

This feedback may need to be quite regular, and on relatively small chunks of course content, to be useful. One piece of detailed feedback on an extended essay or design task after ten weeks of study is unlikely to support learning across a whole course very well. There has been very widespread adoption of computer-based testing to provide at least some feedback on progress, and in some assessment software it is possible to provide 'remedial feedback' when incorrect answers are selected. Cook (2001) has reported that students' final exam marks were closely related to the number (and therefore frequency) of computer-marked assignments they had tackled. The frequency and speed of response of such feedback, which it is possible to provide reasonably economically, may compensate for its relatively poor quality and lack of individualisation.

Feedback has to be quite specific to be useful. The Open University trains its 7,500 part-time tutors to give quite detailed and extensive feedback. Cole et al (1986) list the characteristics of effective feedback in distance learning, and Roberts (1996) found that students' preferences for feedback closely match this list. The specific forms of feedback that are effective vary from discipline to discipline. Evidence about the most effective forms of feedback in language learning, for example, is summarised in Hyland (2001). In both Psychology (Stephenson et al, 1996) and Mathematics (Rice et al, 1994) students have

been reported as wanting specific, detailed facilitative feedback. Greer (2001) reports a study that illuminates exactly what kind of impact feedback was achieving on the learning of Accountancy.

Much of the feedback to students provided in the rest of the higher education sector would be picked up by the Open University's Staff Tutors (who monitor tutors' marking) as being totally inadequate and would lead to quality assurance and staff development interventions. If this seems excessively interventionist it should be noted that the Open University has been ranked top in the National Student Survey results year after year, and primarily on the basis of its extraordinarily high student ratings for assessment and feedback.

3.6 Feedback should focus on students' performance, on their learning and on actions under the students' control, rather than on the students themselves and on their characteristics

Literature on formative assessment intended to support learning distinguishes between feedback that tells students they are hopeless, or among the bottom 10% of students (a grade D, for example), and feedback that tells students exactly where they have gone wrong and what they can do about it. Grades without feedback may be particularly damaging. A focus of critical feedback on personal characteristics can be demotivating and can negatively affect students' 'self-efficacy' or sense of competence. This is important because self-efficacy is strongly related to effort and persistence with tasks (Schunk, 1984; 1985), predicts academic achievement well and is associated with adopting a deep approach to learning (Thomas et al, 1987). In contrast, feedback concerning content provides students with options for action and is less closely associated with their egos – it is about what they can do next rather than about themselves. Wootton (2002) has written passionately about the negative impact of assessment on 'at risk' students and asks whether the system exists "to encourage learning or to measure failure".

3.7 Feedback should be timely: received by students while it still matters to them and in time for them to pay attention to further learning or receive further assistance

This is another of the 'seven principles of good practice in undergraduate education' (Chickering & Gamson, 1987). It is based on a range of studies of the timing of feedback (for summaries, see Dunkin, 1986; McKeachie et al, 1986). A teaching method which places great emphasis on immediate feedback at each stage of a student's progress through course units, the Personalised System of Instruction (PSI), has been demonstrated in many studies to improve student performance (Kulik et al, 1980).

If students do not receive feedback fast enough then they will have moved on to new content and the feedback will be irrelevant to their ongoing studies and unlikely to result in additional appropriate learning activity, directed by the feedback. Owing to resource pressures and quality assurance worries about grades, feedback is today being provided more slowly, and as courses in the UK are quite short, this may mean that feedback on coursework is not provided until after students have completed their studying for that course. Much such expensively-provided feedback is likely to be wasted. There may be a trade-off between the rapidity and quality of feedback so that, for example, imperfect feedback from a fellow student provided almost immediately may have more impact than more perfect feedback from a tutor four weeks later. Carroll (1995) described 'formative assessment workshops' for classes of 300 medical students which consisted of multiple choice question test items followed immediately by a short remedial tutorial on the questions. There was no individualised feedback in this system but the feedback was very immediate and the workshop sessions were scheduled to allow students time to study more material before moving on to the next section of the course. 85% of students reported wanting more such sessions. Sly (1999) reported the impact of 'practice tests' on subsequent exam performance. Students had the option of taking a practice test, with computer-based feedback, sufficiently in advance of an exam to enable them to use the feedback to undertake some more study to address their weaknesses. 197 weaker students chose to take these practice tests and these students improved their exam scores so much that they then, on average, outperformed 417 stronger students. The benefits were still evident in a subsequent exam.

3.8 Feedback should be appropriate in relation to students' understanding of what they are supposed to be doing

My daughter, while studying Sociology, received a comment on one of her essays which read: "Not Sociological enough". Her response was: "If I'd understood how to be 'Sociological enough' I'd have done it!" My daughter's experience is echoed many times in accounts in the literature. For example Higgins et al (2001) discuss the failures of communication that take place in feedback. They describe a case in which the tutor's entire feedback consisted of: "A satisfactory effort. More critical analysis of key issues would have helped." The student, who wanted to be better than 'satisfactory', was left frustrated by the poor quality of critical analysis by the tutor. Four different constraints on students' understanding of the nature of academic tasks, and hence of feedback on them, are discussed here.

Students' conceptions of the task

Students have to make sense of what kind of a task they have been set when they tackle an assignment. They can easily misunderstand and be confused by whatever briefing and feedback they have been given, as in this example:

Interviewer: "What do you think the tutor was looking for in this essay?"

Student: "Ah ... well!, this is confusing me. I know the tutor likes concise work, but doesn't like generalisations, and doesn't like too much detail, although on the whole I think he'd like more detail than generalisations. And because it was such a general question, I thought 'oh help!', I don't know what he's looking for."

(Hounsell, 1987)

Whatever feedback this student's tutor gives will be interpreted in the light of the student's conceptions of what the tutor really wants or what the task really consists of. Students can have a great deal of difficulty understanding what form of communication an essay is (when the only audience knows more than they do about the topic), or what a laboratory report is for (when it has already been written hundreds of times before in exactly the same format), or what a design task has been set for (when only the product is assessed and not the learning that was involved in creating it). Many academic tasks make little sense to students. This inevitably causes problems when they come to read feedback about whether they have tackled this incomprehensible task appropriately.

Students' conceptions of learning

Underlying the students' confusion about what the tutor really wants could be an unsophisticated conception of learning. Säljö (1982) describes students as having one of five conceptions of learning:

1. Learning as passive receipt of information
2. Learning as active memorisation of information
3. Learning as active memorisation of information or procedures, to be used at some time in the future
4. Learning as understanding
5. Learning as a change in personal reality: seeing the world differently.

A student with conceptions of learning 1, 2 or 3 might have trouble interpreting feedback that stated: "Not enough discussion" if they had accurately provided the tutor with information they had diligently collected. Feedback needs to be sensitive to the unsophisticated conceptions of learning that may be revealed in students' work.

Students' conception of knowledge

Perry's "scheme of intellectual and ethical development" describes how students develop over time, and through academic experience, their understanding of what knowledge itself is (Perry, 1970). He describes students as starting off thinking that there are an enormous number of facts and that their job is to learn these and give as many of them as possible back to the teacher, correctly. Perry describes this learning process with the memorable

phrase "quantitative accretion of discrete rightness". He describes students as moving through a number of stages of increased understanding of the nature of knowledge towards, eventually, a flexible commitment to a particular way of seeing things, in the knowledge that more evidence or better theory might alter this perspective in the future. Along this intellectual journey many students display extreme relativism, in which all answers are seen as equally right. A student who does not draw a conclusion to an essay may be leaving it up to the reader to decide, given that all conclusions are seen as equally valid. Feedback that simply read 'No conclusion' might not help such a student to progress! Teachers' feedback is often (though not always) generated from a more sophisticated epistemological stance than that of the student and this offers plenty of scope for misunderstanding of feedback or blank incomprehension.

Students' conception of the discourse of the discipline

Lea & Street (1998) describe a student who, after submitting an essay on a History course, received the feedback: "I like your conclusions to what is a carefully argued and relevant essay." At the same time the same student received feedback on an essay submitted on a parallel Anthropology course which was so critical of the student's ability to write a clear argument or produce a justified conclusion that they were advised to seek study skills counselling. Lea & Street interpreted this as a consequence of Anthropology involving a very different form of discourse with different forms of argumentation and use of evidence, as it was clearly not a case of generalised essay-writing inadequacies. If the student did not understand the discourse of Anthropology and was unpractised in using it, then generalised essay-writing advice was unlikely to be helpful, whether from the lecturer or from a study skills counsellor. Feedback needs to be sensitive to what kind of writing is expected and what students are likely to understand about it. In modular course structures it is common for students to cross disciplinary boundaries and have to cope with such differences in discourse. Science and technology students often have particular difficulties with social science-type essays even if they can write in an articulate way in their own discipline, but there are also profound differences in discourse within the social sciences, for example between Sociology and Psychology, and within the Humanities, for example between History and Literature. Northedge (2003) provides insightful guidance on how to help students to learn about the discourse they are being asked to use.

3.9 Feedback needs to be received and attended to

A number of studies have described students receiving their assignment back, glancing at the mark at the bottom, and then simply throwing it in the bin, including all the feedback.

"Sometimes I do read the comments but I find that I'll never write the same essay again anyway .... I tend to ignore them in some ways, unless there is something very startling."
(Hounsell, 1987)

Crooks (1988) has summarised a range of research on this issue. Where marks on intermediate tests or coursework assignments count significantly towards final marks, students pay less attention to accompanying feedback. Jackson (1995) found that third-year students were particularly likely only to look at the grade rather than at feedback on essays. He reported that students liked to see the feedback, but more to assure them that their essay had been read carefully and marked fairly. It is not inevitable that students will read and pay attention to feedback even when that feedback is lovingly crafted and promptly provided. Special steps may need to be taken to engage students with feedback, such as:

- asking students to specify, on their assignment, what they would like feedback on, and giving feedback on nothing else
- providing feedback but no marks, so that students have to read the feedback to get any idea how they are progressing
- requiring assignments to be self-assessed before they are submitted (without any marks being involved) so that students pay attention to whether teachers' views correspond to their own. A review of literature on self- and peer-assessment has reported that overt self-assessment has been shown to increase student performance (compared with a control group, in controlled studies) and increase students' control over their learning strategies (Dochy et al, 1999)
- using two-stage assignments with feedback on the first stage, intended to enable the student to improve the quality of work for a second stage submission, which is only graded. Cooper (2000) has reported how such a system can improve almost all students' performance, particularly the performance of some of the weaker students. A comparison of students' learning responses to assessment on different courses at the Open University (Gibbs et al, 2003) found that the course with much the highest level of student use of feedback involved a sequence of eight assignments, each of which fed into the next one. It would have been a very dim student who did not pay attention to such feedback – not because it was better feedback but because the design of the assessment enabled it to 'feed forward'. Much 'feedback' is ignored while 'feedforward' is much more likely to be effective
- providing a grade only after self-assessment and tutor feedback have been completed. Taras (2001) reports the successful use of such a sequence as a component of summative assessments.

3.10 Feedback should be provided in such a way that students act on it and change their future studying

This issue concerns the impact of feedback on future learning. Feedback may accurately correct errors but still lead to no change in the way a student goes about the next assignment or tackles any future learning task. This may occur for a variety of reasons:

- feedback may come too late to be acted on by students
- feedback may be backward-looking – addressing issues associated with material that will not be studied again, rather than forward-looking and addressing the next study activities or assignments
- feedback may be unrealistic or unspecific in its aspirations for student effort (e.g. "read the literature" rather than "for the opposite view, see Smith Chapter 2 pages 24-29")
- feedback may ask the student to do something they do not know how to do (e.g. "express yourself more clearly")
- feedback may be context-specific and only apply to the particular assignment rather than concerning generic issues such as study skills or approaches that generalise across assignments
- feedback may be discouraging and lead to less study effort rather than more
- there may be no follow-up to check if students have taken any action, so they can ignore feedback with impunity.

Ding (1998) suggests that even if students read feedback comments, they often do little with them. In contrast Brookhart (2001) found that successful students use both marks and feedback and actively self-assess, both to learn and to direct their future studying. The most important variables here may be, as so often, to do with the student rather than with the teacher. Teaching students to monitor their own performance is, in Sadler's theoretical analysis of the role of feedback, the ultimate goal of feedback (Sadler, 1989). Research on the impact of the use of 'classroom assessment' in college in the USA again and again stresses the impact not on the learning of specific content but on the development in students of 'meta-cognition' and the ability to gain control over their own learning (see Steadman, 1998, for a summary). Students are likely to need to be taught how to use feedback to develop meta-cognitive control (Sadler, 1998). Improved ability to learn may not have the effects hoped for, however. Ramsden et al (1987), studying the impact of a 'study skills' programme designed to increase the extent to which students adopted a deep approach, found it actually achieved the opposite. Students' increased awareness enabled them to adopt a surface approach to a greater extent in order to meet the perceived low level demands of their courses' assessment! Again this illustrates the way students' perceptions of assessment influence their learning.

4. Assessment tactics that support student learning

4.1 Capturing student time and effort and distributing that effort across topics and weeks

Capturing students' effort across a whole course basically requires regular assignments. It is neither necessary nor productive to mark all these assignments to gain student engagement. Instead a range of tactics is available:

- **Completion of assignments as a course requirement, without marking.** Course requirements can be imposed as a condition for taking a subsequent assignment or examination. For example Forbes & Spence (1991) describe an Engineering course in which students had to complete about 80% of an extensive set of problems on problem sheets or they were not allowed to sit the end of course examination which carried 100% of the marks.
- **Sampling of assignments for marking.** The conventional rationale for setting examinations is that it is impossible to ask questions about everything and so the exam paper samples the curriculum, asking about perhaps 10% – 25% of possible topic areas, on the assumption that this will give a reasonably accurate indication of how much students have learnt across the entire curriculum. If students are clever at guessing what will come up in the exam then this is not a good assumption. Alternatively you can sample assignments for marking. Students might be expected to tackle eight assignments but only two (25%) will be marked. To get students to take all of them seriously you would probably have to sample these two randomly so that students don't know which will be marked. An additional advantage of this tactic is that it is much more risky for the student not to take each of the assignments seriously in case that assignment is the one that is marked, whereas if all of them are marked, then individually they carry fewer marks and it is less risky for a student to submit sloppy work.
- **Mechanised (e.g. computer-aided assessment (CAA)) testing and/or marking.** Multiple choice question testing can be a very economical option and allow a much larger proportion of the curriculum to be tested and student effort to be more widely distributed. However, as was mentioned earlier, this can be at the risk of accidentally mis-orienting students to low-level educational goals, and in particular to memorising facts.
- **Self- and/or peer-marking.** There is a considerable literature on whether you can trust students' own marks – of their own work or of the work of others (see for example reviews by Falchikov & Goldfinch, 2000, and Dochy et al, 1999) – sufficiently to use them as a substitute for tutor marks. In summary, with certain safeguards it is possible to produce levels of reliability from both self- and peer-assessment that are similar to levels of reliability of tutor marks. However, this is largely because the level of reliability of tutor marking is also low, rather than because students are particularly reliable. Students are not very good at judging unconventional or 'professional' tasks, and are poor at producing marks against each of a list of criteria: they are more reliable when making a single overall judgement of a form of academic assignment they are familiar with, such as an essay. But the real issue here is that you almost certainly do not need more marks. The real value of self- and peer-assessment lies in its impact on students' learning to self-supervise in relation to standards that they have come to understand, through marking. More would be gained by making self- or peer-assessed tasks a course requirement and then concentrating on the usefulness of self and peer feedback, than by training students to produce reliable marks.

Example

A Law course that previously had one marked essay in a semester had failed to capture student time and effort and students ended up studying the topic they wrote an essay on, and little else, before revising for the end-of-semester exam. This was replaced with a course requirement to complete six essays – roughly one every two weeks. The first three had to be submitted by set dates and feedback consisted of a 'model answer'. Students briefly discussed with each other the difference between their essay and the model answer, in a peer-review exercise in class. One of the second set of three essays was randomly selected to be marked, but students did not know which one it would be, and so had to take all three seriously. There was no feedback on these essays. In this new assessment pattern, students tackled six topic areas in some depth, instead of one, put in much more effort, and distributed this effort reasonably evenly across the weeks of the semester. As there was no written feedback on essays 4–6 it took the teacher less time than previously.

- **Exam demands** that are unpredictable and/or sample everything, so that students have to study everything. The following two students, studying Social Science in two different universities, had quite different perceptions of what they needed to do to prepare for their examinations, on the basis of looking at past exam papers:

Student 1:

"I had a module that we did, we had three questions to answer and two of them were on specific thinkers. And we had maybe ten thinkers that we could look at, but you only had to answer a question on two of those, and for that I only did three thinkers for it, knowing that if I did three really well I'd be able to answer a question on two of them. I did another option ... and for that we had to answer three questions in three hours, and there [were] ten topics, you would only be asked on three of them so I did four topics. So I learnt four topics really well, and then didn't really revise six topics at all."

Student 2:

"I think to do really well in the papers, or most of the papers, knowing all the syllabus would be very helpful. Actually it's necessary to a certain degree because, I know that in certain papers ... what they do is ask a question on a particular topic, but it's linked into another topic, which means that if you've revised one of the topics but not the other, you can maybe half answer the question, or not answer it fully ... sometimes it's linked in with something else. And if you don't know the other topic then it's problematic. So definitely knowing more would be helpful, and it often helps everything fit into place because when you go through it in the eight weeks when you're doing the tutorials it all seems a bit distinct and isolated, but when you come back to revising it, it's like 'oh, this all fits together'. So I think knowing everything is very helpful."

Student 1 distributed her limited study effort narrowly on a few topics, while student 2 studied everything, and in depth, driven largely by their different perceptions of what was required in the exam.

4.2 Generating high quality learning effort, oriented towards clear and high standards

Large-scale open-ended assignments tend to be experienced as more challenging, improve engagement and induce a deep approach. In contrast small-scale, short and simple assignments that are easy to tackle can produce only very superficial and short-lasting engagement. Project work courses tend to capture not simply more effort, but better quality effort. Comparisons of 'problem-based' and conventional courses, in a variety of subject areas, tend to show marked differences in the extent to which students take a deep approach (see Dochy et al (2003) for a review). Complex problems seem to be inherently intellectually stimulating.

Assignments involving interaction and collaboration with other students, in or out of class, also work consistently well across a range of disciplines to engage time and effort, as well as producing better learning outcomes (see Springer et al (1999) for a review of 383 studies comparing group-based with individual study).

Explicit specification of goals, criteria and standards in course guides may help, but the evidence about students' understanding of written criteria is not very encouraging (e.g. Price & Rust, 1999). Sometimes this is because the specification itself is thoroughly confusing, as in this example:

Interviewer: "Did the feedback that the tutors wrote on it relate to the criteria that they had set for the coursework?"

Student: "I suppose I did get a bit confused about that, because we'd have a topic, and maybe there were three questions. And then they'd show a percent weighting on each question. But then we'd have this criteria sheet, that would be assessing knowledge, content, references and how it was written, and that would have the same sort of percent. Sometimes I never knew which side they were marking from, whether I should focus more on question 1, or should I be focusing on my content overall.'

(Gibbs & Dunbar-Goddet, 2007)

But even apparently unambiguous specification often leaves students none the wiser. The desired effects may be better achieved through showing students high quality exemplars of the kind of work you would hope they could produce (and contrasting this with examples of several different ways of doing the wrong thing to the wrong standard). What is needed here is not so much explicit specification as student internalisation of these goals, criteria and standards. This is more likely to be achieved through student marking exercises, or through students publicly presenting their work and having it critiqued in public. Students may come to understand what the words mean in stated criteria through such exercises, but even teachers have trouble agreeing what criteria really mean, or explaining the words, without referring to instances. The best that can usually be hoped for is that students come to be able to make somewhat similar judgements about standards to those of their teachers, even if they cannot explain the criteria or standards very clearly. Standards are slippery things. While they may be broadly shared within an academic community (such as a group of lecturers who have been marking the same kinds of assignments, and discussing their grades in examination committees, for many years) they cannot be communicated directly at all easily. New lecturers often get standards completely wrong and it takes them a while to gradually adjust their marks and their feedback to what is intended and accepted locally, usually by seeing that colleagues do things differently. Students need to go though a similar kind of induction into the local community: seeing a range of quality of work; seeing how experienced teachers view and mark this work, and why; discussing the qualities of various pieces of work with each other; and gradually calibrating their judgements until they are broadly in line.

Avoidance of tests and exams passable by memorisation, that induce a surface approach in students. Teachers rarely set tests or exam questions with the deliberate intention of inducing a surface approach, but they do often allow students to accumulate enough marks to pass without ever doing anything more sophisticated. For students, that may be all the encouragement they need.

Highly challenging exams requiring 'performances of understanding'. The goal here is to articulate clearly for students the challenging nature of the demands of examination questions in such a way that this re-orients their efforts towards appropriate kinds of learning activity.

Marking exercises. Perhaps the easiest and most effective way of orienting students to allocate the right kind of effort in an appropriately focused way, in relation to assessment demands, is to conduct a classroom exercise in which students are asked to mark three or four good, bad and indifferent assignments from students from the previous year (with their permission, and made anonymous). Students should read and allocate a mark to each example without discussion, then discuss their marks and their reasons for allocating these marks with two or three other students who have marked the same assignments. The tutor should then reveal the marks the assignments actually received, and why, in relation to the criteria and standards for the course. Finally two more examples of assignments should be provided for the students to mark, with their now enhanced understanding of the criteria. Students undertaking such exercises have been shown to gain one grade higher for their course than they would otherwise have done, for the investment of about 90 minutes in the marking exercise, and students continued to gain this benefit of a higher grade on a subsequent course (Rust et al, 2003). There can be few more cost-effective interventions.

Example

A Philosophy of Education course used to set exam questions of the form "Compare and contrast the educational philosophies of X and Y". Having seen that this type of question was common in the past, many students responded to the perceived exam demands by memorising a few facts about X and a few about Y in case they might be useful, and when prompted by such a question in the exam, listed these facts in the order in which they remembered them. Some of the brighter students could spot the difference between X and Y. The exam was changed so that students had to demonstrate their understanding by applying their knowledge of philosophy to teaching – which was, after all, the purpose of the course in the first place. They were told that in the exam they would be shown a video of a teacher that illustrated a philosophical dilemma of some sort, for example involving issues of power and control. They were even told exactly what the exam question would be: "Advise the teacher on the video about their future teaching practice, from a philosophical point of view." What they were not told was what would be on the video. Students responded to this assessment demand by watching teachers, looking at videos of teachers, discussing what they had seen with other students, and then going back to their notes and their books to see what philosophy might have to offer in analysing what they had seen.

4.3 Providing sufficient timely feedback to students on their work

Regular assignments, starting early. Students need early feedback, for encouragement and to orient their efforts throughout the rest of the course, and regular opportunities to use and tune up what they know, and know how to do, through assignments with feedback.

Quality standards for the volume and quality of feedback. The Open University, who have the highest ratings for 'feedback' in the National Student Survey, monitor the standard of feedback that tutors provide to students. An experienced Staff Tutor samples new tutors' marking and if they see feedback that falls below accepted standards (for example, too brief to be understandable) or is of an inappropriate form (for example, overly critical, with little advice on how to improve) they will contact the tutor for a discussion, and that tutor's feedback will be put on a higher level of monitoring until it is seen to improve. Most universities monitor the quality of lectures in this way as a matter of course, but feedback has much more impact on student learning than do lectures. The OU also have strict rules on turn-round times for feedback and if tutors are persistently too slow they are likely not to be re-employed.

Tutor briefing concerning volume and nature of feedback. The Open University also focus their tutor training on how to give feedback. They provide exemplars of good feedback and advice on using the 'OU sandwich' of positive comments, advice on how to improve, followed by an encouraging summary.

Mechanised feedback, where mechanised tests are used. Students can sometimes gain from marks on mechanised testing – mainly about whether they have been working hard enough and which topics they need to spend more time on. Tests in class can work especially well in this respect. Some software also allows tutorial feedback to be associated with the selection of wrong answers to multiple choice questions, with students getting an electronic summary of feedback the moment they press the 'submit' button on the on-screen test. Such feedback may be paid more attention if there is a later opportunity to retake the test (see 'Two-stage tests', below). Providing 'self-tests' online, for feedback purposes only, for students to take when they feel ready can also lead to more thorough and focused studying before a test or exam for marks at a later time, guided by their performance on the 'self-test'.

Audio feedback is being adopted to an increasing extent at Leeds Metropolitan University, allowing tutors to speak their comments into a digital tape recorder as they read assignments and email their digital audio files to their students in a rapid and automated way. Several of the examples and case studies below involve audio feedback in imaginative ways. Information about the technicalities and educational issues involved can be found at: http://sites.google.com/site/soundsgooduk/

Development of student self-supervision. The fastest feedback that can be provided is by students themselves, given to themselves, as they are writing or studying. They become able to 'self-supervise' in this way through practice at self-assessment (for example through being required to add self-assessment comments to their own work when they submit it), through seeing, judging and discussing examples of work of various standards (for example through 'marking exercises'), and through comparing their own self-assessments with assessments of their work by other students and by their tutor (for example being asked to discuss the feedback they have received on their most recent assignment, in a short classroom exercise).

4.4 Providing high quality feedback

Being specific and forward-looking. Good feedback is sufficiently detailed that students understand what, exactly, is meant, and also what to do next time to avoid the same mistake or to improve.

Encouragement. The emotional tone of feedback has a good deal of impact. Good feedback encourages students and increases their 'self-efficacy' – their belief that they are capable of doing well. Negative and personally critical comments are ineffective and damaging.

Feedback relating to educational goals, criteria and standards. Feedback should make it clearer to students what the educational goals of the course are, for example whether greater emphasis is placed on familiarity with the literature or on competence. Students should be able to see how marks are arrived at in relation to the criteria, so as to understand the criteria better in future. They should be able to understand why the grade they got is not lower or higher than it actually is. One way to do this is to use the sentence stems: "You got a better grade than you might have done because you..." and "To have got one grade higher you would have had to ...". Feedback sheets with lists of criteria may help, but if students do not understand the words used (such as 'quality of argument') then a tick or a 'good' next to the criterion will not explain much.

Tutor briefing and monitoring concerning the quality of feedback. New tutors often have a limited feel for what good feedback looks like or what standard of feedback, in terms of length and specificity, is expected. They may concentrate on proving their superior knowledge to the student rather than focusing on improving the students' work in future. It helps to provide new tutors with samples of feedback from exemplary tutors, and also examples of inadequate feedback, and to arrange private meetings in which new tutors can discuss samples of their feedback with experienced tutors. The Open University monitors the quality of feedback tutors provide.

Example

A course includes a seminar where students prepare material to share as a presentation in class, rather than the lecturer providing content. This is usually an ideal opportunity for students to engage in the content in a deeper way and to receive some formative oral feedback from their tutor, and to a lesser extent from their peers.

Students really value this opportunity and usually comment on its usefulness in evaluations. However, there are often errors and omissions repeated in the subsequent summative assessment that had been clearly identified in the formative feedback in the seminar. Tutors now digitally audio record the oral feedback given to each group of students during the seminar and deposit this on X-stream [the University's Virtual Learning Environment] for students to refer back to when they tackle the subsequent assignment. This has proved extremely popular and has resulted in improved performance in this assignment, and improved staff efficiency in not repeating feedback or having to review student work prior to its submission.

Ollie Jones
Faculty of Business and Law
Leeds Metropolitan University

Development of students' ability to understand feedback. Often tutors' feedback is understandable to other tutors, but not to students. It can help to put short periods of class time aside to have students discuss with other students the meaning and implications of the feedback they received, and to see feedback other students have received and hear what sense others make of their feedback. Confusions can be aired and resolved in public, for all to hear.

Example

Feedback to students undertaking translations from Spanish to English uses a colour coding scheme that students get used to interpreting.

Graham Webb
Head of Languages and ELT
Leeds Metropolitan University

Example

Digitally recorded audio feedback can be used successfully to replace written feedback and improve its quantity and quality. It is easy to use and it allows you to say all the things you are thinking in your head about the strengths and weaknesses of an assignment without having to translate it first into 'academic peak' so that it's grammatical and fits on the feedback proforma. I found I could say more in the same length of time as it took me to write feedback. Students appreciated the conversational nature of the feedback and it was more accessible to them. Their understanding of the feedback was significantly better. One of my students told me, having received audio feedback on one of her assignments, she now suddenly realised where she had been going wrong with all the others she had submitted previously (and had received written feedback on).

Mandy Asghar
Former course leader
BSc Physiotherapy in the Faculty of Health
Leeds Metropolitan University

Colour coded feedback on translations

Colour used	Feedback
Good translation: Purple	Really captures the meaning and sense of the original in well expressed English
Weak translation: Turquoise	A bit too literal, not well expressed in English, doesn't sound very English
Poor translation: Brown	Far too literal, poorly expressed in English, doesn't mean anything, syntax (word order) incorrect
Wrong translation: Red	Mistranslated, the original doesn't mean this, comprehension of original poor
Suggested word or phrase: Green	This word or phrase is the correct one or is more appropriate
Other comment or suggestion: Blue	Positive comment or suggestion for improvement
Indicative mark	

4.5 Ensuring that students use feedback

Faster feedback

The most effective way to ensure that students are likely to pay attention to feedback is to provide it quickly: fast enough that they are still thinking about their assignment and soon enough to be useful to them in tackling the next assignment. Quick and imperfect feedback can be much more effective than slow and perfect feedback. Providing fast feedback can be expensive – section 4.3 contains ideas for reducing the resources required to give feedback quickly.

Requests for feedback

If you ask students to make specific requests concerning what feedback they would like when they submit an assignment, then they will already be thinking about feedback before they receive it; you can target feedback just on those things they are interested in, and they are more likely to read it. It could save you time as well, through not giving feedback on anything the student did not request. Students will get better, over time, in making more useful requests.

Discussion of use of feedback

If you have gone to a lot of time and trouble to provide feedback then it makes sense to put aside a small proportion of class contact time to arrange discussion between students. Ask them to bring their last assignment with its feedback to class and ask them, in twos or threes, to take turns to answer these questions:

What did you do well in this assignment?

What would you have needed to have done to have got a better grade?

What can you learn from this assignment that could help you in your next assignment?

What would you like feedback on next time?

Two-stage assignments

If assignments are designed in two stages, with the first stage formative (with feedback but no marks) and the second stage summative (with marks but no feedback), then students are likely to pay a great deal of attention to your feedback on the first stage and to produce better quality work at the second stage. PhD supervision, involving feedback on drafts of chapters, uses this principle. Carless (2002) reports that feedback of this kind can successfully re-orient student effort in appropriate ways.

Example

Previously assessment of a double module involved a single substantial project report (10,000 words) with no formative assessment or feedback before its final submission. This has been supplemented by interim formative-only assessment which involves both a poster presentation and a succinct 4,000-word report, both of which receive feedback which students can use in drafting their final report.

Steve Wilkinson
Level 3 Individual Project
BSc (Hons) Multimedia & Entertainment, Hong Kong
Innovation North, Faculty of Information & Technology
Leeds Metropolitan University

Two-stage tests

Where assessment takes the form of short tests or multiple choice question tests, the feedback that students normally get – which questions they got wrong and right – is not much use as the next test will be on different subject matter. If students get two 'goes' at similar tests, with the first treated as a formative assignment and the second, about a week later, just for marks, they will put extra time into working on those topics where they made errors the first time round. A number of studies have shown substantial improvements in student grades, and

reductions in failure, through the introduction of two-stage tests to replace one-stage tests.

Integrated multi-component assignments

You can set students a substantial assignment or project which is tackled in multiple stages, each of which contributes to the next stage and to the whole report submitted at the end for marks. At each stage you can introduce some kind of feedback mechanism (for example peer feedback using a checklist, or a 'model answer' that is discussed). This provides ideal opportunities for 'feedforward', maximising the likelihood of students paying attention to and using feedback.

Requirement for students to demonstrate response to feedback in subsequent assignments

You can require students to attach to their assignment a cover sheet which explains how they have used previous feedback to tackle this assignment more effectively. This can be made a formal course requirement: if they do not have an adequate cover sheet then you can refuse to mark it. When you provide feedback you can concentrate on how well they have improved their work since last time.

Generic feedback

Much feedback concentrates on correcting errors or highlighting omissions. Students very rarely respond to such feedback by going back to the previous topic and studying it some more. Of much more use to students is more generic feedback that would be of help in tackling the next assignment, or even somewhat similar assignments on other courses, almost regardless of the specific topic. Advice on reading around, writing, using evidence, constructing arguments, referencing sources etc is all more likely to be paid attention to and responded to. See Phil Race's website at **http://phil-race.co.uk/?page_id=13** for a compendium of extracts from his work on feedback.

5. Case studies of the use of assessment to support student learning within Leeds Metropolitan University

5.1 Engaging students with feedback

The situation

At Level 1 of the BSc (Hons) Physiotherapy much of the assessment is practically orientated and the students only submit two assignments during the whole year. This has resulted in students having difficulty in grasping the nuances of writing at Level 1 and often being disappointed with their marks. It is also difficult to evaluate how much the students engage with the feedback that is provided to them subsequently and if they actually then go on to use it to improve their standard of writing in Level 2 assignments. In 2006-07 six out of 37 students failed the skills assignment at first submission, so it was identified as an area that would benefit from the implementation of a formal formative assessment strategy.

The strategy

The skills assignment traditionally involved two 1,000-word sections based on a case study scenario: the students had to justify treatment approaches to address the patient's problems in their answers, using research evidence and clinical reasoning. Students were provided with marking criteria to assist with their essay. The formative strategy employed required the students to submit either of the two sections of the assignment as a draft approximately five weeks before the final submission date.

Students were given feedback (not a mark) using the following proforma within 2-3 weeks. This gave the students a minimum of two weeks to use the feedback to change their assignment and to state how they had used their feedback.

On submission of the final assignment students were required to complete the right-hand box of the proforma and attach it and their draft as an appendix. 10% of the marks for the final assignment were allocated to the students' responses to feedback on the proforma.

The results

As a result of the strategy, in last year's cohort of students two out 40 students failed their assignment: a decrease of four. The marks gained for the section that was submitted for feedback were better than for the section the students did not submit. It was hoped that the feedback for one section would have an effect on the quality of the un-submitted section but that does not seem to have been the case. The students varied in their responses to the feedback as detailed in their comments on the proforma but it was evident that many had used the feedback in a positive way to change their final piece of work.

It is often said that this formal submission of drafts for feedback is time-consuming

Name of student: Section 1/Section 2 Delete as appropriate		
Strengths of the submitted section	Tutor comments:	
Areas to develop	Tutor comments:	Student comments: What did you do to address the areas suggested by the tutor? How did you incorporate the suggestions made by the tutor into your final submission? What do you think you have learned from the feedback you have received?
1		
2		
3		

but from a staff point of view, commenting on the 1,000-word drafts did not take long as many of the students made similar mistakes. The value of it for students vastly outweighs the extra effort required.

Mandy Asghar
Physiotherapy

Comment

The tactics used here were:

1. *to use a two-stage assignment, so that feedback on the first stage could help students with the second stage submission*
2. *to use a structured feedback form that looked forwards, concentrating on what students could do to improve the work, rather than backwards (and concentrating on what was right or wrong)*
3. *not to allocate marks to the first stage, so that students concentrated on the feedback*
4. *to provide the feedback on the first stage quickly enough to allow students two weeks in which to act on it*
5. *to require students to state, on their second stage assignment, how they used the feedback on the first stage to improve their work.*

5.2 Multiple assignments and early feedback

The module on Animation Concepts had a previous life as 3D Animation Technology and suffered from student attrition. Previously the formative assessment was an end-of-semester upload to the Virtual Learning Environment (VLE) and marking and feedback were entirely online. This was unsatisfactory as students could not make use of the feedback in order to improve their work, neither was there a requirement for work to be submitted at an earlier stage.

The new module started by making the assignments more engaging, by asking the student to create an 11-second emotive animation, and by increasing the number of assessment points. A first-week assessment took the form of an online quiz, called the 'animation audit', which asked a number of questions about students' skills and knowledge in each of the different animation topic areas. It also got them to view previous students' work and asked to which broad classification each piece of work corresponded. Other questions encouraged students to explore the VLE in order to review the assignment details, such as hand-in dates.

A week 3 assessment asked the student to show planning through the creation of a moving storyboard, called an 'animatic'. The assessment explored the sequence and duration of each part of the story and how cinematography, lighting and audio would enhance the emotion being portrayed. Feedback for the animatic was given in class and via the VLE. This helped the students improve their final animation and gave them an opportunity to branch into different forms of animation.

The final assessment was a complete animation and involved assessment in more detail than previously.

The benefits of this new three-point formative assessment are that student engagement can be assessed from week 1,

and that an interim assignment gives the student an opportunity to plan and receive constructive feedback before investing large amounts of time in using the technology.

There has been a reduction in non-submissions and an improvement in grades.

Steve Wilkinson
Innovation North Faculty of Information and Technology
Leeds Metropolitan University

Comment

This revised assessment pattern:

1. *confronts students with a required and assessed task in week 1, ensuring they engage with the course early*
2. *shows students exemplars of past student work early on, orienting them to the standard required and the nature of the work they are eventually required to do*
3. *uses assignments that students find more engaging*
4. *provides formative feedback at two stages (and in two formats), building up component knowledge and skills that students can use in tackling a larger and more complex later assignment.*

5.3 Improving skills and reducing student errors through linked assignments and oral feedback in tutorials

This case study concerns a Level 1 undergraduate module undertaken by approximately 300 students across a variety of Construction disciplines. The module is assessed by an end examination (50%) and coursework (50%). In previous years, the coursework required the students to produce three sketch details of various parts of a dwelling (foundation, external wall and a pitched roof). The assignment was handed out at the beginning of the module and students were given ten weeks to undertake all three sketch details.

Structuring the coursework in such a way presented a number of problems:

- There was a lack of motivation from students in undertaking the coursework. Students frequently left the coursework to the last minute, resulting in poor quality, rushed sketches and low coursework marks
- It was common for students to make the same mistake on all three sketch details (for instance, incorrect annotation or use of an incorrect graphical symbol), resulting in them being penalised for the same mistake three times
- The students gained little benefit from the detailed feedback that was provided on their work. Students often commented that they ignored the feedback, as they were not required to undertake any more sketch details within the module.

The assessment was restructured and divided into three separate elements, corresponding to each of the sketch details. Each element of coursework was issued early at the beginning of the module but the hand-in dates for each coursework element were distributed evenly across the semester, giving the students approximately three weeks to complete each sketch detail. The tutorials were also re-structured.

Students were encouraged to present draft sketch details to the tutors for comment during the tutorial sessions, and in some of the sessions there was a requirement for the students to spend some of their time undertaking draft sketches in class.

The impact of redesigning the coursework in such a way was as follows:

- Feedback on the sketch details was valued by the students and used to inform future submissions
- Very few students repeated errors from one sketch detail to another
- Student motivation increased and draft details were presented during the tutorial sessions
- Student marks improved overall. On average, the marks obtained for each of the sketch details improved incrementally, with the highest mark being obtained for the final sketch detail.

David Johnston
Introduction to Construction Technology

Comment

The revised assessment pattern captured more student effort and distributed it more evenly across weeks and topics. It provided feedback on early assignments that was useful to students on later assignments: feeding forwards and grabbing students' attention. By providing oral feedback on drafts it turned each of the three assignments into a two-stage assignment – providing six opportunities to get better instead of the previous one.

5.4 Using formative assessment to help students to tackle unfamiliar forms of assignment

Principles and Practice of Forensic Analysis is a Level 3 module for students on BSc (Hons) Biomedical Sciences and BSc (Hons) Psychology with Forensic Biology. Summative assessment involves a report of a case (real or created) together with a structured discussion in essay form designed to ensure students address each of the learning outcomes of the module. As this type of assessment is unfamiliar to the majority of students, they are offered the opportunity of submitting their draft case report plus essay plan. Feedback is then provided on the suitability of the case, the appropriateness of the reporting style and whether their plan is in line with the assessment criteria.

Typically just over 50% of the group take up the opportunity, and for the last two years records have been kept to allow assessment performance to be compared. On each occasion, the average mark for those who have submitted work for formative feedback has been more than 10% greater than the average mark for those who have not. Clearly, there may be differences in the level of engagement between those who do and do not take up this opportunity, but overall, this exercise does appear to improve performance and appears to make the final marking easier.

Alison Caswell
Public Health Safety and Environment

Comment

This innovation provides students with the opportunity for turning a single stage assignment into a two-stage assignment with formative-only assessment of the first stage. Students can use this opportunity to avoid doing unproductive or unsuitable work later on and can re-orient their efforts towards the learning outcomes for the course. There is evidence from studies elsewhere of similar opportunities to understand unfamiliar forms of assignment leading to one grade (or 10%) better performance for those students who have availed themselves of the opportunity.

5.5 Peer coaching and public formative assessment of groups

In the BSc (Hons) Physiotherapy course students are allocated to teams of four to work in four peer coaching sessions throughout their first year. The assessment is designed to test their competences formatively in the execution of practical skills such as massage and exercise prescription. A two-stage formative assessment process is then employed at each session to evaluate the students' ability. At the outset, students are alerted to which three or four skills will be tested at each of the sessions and of the need to ensure that the relevant short written component of an accompanying skills log is completed prior to the assessment.

At each session students are given short clinical case scenarios to which they have to apply a particular physiotherapeutic skill. In the first stage of the formative assessment process students practise together as a team, ensuring, through feedback to and from each other, that everyone is able to perform the skill and to answer questions concerning clinical reasoning, safety, physiological effects, and so on. In the second stage the tutor chooses, at random, one member of the team to single-handedly demonstrate the skill and to answer questions. If that student is successful the entire group have their individual competence section of the reflective skills logs signed off by the tutor. Should, however, the student fail, the group (in recognition that it is a group rather than an individual responsibility) is given feedback and further practice time to rectify mistakes prior to a re-assessment. This subsequently occurs, without penalty, in the same session with the aim of promoting assessment as a learning opportunity. This learning benefits the students in two ways: first by providing a record of competence in the individual skills tested, and second through the learning feeding forward to the exam at the end of the semester when each student is summatively assessed individually. The crux of the strategy is that each member is individually accountable and the group have to work together because to achieve their goal they must all have input and work co-operatively as a team. This means that each student has a vested interest in his or her own learning and in that of the other members of the group.

Students valued the feedback about their knowledge and abilities from the formative assessment process but they also expressed frequently a willingness to engage with reciprocal peer coaching as it provided the pressure that made them study. It increased time on task during the activity, helped them to set short-term goals and created a culture where learning was shared (Asghar, 2009).

Mandy Asghar
Physiotherapy

Comment

This assessment method employs leverage from several social processes:

- *interdependence in a group in such a way that one student can let their group down (and lead to the group having to re-do an assessment) if they do not engage*
- *assessment and feedback in public, which has quite a different emotional dynamic from assessment in private*
- *very informal peer feedback leading up to the formal assessment, in such a way that students take the business of providing useful peer feedback seriously as the student they give the feedback to might be representing the group during the formal assessment.*

6. Evaluating the impact of assessment on student learning

The definitions in Table 1 have been used in research studies to characterise programme level assessment environments (Gibbs & Dunbar-Goddet, 2007) and you can use these definitions to characterise your own programme and compare it with others.

Table 1: Coding characteristics of assessment environments

Characteristic of assessment environment	Low	Medium	High
% marks from examinations	below 40%	between 40 and 70%	more than 70%
Variety of assessment methods	1-3 different methods	4-6 methods	6+ methods
Volume of summative assessment	mark allocated less than 15 times	15-40 times	more than 40 times
Volume of formative only assessment	less than 15 times	15-40 times	more than 40 times
Volume of (formal) oral feedback	less than 15 hours	15-40 hours	more than 40 hours
Volume of written feedback	less than 3,000 words	3,000-6,000 words	more than 6,000 words
Timeliness: average days after submission before feedback provided	more than 14 days	8-14 days	1-7days
Explicitness of criteria and standards	Explicit criteria and standards rare and/or nebulous; marks or grades arrived at through global judgment in tacit way; no effort to enable students to internalise criteria and standards	Criteria for some assignments and exams; weak link to marks or grades; little effort to enable students to internalise criteria and standards	Clear criteria for most or all assignments and exams; link made to grades; effort made to enable students to internalise criteria and standards
Alignment of goals and assessment	Learning outcomes rarely or weakly specified at either programme level or course level; very weak or rare link between learning outcomes and choice of assessment methods; no explicit link between learning outcomes and allocation of proportions of marks; only overall grades recorded	Learning outcomes specified at programme level but weakly specified at course level; no explicit link between learning outcomes and allocation of proportions of marks; only overall grades recorded	Learning outcomes specified at programme level and for most or all courses; documentation shows how each assessment links to each learning outcome at the course level; some link to marking procedures; student performance recorded in relation to outcomes

Degree programmes differ enormously on these variables and Table 2 illustrates the range that has been found so you can compare your own degree programme with others.

Table 2: Range of characteristics of assessment environments found in different degree programmes

Characteristic of assessment environment	Minimum	Maximum
Percentage of degree marks derived from examinations	17%	100%
Percentage of degree marks derived from coursework	0%	83%
Total number of times work marked per student	11	95
Variety of assessment methods	1	18
Total number of formative-only assessments per student	2	134
Total number of words of written feedback per student	2,700	10,350
Total number of hours of oral feedback per student	3	68
Average number of days between submission of assignment and feedback	1	28

The following programme level assessment characteristics have been found to be associated with a variety of positive student learning responses:

- a high volume of formative-only assessment
- a high volume of oral feedback
- timely feedback

... and to a lesser extent:

- a high volume of written feedback
- a low proportion of marks from coursework.

The following programme level assessment characteristics have been found to be associated with a variety of negative student learning responses:

- a high volume of summative assessment
- a wide variety of types of assessment
- highly explicit criteria and standards
- highly 'aligned' assessment (with different assessments and criteria associated with each learning outcome).

6.2 Effort graphs

Students can be asked to estimate their weekly study effort in relation to assessment demands on a graph with the weeks and assignments marked along the bottom axis, as in the examples below. These estimates may not be accurate in terms of the total number of hours expended (students tend to exaggerate this) but will reflect the students' perceptions of the distribution of their effort. Two illustrative examples can be seen in Figures 1 and 2 which have different patterns of assessment and correspondingly different patterns of student effort.

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

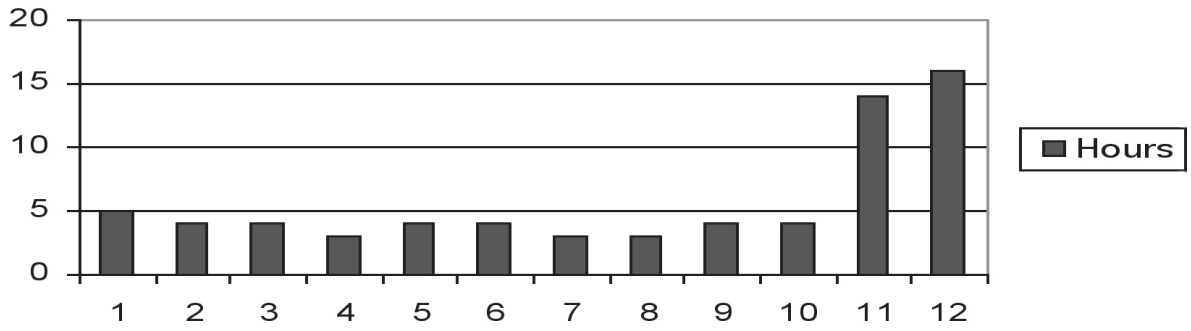
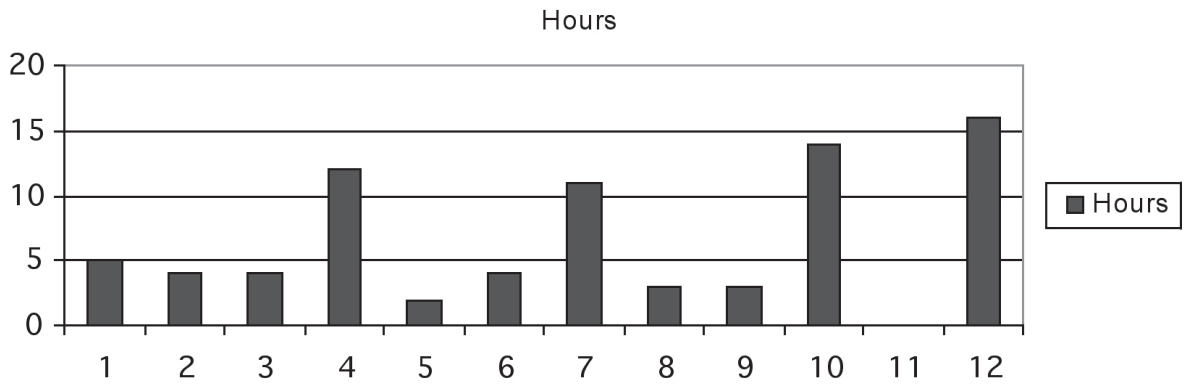


Figure 2: A student's estimate of her weekly study effort (in hours) on a course with three assignments due in weeks 4, 7 and 10 and an exam in week 12. Total study effort = 78 hours



It may be possible for students to sketch rough graphs like this without keeping a log as they go along, and it may be possible for teachers to guess the shape of such graphs themselves.

6.3 Assessment Experience Questionnaire

This questionnaire has been widely used to evaluate the impact of assessment environments, in the UK and internationally. Versions have been translated into Spanish, Swedish and Mandarin. It distinguishes well between programmes and identifies patterns of student response. It takes just a few minutes for students to complete. To interpret results, add up the ratings for each question within each scale (see 'Scoring the AEQ' below to see which questions belong to each scale). Table 3 below shows the range of scale scores that have been found for a wide range of degree programmes in various universities, so you can compare your own scale scores.

Table 3: Means (and standard deviations) for the degree programmes with the highest and lowest mean scale scores (range of possible scores = 1 to 5)

Scale	Degree programme with the lowest mean scale score	Degree programme with the highest mean scale score
Quantity of effort	3.14	3.91
Coverage of syllabus	2.21	4.05
Quantity and quality of feedback	2.68	3.76
Use of feedback	3.21	3.90
Learning from the exam	3.02	3.91

Assessment Experience Questionnaire (V3.3)

Degree course:	
Please respond to every statement by circling sa, a, ?, d or sd to indicate the strength of your agreement or dis Please respond with respect to your experience so far of the entire degree course named above, including all its as	
1	I used the feedback I received to go back over what I had done in my work
2	The feedback I received prompted me to go back over material covered in the course
3	I received hardly any feedback on my work
4	You had to study the entire syllabus to do well in the assessment
5	The assessment system made it possible to be quite selective about what parts of courses you studied
6	The way the assessment worked you had to put the hours in regularly every week
7	It was always easy to know the standard of work expected
8	I paid careful attention to feedback on my work and tried to understand what it was saying
9	The teachers made it clear from the start what they expected from students
10	The staff seemed more interested in testing what I had memorised than what I understood
11	It was possible to be quite strategic about which topics you could afford not to study
12	It was often hard to discover what was expected of me in this course
13	On this course it was necessary to work consistently hard to meet the assessment requirements
14	Too often the staff asked me questions just about facts
15	I didn't understand some of the feedback on my work
16	Whatever feedback I received on my work came too late to be useful
17	The way the assessment worked on this course you had to study every topic
18	To do well on this course all you really needed was a good memory
These questions are about the way you go about your learning on the course:	
19	When I'm reading I try to memorise important facts which may come in useful later
20	I usually set out to understand thoroughly the meaning of what I am asked to read
21	I generally put a lot of effort into trying to understand things which initially seem difficult
22	I often found myself questioning things that I heard in classes or read in books
23	I find I have to concentrate on memorising a good deal of what we have to learn
24	Often I found I had to study things without having a chance to really understand them
Learning from the exam (only to be completed if there were exams on the course)	
25	Doing exams brought things together for me
26	I learnt new things while preparing for the exams
27	I understood things better as a result of the exams
Overall satisfaction	
28	Overall I was satisfied with the quality of this course
Please write comments about your experience of assessment on the degree programme and how it has affe	

sagreement ssessment components	strongly agree	agree	?	disagree	strongly disagree

cted your studying:

Scoring the AEQ

Work out scale scores by calculating average student ratings for each question (sa = 5, a = 4, ? = 3, d = 2, sd = 1). Some questions have this scoring the other way round (i.e. sd = 5, marked 'negative' scoring' below). Each question is part of a scale so then calculate the average scores for all questions in each scale. Compare your scale scores with the typical range of scores for each scale which can be found in Gibbs & Dunbar-Goddet (2007).

Quantity of effort

6. The way the assessment worked you had to put the hours in regularly every week
13. On this course it was necessary to work consistently hard to meet the assessment requirements

Coverage of syllabus

4. You had to study the entire syllabus to do well in the assessment
5 . The assessment system made it possible to be quite selective about what parts of courses you studied (Negative scoring)
11. It was possible to be quite strategic about which topics you could afford not to study (Negative scoring)
17. The way the assessment worked on this course you had to study every topic

Quantity and quality of feedback

3. I received hardly any feedback on my work (Negative scoring)
15. I didn't understand some of the feedback on my work (Negative scoring)
16. Whatever feedback I received on my work came too late to be useful (Negative scoring)

Use of feedback

1. I used the feedback I received to go back over what I had done in my work
2. The feedback I received prompted me to go back over material covered in the course
8. I paid careful attention to feedback on my work and tried to understand what it was saying

Appropriate assessment

10 The staff seemed more interested in testing what I had memorised than what I understood (Negative scoring)
14. Too often the staff asked me questions just about facts (Negative scoring)
18. To do well on this course all you really needed was a good memory (Negative scoring)

Clear goals and standards

7. It was always easy to know the standard of work expected
9. The teachers made it clear from the start what they expected from students
12. It was often hard to discover what was expected of me in this course (Negative scoring)

Surface approach

19. When I'm reading I try to memorise important facts which may come in useful later
23. I find I have to concentrate on memorising a good deal of what we have to learn
24. Often I found I had to study things without having a chance to really understand them

Deep approach

20. I usually set out to understand thoroughly the meaning of what I am asked to read
21. I generally put a lot of effort into trying to understand things which initially seem difficult
22. I often found myself questioning things that I heard in classes or read in books

Learning from the examination

25. Doing the exams brings things together for me
26. I learn new things while preparing for the exams
27. I understand things better as a result of the exams

Satisfaction

28. Overall I am satisfied with the teaching on this course.

References

Angelo, T.A. & Cross, K.P. (1993) *Classroom Assessment Techniques: a handbook for college teachers.* San Francisco, Ca: Jossey-Bass.

Asghar, A. (in press) Reciprocal Peer Coaching as a Formative Assessment Strategy for First Year Students. *Assessment and Evaluation in Higher Education.*

Baird, L.L. (1985) Do grades and tests predict adult accomplishment? *Research in Higher Education* 23(1), 3-85.

Beaty, L., Gibbs, G. & Morgan, A. (1997) Learning orientations and study contracts. In: Marton, F., Hounsell, D. & Entwistle, N. (eds) *The Experience of Learning.* Edinburgh: Scottish Academic Press.

Berliner, D.C. (1984) The half-full glass: a review of research on teaching. In: Hosford, P.L. (ed) *Using What We Know About Teaching.* Alexandria, Va: Association for Supervision and Curriculum Development.

Biggs, J.B. & Collis K.F. (1982) *Evaluating the Quality of Learning: the SOLO taxonomy.* New York: Academic Press.

Black, P. & Wiliam, D. (1998) Assessment and classroom learning. *Assessment in Education* 5 (1), 7-74.

Bridges, P., Cooper, A., Evanson, P., Haines, C., Jenkins, D., Scurry, D., Woolf, H. & Yorke, M. (2002) Coursework marks high, examination marks low: discuss. *Assessment and Evaluation in Higher Education* 27(1), 36-48.

Brittingham, B.E. (1988) Undergraduate students' use of time: a classroom investigation. *To Improve the Academy* 7, 45-52.

Brookhart, S. M. (2001) Successful students' formative and summative uses of assessment information. *Assessment in Education: Principles, Policy and Practice* 8(2), 153-169.

Bruner, J.S. (1974) *Toward a Theory of Instruction.* Cambridge, Mass: Harvard University Press.

Carless, D.M. (2002) The 'mini-viva' as a tool to enhance assessment for learning. *Assessment and Evaluation in Higher Education* 27(4), 353-363.

Carroll, M. (1995) Formative assessment workshops: feedback sessions for large classes. *Biomedical Education* 23(2), 65-67.

Chansarkar, B.A. & Raut-Roy, U. (1987) Student performance under different assessment situations. *Assessment and Evaluation in Higher Education* 12(2), 115-122.

Chickering, A.W. & Gamson, Z.F. (1987) *Seven Principles to Good Practice in Undergraduate Education.* Racine, Wi: The Johnson Foundation Inc.

Chickering, A.W. & Gamson, Z.F. (1991) *Applying the Seven Principles to Good Practice in Undergraduate Education.* San Francisco: Jossey-Bass.

Cole, S., Coates, M. & Lentell, H. (1986) Towards good teaching by correspondence. *Open Learning* 1(1), 16-22.

Conway, M.A., Cohen, G. & Stanhope, N. (1992) Why is it that university grades do not predict very long-term retention? *Journal of Experimental Psychology: General* 121(3), 49-57.

Cook, A. (2001) Assessing the use of flexible assessment. *Assessment and Evaluation in Higher Education* 26(6), 539-549.

Cooper, N.J. (2000) Facilitating learning from formative feedback in level 3 assessment. Assessment and Evaluation in Higher Education 25(3), 279-291.

Crooks, T.J. (1988) The impact of classroom evaluation practices on students. *Review of Educational Research* 58(4), 438-481.

Curtis, S. & Shami, N. (2002) The effect of taking paid employment during termtime on students' academic studies. *Journal of Further and Higher Education* 26(2), 129-138.

Ding, L. (1998) Revisiting assessment and learning: implications of students' perspectives on assessment feedback. Paper presented to Scottish Educational Research Association Annual Conference, University of Dundee, 25-26 September.

Dochy, F., Segers, M. & Sluijsmans, D. (1999) The use of self-, peer- and coassessment: a review. *Studies in Higher Education* 24(3), 331-350.

Dochy, F., Segers, M., van den Bossche, P. & Gijbels, D. (2003) Effects of problem-based learning: a meta-analysis. *Learning and Instruction* 13(5), 533-568.

Dunbar-Goddet, H. & Gibbs, G. (in press) A research tool for evaluating the effects of programme assessment environments on student learning: the Assessment Experience Questionnaire (AEQ). *Assessment and Evaluation in Higher Education.*

Dunkin, M.J. (1986) Research on Teaching in Higher Education. In: Wittrock, M.C. (ed) *Handbook of Research on Teaching* (3rd edn). New York: Macmillan.

Entwistle, N. & Entwistle, A. (2004) Contrasting forms of understanding for degree examinations: the student experience and its implications. *Higher Education* 22(3), 205-227.

EUROSTUDENT (2007) Available at: **www.his.de/abt2/ab21/Eurostudent/**

Falchikov, N. & Goldfinch, J. (2000) Student Peer Assessment in Higher Education: a meta-analysis comparing peer and teacher marks. *Review of Educational Research* 70(3), 287-322.

Forbes, D. & Spence, J. (1991) An experiment in assessment for a large class. In: Smith, R. (ed) *Innovations in Engineering Education.* London: Ellis Horwood.

Gagne, R.M. (1977) *The Conditions of Learning* (3rd edn). New York: Holt, Rinehart and Winston.

Gardiner, L.F. (1997) Redesigning higher education: producing dramatic gains in student learning. *ASHE-ERIC Higher Education Reports* 23(7). Washington, DC: Association for the Study of Higher Education.

Gibbs, G. (1992) *Assessing More Students.* Oxford: Oxford Centre for Staff Development.

Gibbs, G. (1995) *Assessing Student Centred Courses.* Oxford: Oxford Centre for Staff Development.

Gibbs, G. (1999) Using assessment strategically to change the way students learn. In: Brown, S. & Glasner, A. (eds) *Assessment Matters in Higher Education.* Buckingham: Society for Research into Higher Education and Open University Press.

Gibbs, G. & Dunbar-Goddet, H. (2007) The effects of programme assessment environments on student learning. York: The Higher Education Academy. Available at: **http://www.heacademy.ac.uk/assets/York/documents/ourwork/research/gibbs_0506.pdf**

Gibbs, G. & Dunbar-Goddet, H. (in press) Characterising programme level assessment environments. *Assessment and Evaluation in Higher Education.*

Gibbs, G. & Lucas, L. (1987) Coursework assessment, class size and student performance: 1984-94. *Journal of Further and Higher Education* 21(2), 183-192.

Gibbs, G. & Simpson, C. (2004) Conditions under which assessment supports student learning. *Learning and Teaching in Higher Education* 1, 3-31.

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

Greer, L. (2001) Does changing the method of assessment of a module improve the performance of a student? *Assessment and Evaluation in Higher Education* 26(2), 128-138.

Hattie, J.A. (1987) Identifying the salient facets of a model of student learning: a synthesis of meta-analyses. *International Journal of Educational Research* 11, 187-212.

Hattie, J. & Timperley, H. (2007) The power of feedback. *Review of Educational Research* 77(1), 81-112.

Higgins, R., Hartley, P. & Skelton, A. (2001) Getting the message across: the problem of communicating assessment feedback. *Teaching in Higher Education* 6(2), 269-274.

Higher Education Policy Institute (2007) *The academic experience of students in English Universities* (2007 Report). Oxford: HEPI.

Hounsell, D. (1987) Essay writing and the quality of feedback. In: Richardson, J.T.E., Eysenck, M.W. & Warren-Piper, D. (eds) *Student Learning: research in education and cognitive psychology.* Milton Keynes: Open University Press and Society for Research into Higher Education.

Hutchings, P., Marchese, T., & Wright, B. (1991) *Using Assessment to Strengthen General Education.* Washington, DC: American Association for Higher Education.

Hyland, F. (2001) Providing effective support: investigating feedback to distance language learners. *Open Learning* 16(3), 231-247.

Innis, K. (1996) *Diary Survey: how undergraduate full-time students spend their time*. Leeds: Leeds Metropolitan University.

Jackson, M. (1995) *Making the grade: the formative evaluation of essays.* UtiliBASE. Available at: **http://ultibase.rmit.edu.au/Articles/dec96/jacks1.htm**

Kember, D., Ng, S., Tse, H., Wong, E.T.T. & Pomfret, M. (1996) An examination of the interrelationships between workload, study time, learning approaches and academic outcomes. *Studies in Higher Education* 21(3), 347-358.

Kerr, N.L. & Bruun, S.E. (1983) Dispensability of member effort and group motivation losses: free-rider effects. *Journal of Personality and Social Psychology* 44, 78-94.

Kniveton, B.H. (1996) Student perceptions of assessment methods. *Assessment and Evaluation in Higher Education* 21(3), 229-238.

Kulik, C.C., Kulik, J.A. & Cohen, P.A. (1980) Instructional technology and college teaching. *Teaching Psychology* 7, 199-205.

Latane, B., Williams, K. & Harkins, S. (1979) The causes and consequences of social loafing. *Journal of Personality and Social Psychology* 37(6), 822-832.

Lea, M. & Street, B. (1998) Student writing in higher education: an academic literacies approach. *Studies in Higher Education* 23(2), 157-172.

Macdonald, J. (2002) 'Getting it together and being put on the spot': synopsis, motivation and examination. *Studies in Higher Education* 27(3), 329-338.

Maclellen, E. (2001) Assessment for learning: the different perceptions of tutors and students. *Assessment and Evaluation in Higher Education* 26(4), 307-318.

MacFarlane, B. (1992) The 'Thatcherite' generation of university degree results. *Journal of Further and Higher Education* 16, 60-70.

Marton, F. & Wenestam, C.G. (1978) Qualitative differences in the understanding and retention of the main points in some texts based on the principle-example structure. In: Gruneberg, M.M., Morris, P.E. & Sykes, R.N. (eds) *Practical Aspects of Memory.* London: Academic Press.

McKeachie, W.J., Pintrich, P.R., Lin, Y. & Smith, D. (1986) *Teaching and Learning in the College Classroom: a review of the research literature.* Ann Arbor: National Centre for Research to Improve Postsecondary Teaching and Learning, University of Michigan.

Miller, C.M.I. & Parlett, M. (1974) *Up to the Mark: a study of the examination game.* Guildford: Society for Research into Higher Education.

Moffat, M. (1989) *Coming of Age in New Jersey: college and American culture.* New Brunswick, NJ: Rutgers University Press.

Nicol, D. (2006) Increasing success in first year courses: assessment re-design, self-regulation and learning technologies. Paper presented to ASCILITE Conference, Sydney, December 2006. Available at: **http://tltt.strath.ac.uk/REAP/public/Papers/DNicol_Ascilite_26oct06.pdf**

Northedge, A. (2003) Enabling Participation in Academic Discourse. *Teaching in Higher Education* 8(2), 169-180.

O'Donovan, B., Price, M. & Rust, C. (2008) Developing student understanding of assessment standards: a nested hierarchy of approaches. *Teaching in Higher Education* 13(2), 205-217.

Orsmond, P., Merry, S. & Reiling, K. (2002) The use of exemplars and formative feedback when using student derived marking criteria in peer and self assessment. *Assessment and Evaluation in Higher Education* 27(4), 309-323.

Paton-Saltzberg, R. & Lindsay, R. (1993) *The Effects of Paid Employment on the Academic Performance of Full-time Students in Higher Education.* Oxford: Oxford Polytechnic.

Penny, A.J. & Grover, C. (1996) An analysis of student grade expectations and marker consistency. *Assessment and Evaluation in Higher Education* 21(2), 173-184.

Perry, W.G. (1970) *Forms of Intellectual and Ethical Development in the College Years.* New York: Holy, Rinehart and Winston.

Price, M. & Rust, C. (1999) The experience of introducing a common criteria assessment grid across an academic department. *Quality in Higher Education* 5(2), 133-144.

Ramsden, P. (1991) A performance indicator of teaching quality in higher education: the Course Experience Questionnaire. *Studies in Higher Education* 16, 129-150.

Ramsden, P. (1992) *Learning to teach in Higher Education.* London: Routledge.

Ramsden, P., Beswick, D. & Bowden, J. (1987) Learning processes and learning skills. In: Richardson, J.T.E., Eysenck, M.W. & Warren-Piper, D.W. (eds) *Student Learning: research in education and cognitive psychology.* Milton Keynes: Open University Press and Society for Research into Higher Education.

Rice, M., Mousley, J. & Davis, R. (1994) Improving student feedback in distance education: a research report. In: Evans, T. & Murphy, D. (eds) *Research in Distance Education 3: revised papers from the third research in distance education conference.* Geelong, Victoria: Deakin University Press.

Roberts, D. (1996) Feedback on assignments. *Distance Education* 17(1), 95-116.

Rowntree, D. (1987) *Assessing Students – how shall we know them?* London: Kogan Page.

Rust, C., Price, M. & O'Donovan, B. (2003) Improving students' learning by developing their understanding of assessment criteria and processes. *Teaching in Higher Education* 28(2), 147-164.

Sadler, D.R. (1989) Formative assessment and the design of instructional systems. *Instructional Science* 18, 119-144.

Sadler, D.R. (1998) Formative assessment: revisiting the territory. *Assessment in Education* 5(1), 77-84.

Säljö, R. (1981) Learning approach and outcome: some empirical observations. *Instructional Science* 10(1), 47-65.

Säljö, R. (1982) *Learning and Understanding.* Göteborg: Acta Universitatis Gothoburgensis.

Sambell, K. & McDowell, L. (1998) The construction of the hidden curriculum: messages and meanings in the assessment of student learning. *Assessment and Evaluation in Higher Education* 23(4), 391-402.

Schunk, D. (1984) Self-efficacy perspective on achievement behaviour. *Educational Psychologist* 19, 48-58.

Schunk, D. (1985) Self-efficacy and classroom learning. *Psychology in the Schools* 22, 208-223.

Scouler, K. (1998) The influence of assessment method on students' learning approaches: multiple choice question examinations vs. essay assignment. *Higher Education* 35, 453-472.

Scouler, K. & Prosser, M. (1994) Students' experiences of studying for multiple choice question examinations. *Studies in Higher Education* 19(3), 267-280.

Sly, L. (1999) Practice tests as formative assessment improve student performance on computer managed learning assessments. *Assessment and Evaluation in Higher Education* 24(3), 339-344.

Snyder, B.R. (1971) *The Hidden Curriculum.* Cambridge, Ma: MIT Press.

Springer, L., Stanne, M.E., & Donovan, S.S. (1999) Effects of small-group learning on undergraduates in Science, Mathematics, Engineering, and Technology: a meta-analysis. *Review of Educational Research* 69(1), 21-51.

Starr, J.W. (1970) Student opinion on methods of assessment. *Educational Review* 22, 243-253.

Steadman, M. (1998) Using classroom assessment to change both teaching and learning. *New Directions for Teaching and Learning* 75, 23-35.

Stephenson, K., Sander, P. & Naylor, P. (1996) Student perceptions of the tutor's role in distance learning. *Open Learning* 11(1), 22-30.

Tang, C. (1994) Effects of modes of assessment on students' preparation strategies. In: Gibbs, G. (ed) *Improving Student Learning: theory and practice.* Oxford: Oxford Centre for Staff Development, 151-170.

Taras, M. (2001) The use of tutor feedback and student self-assessment in summative assessment: towards transparency for students and for tutors. *Assessment and Evaluation in Higher Education* 26(6), 605-614.

Thomas, J.W., Iventosh, L. & Rohwer, W.D. (1987) Relationships among student characteristics, study activities and achievement as a function of course characteristics. *Contemporary Educational Psychology* 12, 344-364.

Tinto, V. (1993) *Leaving College: rethinking the causes and cures of student attrition* (2nd edn). Chicago: University of Chicago Press.

Tynjala, P. (1998) Traditional studying for examination vs constructivist learning tasks: do learning outcomes differ? *Studies in Higher Education* 23(2), 173-191.

Vos, P. (1991) Curriculum control of learning processes in higher education. Paper presented to 13th International Forum on Higher Education of the European Association for Institutional Research, Edinburgh.

Warren, J.R. (1971) *College Grading Practices: an overview, Report No 9.* Washington, DC: ERIC Clearinghouse on Higher Education.

Williams, K., Harkins, S.G. & Latane, B. (1981) Identifiability as a deterrent to social loafing: two cheering experiments. *Journal of Personality and Social Psychology* 40(2), 303-311.

Wootton, S. (2002) Encouraging learning or measuring failure? *Teaching in Higher Education* 7(3), 353-357.

Wojtas, O. (1998) Feedback? No, just give us the answers. *Times Higher Education Supplement,* 25 September 1998.

Yorke, M. (2001) Formative assessment and its relevance to retention. *Higher Education Research and Development* 20(2), 115-126.

About the author

Graham Gibbs is currently an Honorary Professor at the University of Winchester, working on a project to improve student learning through changing the assessment of entire degree programmes and the quality assurance frameworks within which they operate. He has been a Professor at Oxford Polytechnic, the Open University and the University of Oxford, from where he retired in 2007. He has been writing practical guidance about assessment since the 1980s. He has undertaken large-scale research studies into the way assessment affects student learning over the past 20 years. He founded the International Improving Student Learning Symposium, now in its 20th year, and the International Consortium for Educational Development in Higher Education, which now contains organisations from over 30 countries. He has been awarded Honorary Doctorates by Sheffield Hallam University for the development of university teaching in the UK, and by the University of Utrecht for his leadership of the improvement of university teaching world-wide.

Leeds Met Press
Leeds Metropolitan University
Old School Board
Calverley Street
Leeds LS1 3ED
Tel: 0113 812 9257
Email: publications@leedsmet.ac.uk

www.leedsmet.ac.uk/publications

£12.00
ISBN 978-1-907240-06-5
© Leeds Metropolitan University 2010