

From: *Innovative Assessment in Higher Education* (2006) Ed.

Bryan C. & Klegg K., Routledge, Abingdon, with permission from the Editors.)

Chap 2. How assessment frames student learning

Graham Gibbs

Introduction

Students are strategic as never before, and they allocate their time and focus their attention on what they believe will be assessed and what they believe will gain good grades. Assessment frames learning, creates learning activity and orients all aspects of learning behaviour. In many courses it has more impact on learning than does teaching. Testing can be reliable, and even valid, and yet measure only the trivial and distorted learning which is an inevitable consequence of the nature of the testing. This chapter is not about testing but about how assessment leads to effective study activity and worthwhile learning outcomes. It starts by quoting students describing how they respond to perceived assessment demands. It then outlines eleven 'conditions under which assessment supports learning'. These conditions are based on a review of theoretical literature on formative assessment and on a review of published accounts of successful innovations in assessment, across all discipline areas, undertaken in order to identify why they were successful. Economical assessment methods are described that meet these conditions,

each based on published evidence of worthwhile impact on learning and student performance. Associated diagnostic tools have been developed to help faculty to identify how their students respond to their assessment regime, and some uses of these tools will be described. The chapter is intended to provide a conceptual underpinning to the innovations in assessment described elsewhere in this volume.

Students' experience of assessment

The two most influential books I read at the start of my teaching career were from parallel studies of a very similar nature on opposite sides of the Atlantic, focussing on very similar phenomena. In the US Benson Snyder was undertaking an ethnographic study of the experience of students at MIT. He had not intended to focus on assessment but he discovered that assessment completely dominated student experience and so that is what he wrote most about. 'The Hidden Curriculum' (Snyder, 1971) described the way students strategically negotiated their way through impossibly large curricula, trying to work out what faculty were really after and what they could safely ignore.

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

"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."

The central idea in Snyder's work was the gap between the course as presented publicly in course documentation and by faculty, and the narrower and rather different course students experienced and actually studied. The shape and size of this narrower curriculum was determined by students' perceptions of assessment demands. Studying was an exercise in selective negligence. In Sweden, Fransson (1977) reported how students who were unable to understand or work out what to study, and attempted to study everything, quickly became depressed by the impossibility of the task. After initially working diligently, the number of hours they studied each week declined and they eventually performed badly or dropped out. There are few rewards for students who are not strategic.

At the same time as Benson Snyder was being astonished by students at MIT, studies at the University of Edinburgh, an ancient research-intensive university in Scotland, found exactly the same phenomenon, when they interviewed students:

"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 and Parlett, 1974)

This study, 'Up to the mark: a study of the examination game', described some students as 'cue conscious'. These students were aware of cues about what to study and what to neglect. Others were described as 'cue seekers' and took their professors for a beer in the hope of finding out what questions would be on the exam paper. The remainder were described as 'cue deaf' and no matter how often they were advised what to focus on, this information passed over their heads. It proved easy to predict students' grades simply by categorising the extent to which they were tuned in to cues about assessment, and neglected the right things.

Subsequently in my own research I have often glimpsed the world of the student in relation to perceived assessment demands. The following student on Masters course in Oceanography was only too aware of the gap between his learning and what got him good grades:

"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, 1992a, p101)

At Oxford University assessment for grades is almost entirely separated from learning, and from 'assessment for learning'. Most assessment for learning takes place orally in tutorials, which are often weekly one-to-one (or very small group) meetings between a tutor (an academic or a or graduate student) and an undergraduate student. The work the student has been doing (on average 10-14 hours of reading and writing to produce an essay) is discussed. Metaphorically, or in some cases actually, the tutor and student sit on the same side of the table and explore the subject matter presented in the essay together as a joint scholarly exercise. They 'do anthropology' or 'do history' together. Assessment here is all 'formative' and is designed to support learning. Students may even complain after a tutorial that they still do not know how they are getting on, as essays are not usually graded. Exams happen mainly at the end of three years. The tutor has little or no input into the design of the exam questions and is not supposed to be preparing the student for exams. This gives students considerable freedom to explore what confounds them in the subject matter and tutors considerable freedom to explore students' misconceptions. I am not trying to sell the tutorial method: in any case tutorials are cripplingly expensive even for Oxford! But there is a phenomenon that occurs often in other universities but which happens less commonly at Oxford: that of 'faking good'. Faking good is an attempt by a student to present themselves and their work as if they know and understand more than they actually do, for the purpose of maximising grades. Students in

most institutions normally choose those essay questions that they know most about and that they will need to do the least learning for, not those that will result in most learning. I remember, to my shame, that in my undergraduate essays I sometimes cited more references than I had actually read. In the example below an engineering student (not from Oxford!) describes the way he presents his 'problem sheets' to his tutor for marking, not in a way which reveals his difficulties of understanding or the blind alleys he went down as he tackled the problems, but in a way which is designed to trick the tutor into giving a good grade. Here assessment is a hurdle to be negotiated, a game to be played, 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, 1992a)

This is the opposite of an Oxford student choosing to spend most time on what they do not yet understand or a tutor deliberately choosing to discuss what the student does not yet understand fully. 'Faking good' is a direct consequence of the form of assessment.

Students' experience of feedback

It is a truism that learning requires feedback. The importance of feedback is enshrined in the 'Seven principles of good practice in undergraduate education' (Chickering and Gamson, 1991) and is developed by Nicol and Milligan (Chapter 5). But how do students experience feedback? A number of studies have found that students can find feedback incomprehensible, that they glance at the mark and then throw their work away, or even that they do not bother to collect their work from the Departmental office (e.g. Higgins et al 2000; Hounsell, 1987). In interviews I encountered the following statement that was representative of common student perceptions. It concerns another of the 'Seven Principles': that feedback has to be provided promptly if it is to be attended to and be useful.

"The feedback on my assignments comes back so slowly that we are already on the topic after next and I've already submitted the next assignment. It's water under the bridge, really. I just look at the mark and bin it"

The crucial variable appears not to be the quality of the feedback (which is what teachers tend to focus on) but the quality of student engagement with that feedback. For example Forbes and Spence (1991) report a study of innovation in assessment in an Engineering course where peer feedback and marks, of very mixed quality and uncertain marking standards, provided instantly during lecture classes, produced a truly dramatic increase in student performance (in subsequent exams) compared with the previously high quality teacher feedback and reliable marking which came back slowly and which students as a consequence had not attended to.

This second example of a student statement concerns a general problem with feedback associated with objective testing – including computer based multiple choice question testing and open entry forms of computerised feedback. The following student was studying a ‘maths for science’ course where the assessment was on-line. Students could tackle maths assignments in their own time and then type in their maths solutions. A very sophisticated computer programme then generated instant and appropriate qualitative feedback.

“I do not like the on-line assessment method...it was too easy to only study to answer the questions and still get a good mark ... the wrong reasoning can still result in the right answer so the student can be misled into thinking she understands something ... I think there should

have been a tutor-marked assessment part way through the course so someone could comment on methods of working, layout etc.”

This problem with a focus of assessment on outcome rather than process is echoed in reviews of the impact of different kinds of feedback on pupil behaviour in schools. (Black and Wiliam, 1998). It is now clear that feedback without marks leads to better learning than marks only, or even than marks with feedback. Any feedback that focuses on individual's overall performance (in the form of a mark or grade) rather than on their learning, detracts from learning.

Who makes the judgements?

Thirty years ago I was struck by Carl Rogers' principle of learning that stated that learning is maximised when judgements by the learner (in the form of self-assessment) are emphasised and judgements by the teacher are minimised (Rogers, 1969). At the time it seemed a noble but hopelessly idealistic and impractical notion. I now know better. Much research on self and peer assessment appears to be obsessed with the reliability of student marking in the hope that student-generated grades can substitute for teachers' grades and save the teacher a whole lot of work. If you go to enough trouble students are indeed capable of reliable marking (or, rather, as reliable as the rather low level teachers usually achieve). But this completely misses the point. What is required is not more grades but more learning. The value of self and peer assessment is that students internalise academic

standards and are subsequently able to supervise themselves as they study and write and solve problems, in relation to these standards. It is the act of students making judgement against standards that brings educational benefits, not the act of receiving a grade from a peer. This issue is explored in much greater depth in Chapter 5 by Nicol and Milligan . There are now many studies of the positive impact of self and peer assessment on student performance. In the US this has been associated with the 'Classroom Assessment' initiative. In Europe and Australia this has been associated with less organised, but no less voluminous attempts by teachers to support learning better through changing assessment. My favourite example comes from a Psychology department where the teachers were exhausted by spending every weekend marking experimental and laboratory reports. They would provide feedback such as "You have not labelled the axes of your graphs", week in, week out, despite abundant guidance to students on lab report writing and repeated feedback of an identical kind. The teachers suspected that their diligence in providing feedback was to little purpose. They devised a feedback sheet which contained about fifty of the most frequent comments they wrote on students' reports (such as 'Have not labelled axes of graphs'). Next to each was a 'tick box' and they provided feedback in the form of ticks next to comments. While this saved their wrist from repetitive strain injury from writing the same feedback endlessly it did not improve students' lab reports. They then had a brainwave and gave the students the feedback sheet and required them to attach a copy to the front of each laboratory report they submitted, but with a tick next to all the things they had done wrong. Students were then able to submit technically perfect lab reports because

they could undertake useful self-assessment before submission, and the teachers had to develop new, tougher, criteria in order to avoid everyone getting perfect grades. It is not until students apply criteria and standards to judge their own work, as part of self-supervision while working (just as I am doing while writing this chapter) that their work will improve. And this is at no cost to the teacher (or in my case the book's editors).

Conditions under which assessment supports learning

I have written a number of books over the years about assessment methods, with the intention of increasing teachers' repertoire of alternatives to suit different contexts, and because variety is the spice of life (cf. Gibbs, 1992b; Gibbs, 1995; Habeshaw et al, 1993). What I had not done was to provide a coherent rationale for deciding which kind of method suited which kind of context or educational problem. I have recently set out to turn observations such as those in the previous sections into a coherent rationale (Gibbs, 1999). This involved reading theoretical literature (mostly schools-based) about formative assessment. But most importantly I read large numbers of 'case study' accounts of changed assessment set in higher education where claims were made about improved student performance – but where there was usually no explanation of what this improvement was due to. For example in Forbes and Spence (1991) there is a full description of the assessment innovation and full data about the improvement in grades but no articulation of the underlying pedagogic principles involved. I was interested in what 'pedagogic work' was being done by various assessment tactics that resulted

in them being effective. In the case studies it was also rare to find a rationale for selecting the particular innovation the authors chose to implement. I was interested in how you could diagnose a problem so as to guide the choice of an appropriate assessment solution.

This literature review led to the articulation of eleven 'conditions under which assessment supports student learning' (Gibbs and Simpson, 2004). A student questionnaire was then developed, the 'Assessment Experience Questionnaire (AEQ)', (Gibbs and Simpson, 2003; available for free use at <http://www.open.ac.uk/science/fdtl/tools.htm>) which has been used widely to diagnose which of these conditions is being met and which are not. In the UK, South Africa and Hong Kong there is currently quite widespread use of the AEQ as part of action research projects undertaken by science faculty in order to find ways to support student learning better through innovation in assessment. Scores from the AEQ help to diagnose problems and select appropriate assessment solutions, and then the AEQ is being administered again after the innovation has been implemented, to monitor changes to student learning behaviour. A national project is documenting some of these projects (<http://www.open.ac.uk/science/fdtl/>). These eleven conditions are summarised here clustered under the headings used to structure the questionnaire.

Quantity and distribution of student effort

1. Assessed tasks capture sufficient study time and effort

This condition concerns whether your students study sufficiently out of class or whether the assessment system allows them to get away with not studying very much at all. This is the 'time on task' principle (Chickering and Gamson 1991) linked to the insight that it is assessment, and not teaching, that captures student effort.

2. These tasks distribute student effort evenly across topics and weeks

This condition is concerned with whether students can 'question spot' and avoid much of the curriculum, or stop turning up to class after the last assignment is due in. It is about evenness of effort week by week across a course and also across topics. I once saw data on the distribution of students' answers for an examination in which students had to answer three of fifteen questions. Almost everyone answered the same three questions and the topics addressed by the other twelve questions were presumably hardly studied at all.

Quality and level of student effort

3. These tasks engage students in productive learning activity

This condition is partly about whether the assessment results in students taking a deep approach (attempting to make sense) or a surface approach (trying to reproduce) (Marton et al, 1997) and also about quality of engagement in general. Do the things students have to do in order to meet

assessment requirements engender appropriate, engaged and productive learning activity? Examinations may induce integration of previously unconnected knowledge, during revision, or memorisation of unprocessed information. Which approach to revision will be induced depends not so much on the examination demands as on students' perceptions of these demands.

4. Assessment communicates clear and high expectations to students

This condition is again drawn from Chickering and Gamson (1991): 'Good practice communicated high expectations'. This is partly about articulating explicit goals that students understand and can orient themselves towards, and partly about the level of perceived challenge. Can students spot, within ten minutes of the first class of a course or within the first thirty seconds reading a course description, that this is going to be an easy course and that assessment demands will be able to be met without much effort or difficulty? Where do students pick up these clues from? Without internalising the standards of a course students cannot monitor their own level of performance or know when they have not yet done enough to be able safely to move on to the next task or topic or to reallocate their scarce time to another course they are studying in parallel. On the Course Experience Questionnaire, scores on the 'Clear Goals and Standards' scale correlate with the extent to which students take a deep approach to learning (Ramsden, 1991).

The remaining conditions concern feedback. They are not elaborated here as feedback is addressed in depth in Chapter 6.

Quantity and timing of feedback

5. Sufficient feedback is provided, both often enough and in enough detail
6. The feedback is provided quickly enough to be useful to students

Quality of feedback

7. Feedback focuses on learning rather than on marks or students themselves
8. Feedback is linked to the purpose of the assignment and to criteria
9. Feedback is understandable to students, given their sophistication

Student response to feedback

10. Feedback is received by students and attended to
11. Feedback is acted upon by students to improve their work or their learning

Outline ideas for meeting these conditions are summarised later in this chapter and addressed in more detail in subsequent case studies.

Use of the AEQ to diagnose where to innovate

Evidence from the use of the 'Assessment Experience Questionnaire' (Gibbs et al 2003) is cited here to illustrate the way it can be used to diagnose problems with the way assessment supports students learning and in particular the extent to which the eleven conditions outlined above are met.

This data comes from 776 students on fifteen science courses at two UK

universities. The students at the two universities were revealed to have very different perceptions of their assessment systems. In fact there was more variation between the universities than between courses, suggesting that there are institutional assessment system cultures or norms. In response to data such as that in Table 1, Institution B has focussed its efforts on improving feedback to students. Institution A has, in contrast, focussed its efforts on students making more use of the high volume of feedback that they are given. This data comes from a national scale project ('Formative Assessment in Science Teaching' <http://www.open.ac.uk/science/fdtl/>) that is supporting action research into the way assessment supports learning. The 'scale scores' in Table 1 are out of a maximum score of 30 and are derived from five point rating scales on each of six questionnaire items making up each scale. The differences between these institutions in terms of the 'Quantity and timing of feedback' are very marked.

Table 1 Comparison of fifteen Science courses at two universities in terms of the reported volume and distribution of student effort, and students' perception of the quantity and promptness of feedback

	University A	University B		
Scale	Scale Score		t	p [1]
Time demands and distribution of student effort	20.3 (s.d. 3.16)	18.6 (s.d. 2.91)	7.387 (d.f. 772)	p < 0.001
Quantity and timing of	22.0	15.6	19.28	p < 0.001

feedback	(s.d. 4.40)	(s.d. 4.48)	(d.f. 766)	
Sample items	% agree or strongly agree			
<i>I only study things that are going to be covered in the assignments</i>	8%	27%		
	% disagree or strongly disagree			
<i>On this course it is possible to do quite well without studying much</i>	64%	33%		
Sample items	% agree			
<i>On this course I get plenty of feedback on how I am doing</i>	68%	26%		
<i>Whatever feedback I get comes too late to be useful</i>	11%	42%		

[1] two-tailed t-test.

The data also showed marked differences between different courses in the extent to which, for example, students found feedback helpful, or acted upon feedback. Table 2 examines differences between courses within institution 'A' and displays a selection of data from the 'best' and 'worst' course (in terms of scores on the AEQ) in the sample. The data show that the AEQ is capable of distinguishing between courses even within a single institution within a single subject. Note just how unlikely it is for students to be prompted by feedback to

go back over material. What is clear from such data is that there are major differences in how effectively assessment systems work to support student learning and to foster student behaviour that is likely to lead to learning. There is clearly plenty of scope for using methods that improve matters.

Table 2 Comparison of Science courses within University A in terms of students' use of feedback

AEQ items	'Best' course	'Worst' course
	% strongly agree	
<i>The feedback helps me to understand things better</i>	36%	6%
<i>The feedback shows me how to do better next time</i>	31%	4%
<i>The feedback prompts me to go back over material covered earlier in the course</i>	13%	1%

Assessment tactics that solve learning problems

The section summarises assessment tactics that, from accounts in the literature, have the capacity to address particular conditions well. There is

obviously no one-to-one relationship between their use and changed student learning behaviour – that will depend on an interaction of many variables in each unique context.

Addressing problems with the quantity and distribution of student effort

It is possible to capture student time and effort simply by using more assignments or assignments distributed more evenly across the course and across topics. The Open university, for example, traditionally employs eight evenly spaced assignments on each 'full credit' course, to ensure that distance learning students work steadily throughout the year and on all course units.

To cope with the consequent marking load it is possible to make the completion of assignments a course requirement, or a condition to be met before a summative assessment is tackled at a later date, without marking any of these assignments. It is also possible to sample assignments for marking (e.g. from a portfolio) such that students have to pay serious attention to every assignment in case they are selected for marking. Mechanised and computer-based assessment can obviously achieve similar ends (of high levels of assessment without tutor marking), though often without meeting the other conditions very fully and sometimes at the cost of quality of learning and mis-orienting of learning effort. The use of self and/or peer assessment (provided that it is required) can also generate student time on task without generating teacher time on marking. It is also possible to design examinations

that make demands that are unpredictable, or which sample almost everything, so that students have to study everything just in case, though this too can result in other conditions not being met, such as failing to generate high quality and level of learning effort through students taking a surface approach as a result of anxiety and excessive workload.

Problems with the quality and level of student effort

Assignments that are larger, more complex and open-ended, requiring 'performances of understanding', are more likely to induce a deep approach to study than are short answer tests or multiple choice questions. Assignments involving interaction and collaboration with other students, in or out of class are usually more engaging. Social pressures to deliver, for example through making the products of learning public (in posters, or through peer assessment) may induce more care and pride in work than 'secretly' submitted assignments to the teacher. Clear specification of goals, criteria and standards, and especially the 'modelling' of the desired products, for example through discussion of exemplars, will make it less likely that students will set themselves low or inappropriate standards. If students internalise these goals, criteria and standards, for example through student marking exercises, and public critique of work, they are likely to be able to use these standards to supervise their own study in future.

Problems with the quantity and timing of feedback

Regular feedback requires regular assignments, ideally starting early in a course so that students are oriented to the standard required as early as possible. Some institutions or departments have quality standards for the volume and turn-around time of tutor feedback and also have the means to monitor the achievement of these standards. Mechanised feedback can be used to increase its volume at low cost and much contemporary innovation in assessment is concerned with computer-generated marking and feedback, where mechanised tests are used. The challenge then is to meet other conditions at the same time. The quality of feedback can be traded off against speed of return: for example using peer feedback or model answers, or the tutor sampling students' assignments to produce generic feedback based on the first five assignments assessed, but not reading the rest. The balance of gains and losses from such practices are a matter for empirical study to explore. Ultimately the fastest and most frequent feedback available is that provided by students to themselves from moment to moment as they study or write assignments in 'learning conversations', and investing effort in developing such self-supervision may be much the most cost-effective use of tutors' time.

Problems with the quality of feedback

If students receive feedback without marks or grades, they are more likely to read the feedback as the only indication they have of how they are getting on. This has been demonstrated to have a significant positive impact on learning outcomes (Black and Wiliam, 1998). If feedback is structured around the

goals of the assignment, and relates clearly to criteria and standards, this is more likely to result in clarity and impact than unstructured arbitrary feedback that focuses on student characteristics.

The quality and impact of feedback can be improved through clear briefing to teachers. The Open University in the UK trains all its tutors in how to give thorough and motivating feedback, and also periodically monitors the quality of this feedback (rather than monitoring the quality of their teaching) and provides individualised coaching where there are perceived to be quality problems. Students' ability to make sense of and use feedback can be improved through classroom discussion of what specific examples of feedback mean and through discussion of improvements students intend to make to subsequent assignments in response to the feedback.

Problems with students' response to feedback

If feedback is provided faster, there is more likelihood that students will read and respond to it. If students tell the teacher what they would like feedback on, they are more likely to pay attention to this feedback when they receive it (Habeshaw et al). If students discuss feedback on their assignments, in class, they are more likely to think about it and take it seriously (Rust et al , 2003). If students receive feedback on a draft of an assignment they are likely to use this feedback to improve the assignment. If the assignment allows drafting, with feedback on the draft, students are likely to make good use of this feedback. Students are usually capable of giving each other useful feedback

on drafts. Tutors can also ask students to submit a cover sheet to their assignment explaining how the peer feedback (or the feedback from the tutor on the previous assignment) was used to improve this assignment. If testing is 'two-stage' with an opportunity between a 'mock' test and the 'real' test to undertake more studying on those aspects which were not tackled well in the 'mock' test, then they are likely to use this opportunity (Cooper, 2000). If assignments are 'multi-stage', with each stage building up towards a larger and more complete final assignment, then students will almost inevitably use feedback as they put the whole thing together. If assignments have multiple components, tackled in sequence (e.g. stages of a project, elements of a portfolio of evidence) in which each assignment contributes to a larger whole, feedback on early sections are very likely to be used by students to improve the whole. If students are asked to demonstrate how the current assignment benefits from feedback on the previous assignment, and allocated marks for how well they have done this, they are likely to take feedback seriously. If at least some of the feedback is generic in nature there is more likelihood that it will also apply to subsequent assignments on different topics.

Conclusions

This chapter has provided a conceptual framework for diagnosing the extent to which assessment regimes are likely to create effective learning environments. The 'conditions under which assessment supports student learning' are based on educational theory and empirical evidence concerning either weak student performance where these conditions are not met, or

improved student performance where innovations in assessment have been introduced specifically to address one or more of these conditions. It is clear both that student learning can be poor largely because the assessment system does not work well, and that changes just to the assessment, leaving the teaching unchanged, can bring marked improvements. The chapter has provided a glimpse of the wide variations that exist between courses and between institutions, even within the same discipline area, in terms of how well assessment regimes support learning. Finally the chapter has outlined some of the ways in which these assessment conditions can be met – the tactics that can be adopted to address specific weaknesses. Later sections of this book contain case studies that illustrate some of these tactics in action.

References

Black, P. & William, D. (1998) *Assessment and Classroom Learning*.

Assessment in Education: Principles, Policy and Practice, 5,1, pp7-74.

Brown, E., Gibbs, G. & Glover, C. (2003) Evaluation tools for investigating the impact of assessment regimes on student learning. *Bioscience Education e-Journal*, 2, <http://www.bioscience.heacademy.ac.uk/journal/vol2/beej-2-5.htm>

Chickering, Z.F. & Gamson, A.W. (1991) *Applying the seven principles for good practice in undergraduate education*. San Francisco: Jossey-Bass.

Cooper, N.J. (2000) *Facilitating learning from formative feedback in level 3*

assessment, *Assessment and Evaluation in Higher Education*, vol. 25, no. 3, pp279-291.

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

Fransson, A (1977) On qualitative differences in learning. IV - effects of motivation and test anxiety on process and outcome. *British Journal of Educational Psychology*, 47, pp244-257.

Gibbs, G. (1992a) *Improving the quality of student learning*. Bristol: Technical and Educational Services.

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

Gibbs, G. (1992) *Assessing student-centred courses*. Oxford: Oxford Centre for Staff Development.

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

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

Gibbs, G. and Simpson, C. (2004) Conditions Under Which Assessment Supports Students' Learning. *Learning and Teaching in Higher Education*, 1, pp3-31

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

Habeshaw, S., Gibbs, G. & Habeshaw, T. (1993) *53 interesting ways to assess your students*. (3rd Edn). Bristol: Technical and Educational Services.

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

Marton, F. (1997) Approaches to Learning. In Marton, F., Hounsell, D. & Entwistle, N. (Eds.) *The Experience of Learning*. Scottish Academic Press.

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

Ramsden, P. (1991) A performance indicator of teaching quality in higher education: the course experience questionnaire. *Studies in Higher Education*, 16, pp129-150.

Rogers, C. (1969) *Freedom to Learn*. Columbus: Merrill.

Rust, C., Price, M. & O'Donovan, B. (2003) Improving Students' Learning by Developing their Understanding of Assessment Criteria and Processes *Assessment & Evaluation in Higher Education*, 28,2, pp 147-164.

Scouler, K.M. & Prosser, M. (1994) Students' experience of studying for multiple choice question examinations. *Studies in Higher Education*, 19, 3, pp267-279.

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