Effie Maclellan

Assessment for learning: the differing perceptions of tutors and students

Article
(Refereed)

Assessment for Learning: the differing perceptions of tutors and students

INTRODUCTION

The power of assessment to determine the quality of learning has been established for quite
some time (Ramsden, 1997) with the evidence clearly concluding that the quality of student
learning is as high (or as low) as the cognitive demand level of the assessment tasks (Crooks,
1988; Gibbs, 1999). In other words, if students perceive a need to understand the material in
order to successfully negotiate the assessment task, they will engage in deep learning but if they
perceive the assessment instrument to require rote learning of information, they will be unlikely
to engage with the higher level objectives which may well have been intended by the
programme of study. While a powerful determinant of learning outcome, students' experiences
of assessment do not occur in a vacuum but are contextualised in their overall perceptions of the
goals they have to achieve, the workload they carry, the teaching they experience and the
autonomy they have to direct their own learning. (Prosser & Trigwell, 1999). More than that,
however, since students' perceptions of their learning environment will *vary* within any one
group or class of students (Prosser & Miller, 1989), so the quality of learning will vary, not only
through perceptions of the learning environment but also through what the students do or don't
do in response to their perceptions. Assessment practices, then, play a subtle, complex, and
enormously important role in the students' experiences of learning.

Assessment itself, however, is undergoing a paradigm shift (Gipps, 1994) with a movement
from the measurement model to a standards model (Taylor, 1994). Such a movement comes
from an increasing recognition that the assumptions of traditional learning theory are now very
questionable. Learning is now more commonly recognised as a process of knowledge
construction (rather than of knowledge reproduction), as being situated in particular contexts
(and therefore not necessarily transferable to other contexts) and as being knowledge dependent
(Resnick, 1989). The realisation that learning is not linear and atomistic and that it is not
decontextualised, has led to the desire that assessment should represent meaningful, significant
and worthwhile forms of human endeavour and accomplishment. In other words assessment
tasks should reflect the ways in which knowledge and skills are used in real world contexts
(Newmann & Archbald, 1992).

But while moving to a standards model of assessment may be seen as desirable, it is not automatic since the assumptions which underpin the measurement and standards models are very different. The measurement model is concerned with the relative, reliable performance of individuals on decontextualised, standardised tasks which are deemed to be valid indicators of the domain being assessed. This model seeks to emphasise individual differences. The standards model, on the other hand, is concerned with the level to which knowledge is embedded in deep (and possibly new) understanding, and can be demonstrated in authentic tasks. This model seeks to emphasise the value of education as a means of promoting the development of individuals (Taylor, 1994).

The study reported here had the aim of trying to discern what model underpinned extant in-Faculty assessment practices. It was assumed that the model could be inferred from descriptions of assessment practices as these were perceived by students and staff most immediately involved in assessment.

METHOD

Since the purpose of the study was to describe the student and staff perceptions of assessment, an actuarial survey was deemed appropriate. Data on perceptions of in-Faculty assessment were collected from staff and students by means of a questionnaire. While a questionnaire can be rightly criticised for yielding only superficial information, the questionnaire in this study was seen as being an efficient use of time, as allowing academic peers and students to respond anonymously and as gleaning hitherto unknown information on a range of practices which ultimately shape the assessment 'events' that staff and students experience. Of primary importance were views on the extent to which assessment is concerned to enable learning rather than merely measure learning. Since assessment which is explicitly designed to promote learning is probably one of the most powerful tools that we have in higher education, it was considered important to solicit views on *why* assessment was taking place and on *how useful* the assessment process was. Also of importance were views on how judgements are made since educational assessment is essentially a matter of making valid judgements about the incidence of learning, so *who* makes the judgement, *when* the judgements are made and what the *rules or procedures* are for making the judgements are important matters on which to gain some clarity.

Assessment for Learning: the differing perceptions of tutors and students

Finally it was of importance to capture views on the learning being assessed. What is considered important to assess will strongly determine what is considered important to learn. Moreover, since any mode of assessment has its inherent limitations the mode which is used will underscore the actual learning that is being focused on. It was therefore important to be aware of the *assessment instruments* used and of the *cognitive demand level* of the learning. All of these issues were incorporated into the questionnaire which was derived by the author but which was influenced by the theoretical ideas of Bowden & Marton, (1998,) Biggs (1999) and Prosser & Trigwell, (1999) together with McDowell's (1998) account of old and new assessment practices. The 40 item questionnaire were clustered in 8 variables:

the purpose (4 items), content (6 items), timing (4 items), mode (9 items) and marking of assessment (10 items), feedback on assessment (4 items), the assessor (2 items) and the identity of the participant (1 item). All items, with the exception of the item on participant identity, reflected the range of assessment issues which are documented in the literature. Small scale piloting of the questionnaire was conducted with 5 members of Faculty staff and 5 undergraduate students.

A 4-point itemised rating scale of 'frequently', 'sometimes', 'never', 'don't know' was used for collecting responses. Faculty Staff members were issued with the questionnaire forms and asked to complete them within 10 days. The students were issued with the questionnaire during class time at the start of a module on assessment. Eighty members of Faculty Staff (80% return) and 130 3rd year undergraduates (100% return) completed a 40-item questionnaire on their experiences of assessment. Faculty Staff members were all experienced in teaching and assessment practices. The undergraduates were following a B.Ed. (Hons.) Programme and were selected because:

   a)   they were a sufficiently large and collectively accessible sample;

   b)   they represented a degree course which is a major part of the Faculty's teaching;

   c)   they were deemed to have had a range of assessment experiences within the Faculty.

Assessment for Learning: the differing perceptions of tutors and students

Participants were asked to think about their experiences to date of assessment practices **within the Faculty** and to draw on the totality of that experience when responding to the questions. Staff and students were asked not to include the assessment of field experience/placement in their considerations.

RESULTS

Summary of Questionnaire Findings

The essential purpose of this study was to describe assessment practices as these were perceived by staff and students. Each item in the questionnaire was endorsed in terms of the frequency with which the participant had experienced the practice under consideration. The modal preferences are summarised below.

INSERT Table 1 about here

The most frequently endorsed purpose of assessment, as perceived by both students and staff, was to grade/rank student achievement. A second important purpose for staff was the motivation of learning though the students largely perceived assessment to be only sometimes motivating, with 25% of the students claiming that assessment was never motivating. The diagnosis of strengths and weaknesses was seen as a frequent purpose of assessment by 41% of students and 66% of staff. Staff were divided on the degree to which assessment was used to evaluate teaching, with 41%% saying that this happened frequently and 40% saying 'sometimes'. Students, however, were much less convinced that assessment was frequently used to evaluate teaching.

INSERT Table 2 about here

Amongst staff, assessment was frequently used to judge the development of knowledge (81%) and the application of knowledge (76%). Assessment events were frequently designed to assess students' ability to analyse, synthesise and evaluate information (86%, 76% and 79% respectively). Students too (but to a lesser extent than staff) acknowledged that these aspects of development were assessed frequently. This is consistent with Norton's (1990) finding that

Assessment for Learning: the differing perceptions of tutors and students

students put more emphasis on demonstrating knowledge than on constructing an argued position in response to the set task.

INSERT Table 3 about here

Self and peer assessments were not frequent occurrences. Indeed the staff respondents were fairly evenly divided on whether self and peer assessments never happen (41% endorsing self assessment and 48% endorsing peer assessment) or happen sometimes (50% endorsing self assessment and 48% endorsing peer assessment).

INSERT Table 4 about here

Seventy-seven percent (77%) of students and 73% of staff reported that students were never assessed at the start of a module. Ninety-seven percent (97%) of student responses and 86% of staff responses indicated that assessment typically occurred at the end of a module. Only 12% of students and 24% of staff reported that students' views of their own readiness to be assessed had a bearing on the timing.

INSERT Table 5 about here

The most frequent mode of assessment was the essay (95% of students and 79% of staff). Other forms of written assessment were variously perceived. Eighty percent (80%) of students thought multiple choice questions were sometimes used while 60% of the staff thought the self same mode was never used. Seventy-seven percent (77%) of students thought case/fieldwork notes were frequently assessed while 41% of staff thought these were never assessed. Short answer questions (76% of students and 45% of staff) and reflective logs/diaries (51% of students and 58% of staff) were used sometimes.

In the non-written mode students perceived assessment to be sometimes through seminar contributions, presentation to peers and participation in workshops/labs (75%, 88% and 65% respectively) and while staff respondents also acknowledged these modes of assessment (48%, 59% and 38% respectively) 40% of staff said seminar contributions were never assessed and

5

53% said participation in workshops/labs was never assessed. Ninety-one percent (91%) of students and 63% of staff reported that students were never assessed through audio/video recordings.

INSERT Table 6 about here

Amongst staff, marking was most frequently against explicit criteria (81%) and focused on knowledge (60%), thinking (64%) and presentation (54%). As a result of marking a summative grade was frequently awarded, according to 85% of the staff. Only in 18% of the staff sample was there the view that assessed work is marked by second assessors as a matter of routine although this number increased to 50% of the sample when considering 'fails'. Alarmingly, perhaps, 18% of the staff reported that the second marking of 'fails' never happened. Internal consistency was checked frequently (38%) or sometimes (44%) according to respondents and while work for assessment was frequently submitted anonymously according to 39% of the staff another 43% maintained that anonymous marking never happened.

Amongst students, marking was also most frequently against explicit criteria (54%) although 50% of the students thought that marking was sometimes against implicit criteria. Marking sometimes focused on knowledge (60%), thinking (57%) and presentation (55%) and as a result a summative grade was frequently awarded according to 59% of the students. Sixty-one percent (61%) of students thought that work is sometimes marked by a second assessor while 25% of the students did not know if 'failing' work is marked by anyone else. Internal consistency was sometimes checked (64%) and almost half of the sample (49%) reported that anonymous marking never happened.

INSERT Table 7 about here

The majority of staff respondents considered feedback to be helpful in its detail (93%) and to improve learning (94%) either sometimes or frequently. Most of the students too recognised the value of feedback although they didn't consider the value to be as frequent as staff claimed. Differences in student and staff perceptions were noted in the role played by feedback in

6

prompting discussion between staff and student(s) and in enabling students to understand assessment.

INSERT Table 8 about here

Although the modal responses summarised in Table 1 might suggest a fair degree of correspondence in the perceptions of staff and students, this apparent agreement between staff and students is more likely to be an artefact of the somewhat crude rating scale. Statistically, there were considerable differences between staff and students. On a Mann-Whitney U Test, perceptions of assessment were significantly different on 32 out of the 39 items.

INSERT Table 9 about here.

DISCUSSION

In discussing assessment practices as these were perceived by staff and students, there will be an attempt to deduce the respective 'views' of the two groups. These views will be in very general terms since the measuring instrument could access fairly blunt perceptions only.

The staff view of assessment

For staff the primary purpose of assessment was to grade or rank students, but the more developmental purposes of motivating students, diagnosing learning and evaluating teaching were not discounted. The importance given to grading/ranking achievement (confirmed in a number of the items concerned with the marking of students' work) is perhaps not misplaced given the need for universities to be able to establish students' levels of achievements and to communicate these with professional regulating bodies and potential employers (Atkins et al, 1993). That staff perceive the developmental function of assessment to be important seems to be corroborated in the valuable role they see being played by feedback in strengthening the students' knowledge base, in developing student thinking and in improving presentation. However, the importance allegedly placed on the developmental or formative function of assessment is not internally consistent with other views endorsed by staff. For example, staff reported that assessment neither took place at the beginning of a module nor could students be assessed when they themselves felt ready. Furthermore, staff reported that self and peer

assessment were infrequent occurrences. There are at least three important educational implications arising from this. Firstly, the practice of not assessing at the start of a module precludes the opportunity to modify/design teaching in response to student understanding (Prosser & Trigwell, 1999). Secondly, the practice of not allowing students to be assessed when they feel ready for assessment denies that students may need differential amounts of time to achieve desired learning outcomes (Boud, 1995). Thirdly, to discount students' judgements is to fail to appreciate that effective learning is in large measure a function of strategic metacognitive behaviour (Biggs, 1999). All three practices, which discount the status of the students' learning, are inconsistent with a constructivist view of learning and, by extension, with a standards model of assessment.

Amongst staff, assessment was frequently used to judge the development and application of knowledge together with the skills of analysis, synthesis and the evaluation of information. This finding is reflected in the perceptions of both the content and marking of assessment. That declarative knowledge per se was not the sole focus of assessment suggests that students were being assessed on their ability to assemble and interpret information, formulate ideas, construct a defensible argument and critique a line of reasoning, an emphasis which Norton (1990) also found among staff. Being able to engage in these particular types of cognitive tasks is important because they have the potential to be generalisable to other learning and problem solving situations in the real world (Messick, 1994). In other words in being assessed on their ability to engage in these various cognitive tasks the students were being required to demonstrate authentic academic achievement (Newmann & Archbald, 1992). However, the extent to which assessment genuinely focused on students' capacity to apply, transform or evaluate the relevance of declarative knowledge in new situations could be viewed as questionable when considering the processes through which the assessment information was gathered. Initial perusal of the different 'instruments' suggested some to more obviously mirror real life than others. For example tutorial contributions and oral presentations might be seen as manifestations of the practical, real world skills of group interaction and problem solving (Taylor, 1997). Similarly reflective diaries are promoted as evidencing continuing professional development (Brockbank & McGill, 1998). Finally the creation of videotapes, the maintenance

of case notes and participation in labs/workshops can directly map on to real life activity. The extent to which these real life modes of assessment were used was, by staff's own admission, not frequent. Conversely, while the essay mode and short answer mode may well assess the use of cogent argument and the expression of complex ideas, writing about one's ideas is somewhat removed from actively demonstrating one's knowledge. The extent to which assessment tasks made authentic demands of students is then questionable. Staff believed that they were assessing a full range of learning but the heavy emphasis on one particular mode suggests a more limited range of learning was actually being assessed. This would not be fully consistent with the standards model.

The student view of assessment

Students agreed that a frequent purpose of assessment was to make a summative judgement in the form of grading or ranking student performance. However, students claimed that formative purposes were served only some of the time. For example, most students did not view feedback on their learning as either routinely helpful in itself or as a catalyst for discussion. That students primarily perceived assessment to be about judging levels of achievement rather than about enabling learning may be partly a function of what formative assessment can mean. Its most essential meaning is that it provides the teacher with information which can be fed back into the teaching/learning process (Crooks, 1988; Gipps, 1994). However, a conception of formative assessment which focuses on the teacher's role but discounts that of the learner is increasingly being understood as incomplete. Sadler (1989) noted that even when learners are given valid and reliable information about the quality of their work, there is no necessary improvement in the work. Sadler's analysis has led to the realisation that assessment can only have a formative influence if learners are involved in the process (Pryor & Torrance, 1996, Tunstall & Gipps, 1996; Wiliam & Black, 1996). The implication of this is that if students are not actually monitoring and regulating the quality of their own learning, feedback of itself, regardless of its degree of detail, will not cause improvement in learning. That students did not view assessment as offering them opportunities within which to advance their own learning is further evidenced in the perception by almost 80% of the students that assessment was frequently or sometimes carried out using implicit criteria. The dangers of this do not need to

be exaggerated. If they believe the criteria are implicit, students will, by definition, be unclear as to what to do to achieve the desired standard. Further if students believe the criteria to be implicit, then they may see assessment as some sort of lottery in which they experience inequable treatment from idiosyncratic staff. Such a perception is not impossible given the subjectivity of staff in the marking process (Norton, 1990). In not recognizing their own role in formative assessment, students see staff as having the power to determine either the veracity of student performance per se or the validity of the evidence from which performance is inferred. Such a view is consistent with the measurement model of assessment.

Like staff, students perceived a range of learning to be assessed although they were of the view that the full range was assessed only some of the time rather than frequently. Unlike staff, the students considered this learning to be assessed through various modes. Although the essay mode was the most frequently endorsed, seminar contributions, presentations to peers, case/fieldwork notes and multiple-choice questions were acknowledged by large numbers of students as being modes of assessment which they experienced. Quite at odds with the staff perception, students seemed to perceive assessment as being ubiquitous. Furthermore, that large numbers of students saw themselves engaging in self assessment at least some of the time while 41% of the staff denied student self assessment, suggests that there was a lack of shared meaning between students and staff as to the status of assessment. This is perhaps an example of the general finding that there are varied conceptions of learning and teaching amongst university staff and student populations (Prosser & Trigwell, 1999) and may be partly attributed to the students' lack of appreciation of the distinction between learning *activities* and learning *goals* (Dwyer, 1998) and partly due to a failure to *align* learning goals, learning activities and the assessment of learning outcomes (Biggs, 1999). Given that (albeit small numbers of) students were unaware of marking and moderation procedures and given that the students appear not to have any real grasp of the power or value of formative assessment, it does seem likely that the student conception of assessment was somewhat primitive. This is not to suggest that the student view was wrong, merely that it was underdeveloped. But this is perhaps not surprising since, by dint of professional power, the staff view of assessment will influence the students' attitudes towards, and perceptions of, assessment (Sadler, 1998). And since the staff

view of assessment did not fully espouse the philosophy of the standards model, thereby presenting a somewhat confusing picture of assessment, it should not be surprising that the student view of assessment was somewhat incoherent.

CONCLUSION

The standards model of assessment is the desirable model in formal education (Biggs, 1999) because it attempts to reflect *what has been learned* in criterion referenced terms. However, the historical dominance of the measurement model together with the difficulties of implementing assessment which is premised on a standards model means that extant practices in educational assessment may not be consistent with a standards model (Taylor, 1994). In this study such inconsistency was indeed evidenced. Staff declared a commitment to the formative purposes of assessment but engaged in practices which militated against formative assessment being fully realised. Similarly staff maintained that the full range of learning was frequently assessed yet the dominant mode of assessment was the traditional, academic essay, thereby attenuating the idea that students were engaging in authentic assessment which could enhance their learning. In other words the staff view suggested aspirations towards the standards model of assessment but these aspirations have yet to be fully realised.

Overall the student view of assessment is a depressing one. The students do not exploit assessment to improve their learning and, furthermore, appear to have a very underdeveloped conception of what assessment is. Given that assessment practices may or may not precipitate powerful or transformative learning it seems important to appreciate the central involvement of students themselves in the assessment process. Such understanding may well have to be developed in staff who in turn could identify the changes which they need to make in their practice in order to help students take greater responsibility for their own learning. Only when all assessment tasks can be fully authentic and only when staff and students can put the students' learning at the very centre of the educational enterprise, can the assessment practices be consistent with the standards model.

REFERENCES

Atkins, M., Beattie, J. & Dockrell, W. (1993) *Assessment Issues in Higher Education*, (Great Britain, Employment Department).

Assessment for Learning: the differing perceptions of tutors and students

Biggs, J. (1999) *Teaching for Quality Learning at University,* (Buckingham, The Society for Research into Higher Education &The Open University Press).

Boud, D. (1995) *Enhancing Learning through Self Assessment*, (London, Kogan Page).

Bowden, J. & Marton, F. (1998) *The University of Learning,* (London, Kogan Page).

Brockbank, A. & McGill, I. (1998) *Facilitating Reflective Learning in Higher Education,* (Buckingham, The Society for Research into Higher Education &The Open University Press).

Crooks, T. (1988) The impact of classroom evaluation practices on students, *Review of Educational Research,* 58 (4), pp. 438-81.

Dwyer, C. (1998) Assessment and classroom learning: theory and practice, *Assessment in Education*, 5(1), pp. 131-137

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

Gipps, C. (1994) *Beyond Testing,* (London, The Falmer Press).

McDowell, L. (1998) Editorial, *Assessment and Evaluation in Higher Education*, 23(4), pp. 335-338.

Messick, S. (1994) The interplay of evidence and consequences in the validation of performance assessments, *Educational Researcher*, 23(2), pp. 13-23.

Newmann, F. & Archbald, D. (1992) The nature of authentic academic achievement, in: H. Berlak, F. Newmann, E. Adams, D. Archbald, T.Burgess, J. Raven, & T. Romberg (Eds) *Towards a New Science of Educational Testing and Assessment,* (New York, State University of New York Press).

Norton, L. (1990) Essay-writing: what really counts? *Higher Education,* 20, pp411-442.

Prosser, M. & Miller, R. (1989) The 'how' and 'what' of learning physics, *European Journal of Psychology of Education*, 4(4), pp. 513-528.

Prosser, M. & Trigwell, K. (1999) *Understanding Learning and Teaching*, (Buckingham, The Society for Research into Higher Education &The Open University Press).

Pryor, J. & Torrance, H. (1996) Teacher-pupil interaction in formative assessment: assessing the work or protecting the child? *The Curriculum Journal*, 7(2), pp.205-226.

Ramsden, P. (1997) The context of learning in academic departments, in: F. Marton, D. Hounsell & N. Entwistle (Eds) *The Experience of Learning*, pp.198-216 (Edinburgh, Scottish Academic Press).

Resnick, L. (1989) Introduction, in: L. Resnick (Ed) *Knowing Learning and Instruction*, (New Jersey, Lawrence Erlbaum Associates).

Sadler, R. (1989) Formative assessment and the design of instructional systems, *Instructional Science*, 18, pp. 119-44.

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

Taylor, C. (1994) Assessment for measurement or standards: the peril and the promise of large scale assessment reform, *American Educational Research Journal*, 31(2), pp. 231-62.

Taylor, I. (1997) *Developing Learning in Professional Education* (Buckingham, The Society for Research into Higher Education &The Open University Press).

Tunstall, P. & Gipps, C. (1996) 'How does your teacher help you to make your work better?' Children's understanding of formative assessment, *The Curriculum Journal*, 7(2), pp. 185-203

Wiliam, D & Black, P. (1996) Meanings and consequences: a basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, 22(5), pp. 537-548.