

chapter 8

ASSESSING: STUDENT ASSESSMENT

If we wish to discover the truth about an educational system we must look into its assessment procedure (Rowntree, 1987: 1).

As we develop through the markers, the examiners are expecting more of you, to show that you've understood the material. They're not just wanting the facts regurgitated.... We can't go on preparing for exams the way we have been (engineering student cited in Thomson and Falchikov, 1998: 382).

In this chapter we explore assessment, beginning by looking at its critical relationship to learning. We then review a number of key aspects of assessment, including issues of reliability and validity, differences between formative and summative assessment and contrasting systems of assessment, such as norm-referenced and criterion-referenced. A wide range of assessment methods are discussed in relationship to the different dimension of the critical matrix.

INTRODUCTION

Assessing students is perhaps the most emotionally sensitive part of teaching. It is intellectually demanding for teachers and can be socially disturbing and divisive for students. Students may easily feel that it is not just their learning being assessed, but their developing identity as

persons. Associations with right and wrong can trigger the more primitive associations with good and bad, creating fear and loss of confidence. This may be particularly palpable in subjects such as mathematics, where being wrong can be painfully obvious.

Thus, whether assessment appears to be valid is no trivial issue. Students must feel that they have received the best opportunity to demonstrate their ability in a course, and that they have had the chance to convey something of themselves and what the subject means to them. Otherwise, students may associate assessment and marking with a system of control. With some students this may produce either conformity or alienation, while other students may see it as a rejection of what they have to offer as mature adults. For these reasons, the whole experience of assessment must be matched with what the course is trying to achieve and the culture it is trying to create. Certainly, assessment and course design must be integrated with key issues of student learning.

In this chapter, we first address a number of general concepts and issues pertaining to assessment. We then explore assessment as it relates to the intellectual, personal, social and practical dimensions of learning within the critical matrix. Throughout, we discuss a wide range of methods designed to aid in the development of a balanced assessment strategy for effectively gauging student achievement and enhancing learning.

DEFINITIONS

While there is some variation in the common use of the term ‘assessment’, particularly in conjunction with the term ‘evaluation’, in this book we understand assessment to mean the gathering of information about student learning, which may be qualitative or quantitative in nature and used for some purpose. These purposes include providing feedback to students about their progress and development, assigning grades and making instructional or pedagogical choices (Brookhart, 2004). As such, assessment includes, but transcends, the simple measurement of student achievement.

We distinguish assessment from ‘evaluation’, which forms the substance of the next chapter. We reserve the term ‘evaluation’ for making informed judgements on the value or quality of an educational entity – including such things as programmes, courses, sessions, teachers, books, etc. – aimed at fostering student learning. Assessment of student learning thus plays a key role in evaluation, but they are distinct.

Table 8.1 *SOLO and assessment*

Level	Pre-structural	Uni-structural	Multi-structural	Relational	Extended abstract
<i>Associated verbs</i>	Misses the point	Identify, do simple procedure	Enumerate; describe, list, combine, do algorithms	Compare/contrast, explain, analyse, relate, apply	Theorize, generalize, reflect, hypothesize
<i>Example of question in each level</i>	Question will be irrelevant or nonsensical	'What is the capital of France?'	'Identify the major powers in the First World War and outline the key foreign policy associated with each'	'Explain how the government's nutritional requirements concerning food groups have changed over time and discuss how different social groups have been affected by these changes'	'Design an outreach programme that will help educate a diverse community about the government's new nutritional guidelines'
<i>Example of response</i>	Misses the point and/or is irrelevant	Focuses on only one relevant conceptual issue in a complex case	Focuses on more than one issue but they may be a disorganized collection of items; may reproduce a 'shopping list'	Shows understanding, applies or uses a concept(s) which integrates a collection of data, issues, etc.	Goes beyond existing principles; higher-order principles are used to bring in a new broader set of issues

Source: Adapted from Biggs, 2003

THE CENTRALITY OF LEARNING OUTCOMES

Learning is a critical aspect of assessment. For assessment to be effective and meaningful, it must be clearly aligned with the learning objectives and teaching methods and activities employed in the learning context. As we discussed in Chapter 3, Bigg's SOLO taxonomy provides a useful framework for designing both learning outcomes and assessments (Biggs, 2003). With five stages of rising complexity (pre-structural, uni-structural, multi-structural, relational and extended abstract), SOLO also offers a systematic way to demonstrate how a learner's performance grows as tasks grow more complex (see Table 8.1).

To a large degree, Bigg's SOLO maps on to Bloom's taxonomy, with the movement from lower-order to higher-order thinking, in complexity and level of skill. Unlike Bloom's taxonomy, however, Bigg's SOLO also offers a means to assess learning outcomes, exploring the breadth and depth of learning. SOLO can be used by instructors to elicit the critical thinking and learning they want from their students.

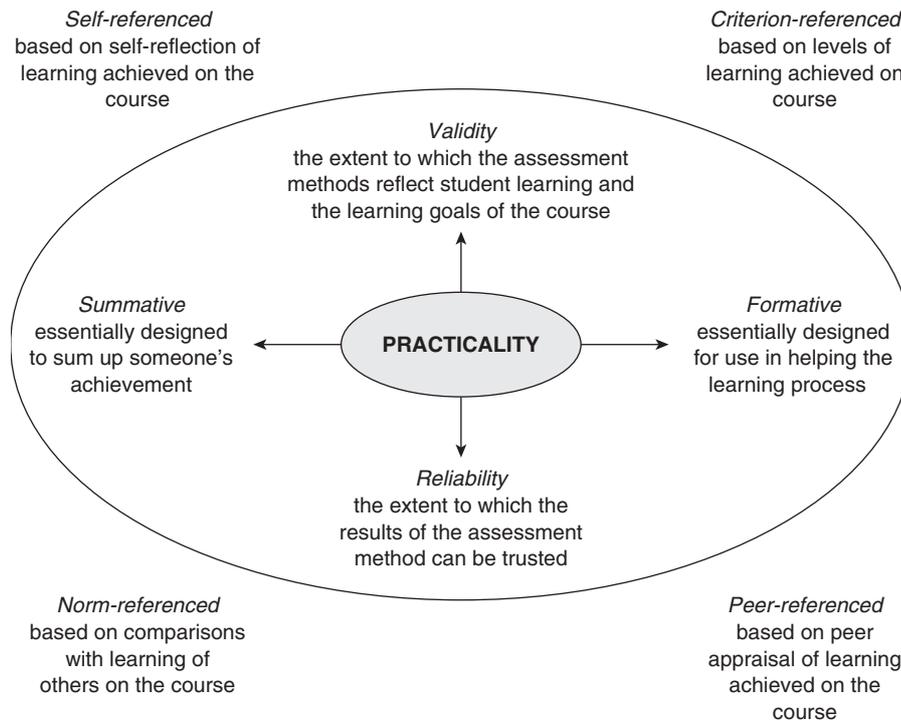


Figure 8.1 *Aspects of assessment of student learning*

KEY ASPECTS OF ASSESSMENT

There are several key aspects of assessment, each with its own set of tensions (see Figure 8.1). Whether the assessment is intended to be formative or summative, where the point of reference is (instructor-referenced or peer or self-referenced) and how valid and reliable the assessment is (or perceived to be) are all crucial in assessing learning. Practicality is at the heart of all tensions as there are usually significant practical implications limiting the full achievement of these aspects.

Formative and summative assessment

The first tension focuses on the overall purpose of assessment. Two general purposes of assessment – enhancing learning and measuring student achievement – are often referred to as formative and summative assessment, respectively. Simply, formative assessment concerns development, improvement and learning, while summative assessment concerns accountability and performance. Formative assessment, for example, can be used to:

- provide feedback to improve learning;
- motivate students;
- diagnose a student's strengths and weaknesses; or
- help students reflect critically on their own learning.

Summative assessment may be used to:

- pass or fail a student;
- grade or rank a student; or
- predict a student's success in other courses or employment.

Although different methods of assessment may fulfil one or the other of these two purposes – an exam may be summative, while a first draft of an essay is formative – almost all methods can be applied to either of these two purposes. For example, Wininger (2005) describes how he teaches his students to study their graded (summative) exams to improve their learning and methods of studying in a formative way. Neither should the two purposes be regarded as mutually exclusive. 'Enhancing learning', as we saw in Chapter 2, could mean either deep (transforming ideas of learning) or surface (reproducing ideas of learning). Similarly 'measuring achievement' could suggest a measurement of either deep or surface learning. 'What' we assess becomes more important than 'why' we are assessing.

This distinction is important, particularly if we agree with Rowntree (1987: 1) that 'to discover the truth about an education system we must look to its assessment procedures'. Working purely for marks or grades can indicate a cynical and purely strategic approach to assessment. Traditional finals, for example, have often elicited surface or strategic approaches, especially in terms of course content, enabling students to evade the harder or deeper demands of their courses and still pass.

If, however, assessment goes beyond basic information and techniques to include higher-level intellectual demands and reflects a commitment and enthusiasm for education and the deeper values of the discipline, then 'high grades' will reflect the deeper issues that the course may really be about. In this respect, 'high grades' might reflect the students' appreciation of alternative perspectives, their recognition of the need to change and to reconceptualize and construct knowledge, and their capacity to cope with unexpected and new complexities. In addition, college offers many valuable experiences that should neither be incorporated into grades nor marginalized by them, experiences which are more appropriate to formative assessment.

Classroom assessment

Classroom assessment provides a relatively simple form of formative assessment for teachers. It is a process of critical inquiry and analysis that instructors use to assess student learning with an eye on the impact of their teaching. It is, moreover, learner-centred, teacher-directed, mutually beneficial, formative, context-specific, ongoing and rooted in good teaching practice (Angelo and Cross, 1993). As such, more instructors than ever have begun to use classroom assessment better to understand the learning that transpires in their classrooms.

Rather than simply assuming that students are learning what they are teaching, instructors can use classroom assessment to assess their students closely as they learn, collecting frequent feedback and asking critical questions to gauge student comprehension, skill acquisition, application of materials and knowledge construction. Several common classroom assessment techniques (CATs) are discussed more fully throughout the chapter in terms of the critical matrix.

Reliability and validity

The second tension considers the nature of reliability and validity. Reliability refers to the extent to which the results of the assessment method can be trusted, whereas validity refers to the meaning and value of the assessment, and the extent to which the assessment methods reflect student learning and the learning goals of the course. Lack of reliability or consistency of marking and grading, both perceived and real, has contributed to students' lack of confidence in marking systems, particularly as wide discrepancies can occur in individual cases (Montero et al., 2006). Indeed, as Pieron observed in the early 1960s, 'assessment by different examiners produces marks with considerable variability such that in the determination of these marks the part played by the examiner can be greater than that of the performance of the examinee' (1963: 140). Even today, academics may still discuss marks as if they were accurate and absolute.

Although marking can sometimes be reasonably reliable, accuracy depends on the hypothetical mark that might be obtained if the marks given by a large number of markers were averaged. This does not mean that there is an absolute, true and accurate mark; rather, that the mean of one large group of examiners will tend to agree with the mean of another large group of examiners.

There are similar problems associated with a lack of validity, particularly if students cannot clearly see or understand the connection to the rest of

the course or if the learning objectives and teaching approaches are not clearly and constructively aligned. Improving reliability and validity, then, is crucial. Both may be improved by encouraging teachers to reflect more fully on their approaches and interpretations of marking, developing thoughtful criteria based on intended learning outcomes; using appropriate teaching activities to meet those outcomes; and communicating those criteria to students ahead of time (Brookhart, 2004). A few methods follow.

Double marking

In this activity, two examiners rate the same work, without knowing the other's mark. Each marker is, in effect, blind. In one variation, the two marks are simply pooled and averaged, which may contribute a partial solution to the above problem. However, the real use of double marking is to encourage the discussion of issues that the discrepancies raise, and to encourage a clarification of the criteria and their interpretation. Real improvement will only occur when these issues are resolved through the development of a genuine community of assessors with a shared culture of standards. While such approaches to marking are undermined by the increasing modularization of courses and pressures on time, examiners' meetings need to focus on developing shared contexts and criteria for marking as well as determining the actual marks.

Rubrics

A criterion assessment grid, commonly called a rubric, is a scale or set of scales designed to evaluate student work, by identifying different degrees of quality for individual criteria. There are two common types of rubrics:

- *Analytical*, in which several scales are applied to the same work, with each criterion examined separately (see Table 8.2).
- *Holistic*, in which only one overall judgement is made, with all the criteria considered together (Brookhart, 2004).

There are, however, advantages and disadvantages to each rubric type. For example, an analytic rubric provides more specific feedback but may be less practical as it generally takes longer for the instructor to use, with more criteria to consider. Conversely, a holistic rubric may not provide the student with much information about how to improve, but may speed the grading process for the instructor. As such, holistic rubrics, unlike analytic rubrics, tend to be more summative than formative in nature (Brookhart, 2004).

Table 8.2 *Sample rubric*

	1) <i>Weak</i> Little or no evidence of outcome	2) <i>Basic</i> Beginning of/some evidence of outcome	3) <i>Proficient</i> Detailed and consistent evidence of outcome	4) <i>Strong</i> Highly creative, mature outcome
<i>Criterion 1</i> Analysis of texts				
<i>Criterion 2</i> Organization of key concepts				
<i>Criterion 3</i> Clarity of prose and writing style				

Exemplars

Offering examples of student work that can represent or ‘anchor’ each level of performance quality may also enhance validity (Brookhart, 2004). By comparing the model with each level, the instructor will maintain consistency. This is also helpful when working with multiple graders or examiners.

Norm, criterion and peer/self-referenced assessment

The third tension concerns the point of reference for the assessment: whether it is instructor-referenced or peer or self-referenced. There are two main types of instructor-referenced assessment: norm and criterion-based. Marks and grades are, of course, essentially norms by which to judge the difference between students. This kind of assessment, often referred to as ‘norm-referenced’ assessment, aims to enable effective and reliable discrimination among students. While students often seek information about these differences, this does not necessarily tell them much about the quality of their thinking or what they are able to do.

The issue is not so much about what they achieve but more about what their status is in relation to other students. If too many students achieve the required outcome, then the norm-referenced assessment has been a failure since it will not discriminate. Assessments which grade against sets of pre-determined criteria, on the other hand, are ‘criterion-referenced’. With criterion-referenced assessment, every student can realize the levels established

by the set criteria. In this respect, criterion-referenced achievement helps students to understand how far their thinking and their performance have progressed.

Finally, while assessments are often instructor-referenced (generated by the instructor), they may also be generated by individual students (*self-referenced*) or by their classmates (*peer-referenced*). In so far as they offer different perspectives on the student's growth and development, they also tend to be more criteria-referenced than norm-referenced in structure.

Issues arising from traditional assessment

Although there have been more recent moves towards a diversity of assessments, traditional assessments are still pervasive and automatic, often still weighted heavily in many classes. Despite the intentions of faculty, and even students, traditional timed, in-class, cumulative or partially cumulative exams do not give an adequate picture of the many and varied abilities which are developed in higher education. Distortions are frequently encouraged by exaggerated and stereotyped perceptions that students have of traditional exams. They are sustained by the idea that exam results provide a grand verdict on a student's academic worth.

This verdict, however, is rarely based on evidence accumulated from a variety of settings that resemble the sort of tasks and situations they will face in later life. Instead, under the traditional system, it is based on a situation – typically several hours of writing with very tight time constraints and a heavy memory load – that they will probably never face again. Criticisms of how such traditional assessment can disadvantage the student will be familiar to many teachers in higher education. These critiques include:

- too much emphasis on memory;
- too much stress on factual knowledge;
- too great a reliance on speed of writing and thinking;
- too great an element of luck;
- too much pressure of a kind seldom found in later life;
- too little scope for originality and sustained writing; and
- too little opportunity for constructive feedback.

A major difficulty with the traditional exam system is that it gives teachers so few leads on how to learn more about the student's experience of learning in their course. A great deal can be done to diversify our assessments

so that different desired outcomes are evaluated separately. While the administrative functions of assessments may be simplified by overall one-dimensional grading, they are not necessarily effective predictors of a student's future work, even as predictions for graduate work (Kuncel et al., 2001). Overall assessments may differentiate students reasonably consistently, but the purpose of this particular way of categorizing students is too general. What is needed is a range of information about the very different abilities of students that can be used in different ways, depending upon why that particular selection is being made.

Formative assessment is also not well served within the traditional system, primarily because of this lack of differentiation of abilities, but for other important reasons as well. The timing of traditional assessments is usually too late for early feedback and information is often regarded for some reason as confidential, but perhaps even more important are the attitudes of faculty and students. Few students regard their exam performance as an accurate indication of their ability, and faculty may not encourage discussion with students who do not seem to want it.

Traditional exams can be used to direct students' efforts to important areas of the course, but they are generally not good at directing students' efforts towards developing a higher level of intellectual abilities to be used in these areas. The general purpose of providing an incentive to work is often successful, but extrinsic motivation is an unconvincing part of higher education. Indeed, it conflicts with the often stated aims of education which rank love of learning and intellectual development above all else. This distinction may not always be very clear, but intrinsic motivation is generally considered more valuable. It can certainly make a greater contribution to the important question of student identity mentioned earlier.

Ultimately, how we assess students affects the way they learn and develop. Yet, any system of assessment must be practical and acceptable to faculty and not unduly increase their workload. As assessment becomes more closely integrated with learning and teaching, the time spent on it is not just concerned with grading but is an integral part of teaching. The acceptance, moreover, of peer and self-assessment methods can save faculty time, as well as contribute to students becoming more independent professionals. It can counter, for example, some of the unfortunate consequences of objective, machine-scored exams which, while time saving, can encourage routine low-level learning and dependency. Indeed, we might heed the warning put forth by the Senate Committee on Examinations at the University of Edinburgh in the nineteenth century: 'The excessive employment of selective examinations is gradually subverting all that is best

in the education of our youth and a reaction is threatened which may bring the use of examinations altogether into discredit if remedies be not found for the worst abuses of the system.'

Student views on traditional assessment

Although many students may take a strategic attitude towards traditional assessment, and do what they need to do to get high marks, they are not necessarily naïve about the nature of academic work. Early work on face (perceived) validity (Jones et al., 1973) suggested that students were very concerned about what was being assessed within traditional examinations. Two thirds of students in a wide range of disciplines agreed with the statement that 'in planning my work I frequently found my real intellectual interest had to take second place in the need to get good marks'. A high proportion felt that the intellectual qualities that the lecturers valued were not tested in the examination. Despite claiming to know what lecturers wanted, almost two thirds of students felt that they were not given a clear idea of the academic qualities faculty expected in their work.

Moreover, students often found their experience of assessment contradicted their own personal and educational aspirations. Over thirty years ago, a student answering an interview question about his experience with traditional assessment observed: 'Writing four questions in a three hour exam paper is a total contradiction of what you've been training yourself to do throughout the year.' Answering the same question, another commented: 'It's a great mistake to be over sophisticated in exams: its part of the technique not to be: one has to limit oneself' (Cox, 1975). Both these comments suggest the students' realization of, and possibly frustration with, an assessment system that asked them to take a strategic approach to studying, rather than a deep, perhaps more satisfying, approach to learning.

In a more recent study, students revealed similar attitudes towards traditional assessments (Thomson and Falchikov, 1998). One student remarked on the futility of delving into course content deeply: 'You skim over the surface... it's just a rush job all the time.' Another student, agreeing, added: 'You get the situation where people are not learning anything, they're just doing the assessments... and doing enough work to get through the exam' (Thomson and Falchikov, 1998).

Assessment, the dominating instrument of control in higher education, gets bound up with the struggle for independence. Conformist students may lose out in this battle and become passive learners whose view of education is dominated by memorizing. Other students can become rebellious

and find their energies are dissipated in hostility or may withdraw in apathy from a system they disagree with. Only the more independent students can cope with the system and yet maintain a sense of integrity.

For an increasing number of students, however, integrity dissolves into simply coping and/or the quest to attain higher grades. And as students so respond, ensuing pressures build within the assessment system towards the giving of even higher grades. Indeed, in recent years, as issues like 'grade inflation' have become a more pronounced concern across academia, higher education's very marking practices have been called into question (Hu, 2005).

Faculty conceptions of assessment

It is not surprising, perhaps, that the differences in students' approaches to study and assessment are somewhat reflected by contrasts in how teachers understand the role and meaning of assessment in their courses. Watkins et al. (2005) found that faculty members' conceptions of assessment fall into three main categories. At a less sophisticated level, assessment is seen as external to teaching. The teaching goal is knowledge gain, and the assessment is viewed as conveying to the instructor what knowledge students have gained and what they will be able to do as a result of the course.

In contrast, at the next, transitional level, instructors believe assessment will convey whether their students have the basic knowledge required to seek more sophisticated knowledge. At the most sophisticated level, assessment is seen as internal to teaching, and the teaching goal which is developing learners. Assessment for these teachers is understood as helping students develop sophisticated strategies for learning - e.g. reflecting, analysing, interpreting (Watkins et al., 2005).

To a some degree, the contrast between these conceptions of assessment reflects potential tensions inherent in the aspects of assessment depicted in Figure 8.2. Less sophisticated teachers' conceptions appear to focus on the summative, reliability and norm-referenced aspects of the figure which are more external to the teaching, while the more sophisticated tend towards the more formative, validity and criteria-referenced aspects of the figure which are more internal to teaching.

THE INTELLECTUAL DIMENSION

Communicating assessment expectations

This section focuses on how we can support student learning through assessment. It provides a rationale for the way we assess students that will:

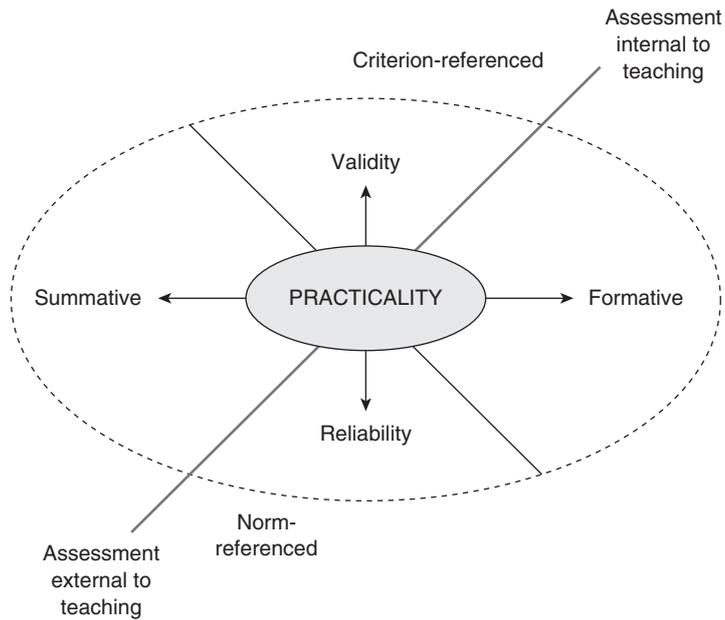


Figure 8.2 Faculty conceptions of assessment by aspects of assessment

- enable them to achieve a clearer understanding of the criteria and standards;
- give them more confidence in the reliability, validity and fairness of the system; and
- provide better feedback and reporting.

Often, students derive more understanding of a course from the demands of its assessments in the course than from their teachers and course handbooks. If we want students to understand what their courses are meant to achieve, we need to offer more than course descriptions and objectives. We need to convey clearly the ‘real’ demands of how they are being assessed in the course.

Making criteria readily understandable and transparent are an important component in helping students understand what is expected of them in a course. But merely telling students may have little effect if they do not share with their teachers the same assumptions and understanding of the criteria used to assess their exams and assignments (O’Donovan et al., 2004; Hounsell, 2005). They interpret the written criteria and feedback comments in a way that fits their assumptions rather than the teacher’s intentions.

Helping students develop a deeper understanding of the course by looking at the demands of the assessments is challenging and involves a gradual development of skills that require practice. One of the most pressing challenges in this respect, and one which incites the ire of most teachers more than any other, is the issue of plagiarism.

Avoiding plagiarism

Despite our warnings and strict punishments, student plagiarism still remains prevalent in higher education (Lieberg, 2008). Technology may have exacerbated the problem as well: with a few clicks, students can cut and paste online text into a paper or simply access an essay or report readily available on the Internet (Scanlon, 2003; Martin, 2005; Purdy, 2005).

But there are several layers underpinning the practice of plagiarism that need to be sorted out. On one level, the teacher might consider whether the student truly intends to steal the words of others and pass them off as his or her own. In some cases, plagiarism may have been an act of:

- deliberate treachery ('this is the easiest way to a high mark', 'I won't get caught because my professor isn't likely to look at this book or at this website', 'I know my friend wrote this for a different class');
- laziness ('I don't feel like doing this assignment; this person expresses everything much better than I ever could anyway'); or
- desperation ('I can't cope with the expectations of this assignment').

Other students may have committed plagiarism unwittingly, for a variety of reasons. They may erroneously believe that simply changing a few words or reordering a list, for example, is sufficient to avoid the charge of plagiarism. Plagiarism may also be culturally conditioned: meaning what is understood to be academic theft or the unwarranted borrowing of words and ideas may vary across cultures (Sowden, 2005; Gu and Brooks, 2008). Lastly, what constitutes plagiarism may differ widely across disciplines. For example, it may be common for students in the sciences to work together on lab reports, but the point where they must work independently may vary.

On a different level, teachers will want to consider the overall course design in deciding the strategies and methods they can take to minimize plagiarism, whether deliberately or unwittingly committed (Lieberg, 2008). They will want to define and discuss their expectations for their students and to explain clearly what constitutes plagiarism in the course and for specific assignments. They shouldn't assume, either, that students will enter university knowing

what plagiarism entails or that they should have learnt it on a specific course (e.g. English composition) at some point in their undergraduate career.

To minimize plagiarism, teachers will want to tailor assignments in a way that allows specificity of topic (to avoid easy replication of general ideas) and which requires students to show independent creative thinking. If students are asked to apply principles or concepts to something specific and unique, engaging their own experience, there is less likelihood that someone has already written on the exact same subject. Giving students autonomy and choice – by allowing them to pick topics, materials and methods that are meaningful to them, or which allow them to relate material to new contexts (as appropriate) – will also help motivate them and spur deeper learning.

Engaging students in assessment

Marking exercises and criteria

Engaging students in marking exercises is an important way to improve their confidence in their understanding of the criteria and demands of the course. For example, an instructor may ask her students to mark or comment on an essay or assignment written by a former student, rather than completing the assignment themselves. She could then lead a discussion about what sort of criteria should be used in their own assignments, rather than simply offering an official line on criteria and marking standards. They are likely to learn a great deal more about marking if they attempt to formulate the criteria themselves. By comparing and discussing the marks with the teacher, they can begin to revise and develop their criteria in ways that might improve their own work. Gibbs (1981) provides an interesting exercise where students mark two contrasting approaches to the same essay question that brings out interesting points related to the issues of deep and surface learning.

Comparison of rankings

Another way to increase students' understanding and ability to operate with varied criteria is to give them a range of essays on a particular topic which have already been graded by faculty. Their task is to rank them and compare their rankings with those of the faculty. The instructor can assess their performance on this task by comparing how close their rankings correspond to the faculty ranking. This has two very practical advantages: it enables students to see and consider a range of alternative responses to a set question and it enables faculty to mark large numbers very quickly.

This type of activity can be a useful preliminary to more serious attempts to develop self and peer assessment. It can enable students to move from a quantitative view of the value of their work, expressed in terms of more and more detailed information, to a more qualitative approach based on differences that reflect different levels of intellectual understanding.

Objective testing

When factual information must be assessed, objective testing – such as those employing true–false and multiple-choice questions – can contribute to the overall assessment of student learning. Easily graded and time-saving, objective tests are a staple in many fields, especially in the sciences and health professions. In one sense, they can be very reliable: answers are right or wrong and frequently are machine graded. In another sense, however, the choice of right answer and the nearness of the ‘distracters’ to this right answer can create areas of ambiguity and uncertainty that diminish this ‘objectivity’. In addition, such assessment methods frequently appear artificial and detached from any meaningful context.

While generally used to measure knowledge, comprehension and application, when well designed, objective tests can also gauge higher-level areas of learning, including the student’s ability to hypothesize, predict, analyse and generate new ideas. Indeed, it is now becoming common to insert factual questions within a narrative where the questions are linked to a specific, concrete context, argument or diagnosis (Case and Swanson, 2001).

Certainly, objective tests are in danger of being overused, particularly as faculty get more crunched for time. They may find that their advanced students have to unlearn years of study practices – such as memorization and rote-learning – developed from too much emphasis being put on objective tests. Since the more important challenges of assessment are essentially holistic, demanding the demonstration of understanding and the construction of meaning, objective tests may be best used alongside other forms of assessment that do not require the reproduction of excessive factual information. Such abilities are difficult to assess within objective format, although more objective kinds of questions can be successfully integrated into complex texts that require a more holistic approach from the student (Brown et al., 1997). Asking students to explain or annotate their objective testing responses, even if they selected the wrong choice, may help students address their misconceptions and clarify their thinking to the instructor (McKeachie, 2006).

Feedback to students

One of the key findings which Light (2001) reports in his study of what students felt help them the most during university was receiving timely feedback from their teachers. Feedback serves several important purposes, including:

- helping students understand the mistakes they have made, as well as the underlying causes of those mistakes;
- suggesting ways students can keep tabs on their own learning;
- helping students find ways to help themselves;
- identifying areas of achievement (Hattie and Timperley, 2007); and
- enhancing the learner's professional and personal growth (Bhattarai, 2007).

Frequent feedback to students can be particularly supportive, especially for those who are rather anxious. It is important for students to have a sense of their own progression through the course and how they are coping with requirements. Feedback from summative assessments in the form of grades or marks alone might provide students with a general sense of where they stand in the course and among their peers, but might not contribute much to their sense of intellectual growth. These, of course, can be supplemented with constructive comments where appropriate and possible.

Criterion-referenced assessments at least provide the student the opportunity to reflect upon and understand how far their thinking and their performance have progressed. Summative feedback from norm-referenced assessments provide little more than a sense of peer positioning which might, depending on that position, develop or undermine the student's confidence in their progress. Support for students through formative assessment is generally criterion-referenced, helping the student to achieve the learning outcomes of the course in accordance with its established criteria and standards.

Feedback can, however, be given in different ways with different objectives. In a study of feedback given to Cambridge students in the UK and MIT students in the USA, Cox (2007) mapped out four ways in which faculty approached the giving of formative feedback on assignments. These approaches to feedback are contrasted with respect to the ways in which they may be classified – closed and open – and the ways in which they are pedagogically framed – controlling and non-controlling (Table 8.3).

The four approaches presented in Table 8.3 are described in terms of their respective frame and classification, and they are illustrated by examples of the kind of comments which might be given to students by teachers providing the feedback. Closed-boundary approaches to feedback restrict the feedback to the specific assignment. Comments from open-boundary feedback, on the

Table 8.3 *Approaches to feedback*

Classification: focus on boundaries	Framing: focus on pedagogy	
	Controlling: didactic teaching style	Non-controlling: facilitative teaching style
<i>Closed boundaries</i> Restricted feedback Focus on assignment or topic	<p>Closed directive: Narrow orientation – feedback is focused on generating excellent assignment</p> <p>Teacher-centred – feedback is directive and focused on providing clear directions to meet clear criteria with clear grading</p> <p>Illustrative comment: ‘Your discussion of... is not relevant’ ‘You need to give examples in section two’</p>	<p>Closed explorative: Narrow orientation – feedback is focused within the narrow boundaries of generating an excellent assignment</p> <p>Learner-centred – feedback focused on engaging student and eliciting student-centred responses. Informal</p> <p>Illustrative comment: ‘How could you improve this assignment?’ ‘What would be another approach to this problem?’ ‘What did you mean by...?’</p>
<i>Open boundaries</i> Developmental feedback Focus on wider educational context	<p>Open directive: Wider orientation – feedback is on the assignment but open to crossing the specific assignment boundaries to looking at wider educational context and issues</p> <p>Teacher-centred – feedback is focused on the wider context of the assignment but in a direction and on material controlled by the teacher</p> <p>Illustrative comment: ‘What do you think Bernstein would say about your point of view?’ ‘You need to review more about the history of this problem’ ‘You should read his last two books’</p>	<p>Open explorative: Wider orientation – feedback on assignment is used as a stimulus for educational development: critical reasoning, writing and creative skills</p> <p>Learner-centred – feedback is open, informal and less directive – focused on eliciting student-centred responses</p> <p>Illustrative comment: ‘How do you go about writing assignments?’ ‘What have you learnt about structuring arguments?’ ‘Can you think of other areas where this applies?’ ‘Do you think all your assumptions are appropriate?’</p>

Source: Cox, 2007: 110

other hand, direct the student’s attention outside the assignment boundaries to the wider educational context and issues of the topic.

Comments provided in controlling pedagogical frames are essentially teacher-focused, directing the student’s attention to the solutions and material the teacher feels will achieve the goal of the assignment most effectively. Non-controlling frames, in contrast, are focused on engaging the student in a more reflective, learning-focused activity for which the assignment is essentially the catalyst. Faculty may, indeed, provide comments which span these categories.

The dangers of feedback are that students can become over-dependent on this support. The less controlling approaches are a step in the direction of moving them away from such dependence, encouraging students to internalize critical feedback processes of their own work. The ultimate aim in students becoming professionals is, of course, that they are no longer highly dependent on others for judgements about the quality of their work. As Bruner comments: ‘the tutor must correct a learner in a fashion that eventually makes it possible for the learner to take over the correcting function himself, otherwise the result of instruction is to create a form of mastery that is contingent on the perpetual presence of the teacher’ (1966: 53). This is where the self and peer-assessment techniques (see below) and the necessary training for doing this can play a useful role. It can aid students in developing the skills to reflect upon and assess their own work and progress.

Encouraging creativity, choice and self-critique

If we are too supportive, as we have suggested above, students may become dependent and avoid the more challenging demands that would enable them to become independent learners and professionals. In this section, we consider what methods of assessment might be most appropriate in encouraging intellectual independence and helping students achieve higher-level objectives: creativity, choice and self-critique.

To encourage independent thinking, however, it is necessary to make it clear to students that their work needs to reflect those deeper, higher-level processes. These include the processes of analysis, synthesis and evaluation described, for example, by Bloom’s Revised Taxonomy and in the higher-level outcomes expressed in the SOLO taxonomy mentioned above.

Operating at these levels is not easy and a major obstacle to doing so is lack of time. Overloaded courses are notorious in encouraging surface-level approaches to learning and assignments. Examinations with high time constraints can give little opportunity for reflection, considering alternatives, appreciating different contexts and integration. The tasks set on many courses, especially on foundation or survey-level courses, are very often short term, to be completed in a single session. Work which requires extended concentration over weeks or months is usually restricted to the final year of a programme, but the quality of the thinking that is generated out of long-term commitment can be very important for encouraging independence.

Table 8.4 *Student experiences of projects*

Advantages	Student experiences
Enable the student to explore deeply a field or topic	'It's the only time you do your own work'
Develop initiative and resourcefulness	'It's not just an exercise, it's good that you have to do it all on your own and fit it into a theoretical framework'
Enhance time and project management skills	'The best opportunity we had – other work is other people's thoughts'
Provide personal ownership of learning	'Very much me, my biases and beliefs,' 'It's something you can get your teeth into, something you choose yourself'
Foster independence and creative problem-solving	'Some original thinking is required, essays are more regurgitatory [sic] of other people's opinions'

Sources: Adapted from Cox, 1975; Brown et al., 1997

Projects

Projects have long been recognized as perhaps the most important area for self-expression and commitment (Cox, 1975). For many students, the first years of college may be spent memorizing other people's work and only in final years are they able to begin seriously to engage with the discipline and subject matter. Table 8.4 provides a list of the educational advantages suggested by Brown et al. (1997), augmented by comments from students on project work from seminal research by Cox (1975).

Project work also provides a good opportunity for students to engage in divergent – as opposed to convergent – thinking (Rowntree, 1987). Although this will vary among students and disciplines, the undergraduate in particular is frequently not provided with substantial opportunity to engage with coursework which permits them to diverge across a range of possible answers, concepts, meanings, solutions, approaches and so on. Yet, this is an essential feature of any creative work or in-depth inquiry or research. Many, if not most, of our methods of assessment encourage convergent thinking, asking students to converge upon a right or best answer. While project work, like any coursework, can be derivative, poorly planned, badly referenced, tedious, etc., it does also offer students opportunities for choice, creativity and divergent thinking – indeed, it can be written into the criteria for project assessment.

The disadvantages of project work include concerns that it is time-consuming to set up, monitor and provide feedback on, and that it is difficult to assess failure fairly. Wide variations in the help sought by students and the fear of plagiarism – the latter enhanced by access to the

Internet – are also obstacles to consider. None of these disadvantages is, however, insurmountable, and Brown et al. (1997) regard problems in assessing projects as no more than those involved in assessing coursework, essays or written papers.

Portfolios

Portfolios are another form of assessment that allow students a wider range of choice than more traditional methods, and may be a more accurate reflection of student learning (White, 2004). Students may be asked to provide a portfolio of evidence of achievement in terms both of outcomes specified by the course and a wider range of abilities and achievements which are more personal to the individual students and their particular interests.

Research on the personal development and motivation of students often stresses the importance of them having a sense of control over their own environment, and this may increase the students' sense of responsibility in the learning process (White, 2004). This is typically expressed in terms of how many choices they are able to make in regard to what they learn and how they learn (Cox, 1996). They can show a respect for individual differences and the varied contexts in which people work.

Portfolios are now becoming a widespread way of assessing professional and continuing development – including now the accreditation of teaching in higher education. Nevertheless, their use is usually accompanied by guidelines as to what should be included and how they should be structured. They have similarities with records of achievement and are quite often used for assessing prior learning and prior experiential learning.

Open-book exams

Most assessment that is concerned with encouraging independence is much more open than traditional assessments. Allowing textbooks into examination rooms may encourage independence to an extent, but it depends on how the student uses the textbook. It may assess an ability to locate information quickly rather than the ability to use it in more independent, creative ways. Open-book exams usually specify the texts and resources that may be consulted, but with the expansion of coursework as opposed to examinations, it becomes more difficult to justify making these types of assessment within an examination hall.

Prior-notice exams

A more useful method of assessing the ability to produce work under time pressures and in a more secure environment are examinations where the

topics or the actual questions are given out in advance. Students can carry out research and develop their understanding in libraries beforehand. Cox (1975) found that sociology students reacted differently to the change in constraints. While prior-notice exams relieved the constraint on memory, many students feared that the standards expected would be correspondingly higher. However, they provided the more independent students with the opportunity to explore issues in depth and to take risks and express what they themselves thought about issues rather than reproducing textbook answers. They had the chance to justify their ideas in a way in which they felt they never had in traditional exams. For some dependent students this form of assessment provoked anxiety and they found it difficult to stop preparing for it. The majority, however, found it quite liberating.

Problem-based learning

Problem-based learning (PBL) and inquiry-based learning (IBL) are an important development for encouraging independence in students (Sadlo and Richardson, 2003; Lee, 2004). PBL courses are designed around real-life problems, whereas in IBL they may focus on a specific lecture or assignment. The more abstract and, sometimes, less engaging aspects of the course are learnt in relation to a close involvement with problems which are highly relevant to students' more concrete ideas of what the course is about and what competence would mean within it.

Although PBL and IBL courses use examinations, they are generally less important than in courses that are more traditional. The actual solving and understanding of the problems are the most prevalent form of assessment. This more contextualized and naturalistic assessment enables students to feel that assessment is not a control mechanism, but is a natural feature of learning.

Self-assessment

The development and increasing use of methods of self-assessment is, perhaps, the most important innovation in assessment for the development of intellectual independence. In contrast to the vast majority of assessment methods, it directly addresses the paradox of a highly dependent education leading to the independent responsible status of a professional person. Indeed, self-assessment is a critical aspect of a student's growth and development as a scholar (Cassidy, 2007).

Introducing self-assessment is not without challenge, however, particularly in academic cultures unused to innovation, or where students are unduly competitive. To thrive requires a culture of mutual trust among

students and faculty. Teachers have also frequently resisted the development of self-assessment methods, suggesting that students can be overgenerous in their marking. In fact, there is evidence to suggest the opposite. Generally, student reactions are very positive. Five times as many students found it a worthwhile experience as those who did not, and a similar ratio found the exercise helped them to pinpoint their strengths and weaknesses (Boud, 1995).

A critical issue in the development of self-assessment concerns strategies for generating criteria against which the student will assess his or her work. It is important that teachers do not simply issue criteria from above but rather provide students with a role in formulating and refining them together with teaching faculty. Exercises in peer marking can be helpful in this respect. Although learning from mistakes has its benefits, there can be serious problems in introducing self-assessment into an unprepared environment. Early bad experiences can easily make it extremely difficult to try to introduce it later.

It is important to establish clear rationales for involving students in developing the criteria. In the process they will also learn to make qualitative judgements, including justifications for the assessments they make. Boud (1995) also stresses that self-assessment practices should permeate the total course. They should make an identifiable contribution to formal decision-making and be part of a profiling process in which students are actively involved.

Self-assessment can be viewed, not as a distinct element of teaching and learning, but in relation to reflection, critical reflection and meta-cognitive practices. It is part of that set of activities which encourage students to take responsibility for their own learning, monitor their learning plans and activities, process their studying and assess their effectiveness. Self-assessment then would become something which is embedded in courses designed from the very start to assist students with their learning (Cassidy, 2007).

Interpersonal (peer and group) assessment

Historically, higher education has been primarily focused on the individual, encouraging the student to 'think for himself [*sic*] and work on his own' (Hale, 1964: iii). Collaboration with respect to assessment was highly suspect. The development of interpersonal skills, co-operation and abilities to work in teams is, however, increasingly valued. In this section, we look briefly at the widening the scope of our assessment systems to engage with interpersonal skills.

Peer assessment

The discussion of self-assessment inevitably raises questions of peer assessment since the ability to become an effective self-assessor is often enhanced by assessing and being assessed by peers. It is an important issue in assessment generally to be able to appreciate different perspectives and points of view. Attempting to assess a range of different student assignments can be extremely interesting for students, enabling them to see a wider range of perspectives and solutions. Of course, in seminars, students will be exposed to different views, but this is a very different experience from assessing and analysing them critically for their different strengths and weaknesses.

These skills are also critical in a multicultural society that expects people to understand and tolerate multiple views and perspectives, which values the intellectual virtues of parallel thinking (De Bono, 1994) rather than patterns of critical thinking which encourage the rejection and substitution of intellectual positions. Employers value interpersonal skills and communication in a way they have never done before. Modern companies are less hierarchical and their employees are more likely to work collaboratively on projects rather than simply follow instructions.

Issues of co-operation and competition permeate all levels and activities in society, a point well understood by millennial students long used to working collaboratively in groups and teams (Howe and Strauss, 2003). We think less in terms of single dimensional intelligence (IQ) and more in terms of 'multiple intelligences' (Gardner, 1999), which include the interpersonal which can be enhanced by peer assessment as well as the intrapersonal intelligence which can be enhanced by self-assessment.

Being able to discuss assignments with faculty has long been an invaluable part of higher education. Written comments on assignments followed by discussion can be one of the most effective learning activities in academic life. While still an essential aspect of graduate supervision, it is growing harder at the undergraduate level, given the larger faculty-student ratios. Peer assessment can provide students with new perspectives in a way teacher assessment now seldom can.

As with self-assessment, peer assessment is not something that can be effectively introduced without considerable practice and reflection. This development, however, may be done in larger groups and the investment of considerable time in extending the skills of giving and taking criticism can be an extremely valuable learning activity. It is important to see this as a skill that is developed over the whole range of teaching and learning in higher education.

Traditional student-led seminars, for example, are often painful experiences, either in the sense that criticism is badly given and/or badly taken, or in the sense that serious discussion and criticism are difficult to initiate and sustain. Process-oriented time in seminars – conducted in an atmosphere of mutual trust – can be a useful occasion for developing the interpersonal skills that are essential in peer assessment. Equally, what is learnt in peer assessment may enable group work to function in a more constructive and helpful way. Certainly, research has shown that peer assessment can support the goals of group learning, particularly collaborative learning and co-operation, and can allow a group to assess individual contributions holistically (Lejk and Wyvill, 2002).

Consultants and assessors exercise

An interesting way to combine group work with peer assessment is the simulation exercise of consultants and assessors. This very useful teaching exercise consists of setting up consultant teams to address a specific problem in the particular subject area.

Briefly, the class is divided into consultant teams of about four to six members. Each team prepares a report on the problem for a group of assessors drawn from each team. The assessors' group formulates criteria by which to judge the several consultant team presentations. The exercise can take anything from an hour and a half to several days and may involve purely thinking and discussing in the groups or consulting various resources. Questioning by the assessors can draw out many issues in assessment, as can the judgements themselves.

As important, however, is a debriefing session when the consultants assess the assessors on the criteria they devised and applied. The way in which they are applied can come under very active scrutiny by both the student consultants and the teacher. The debriefing can also be a further time for co-operation both in terms of the intellectual task set and understanding the process.

Group projects

Group projects provide a further opportunity for intellectual and interpersonal development with respect to assessment (McKeachie, 2006). Their increasing use in higher education parallels the growth of project-orientation approaches developed in industry. The traditional focus on assessing the individual is one reason why group projects have not been used as often as they might. How do you 'fairly' allocate marks? Do you give all members of the group the same grade despite the fact you may know that some have actually played very little part in it? Many teachers feel that competition for marks

within the group may not facilitate effective problem-solving, co-operation and learning. Answers to this question are very much dependent on the nature of the projects and the aims which they are designed to achieve. There are at least three different approaches.

In the first approach, the instructor or the group divides the tasks into sections, which are then completed individually and marked separately. Some parts, such as the introduction and the conclusions, might need to be considered by the whole group. In this context, each individual might receive two marks: a common group grade given for the overall quality of the project, and a grade for the each member's individual contribution. This may be an effective approach if the project can be divided into equally challenging sections.

A second approach might be taken if dividing the project into separate sections is too difficult or artificial. Here, each member receives the same group grade, but also writes a reflective commentary on the process, describing their own specific contributions as well different levels of commitment and understanding within the group. This approach has the added advantage of encouraging and rewarding reflective practice.

The last approach involves aspects of peer assessment and asks the group to allocate a proportion of the total mark to each of the participants. Conducted badly, such an approach can be a recipe for antagonism within the group. As with other forms of peer assessment, it needs to be carefully discussed and planned with the students before they actually begin to work on the project (McKeachie, 2006). Students will need to agree beforehand how decisions will be made and the criteria for dividing the joint mark. In any group work, of course, the issues are not purely intellectual. There will be many personal issues raised and the social composition and practical functioning of the group will be crucially important and will materially affect the outcome.

THE PERSONAL DIMENSION

Supporting personal development

In this section, we stress the importance of assessment in providing support for personal development and suggest some ways of doing this by encouraging choice, learning contracts or agreements and reflective commentaries.

Encouraging choice

We have suggested in earlier chapters that higher education is not just about transmitting information and developing particular competencies

and skills. It is a time for self-understanding and personal development, both in the sense of coming to know ourselves better and coming to know what sort of professionals we may be. In professional courses, the match between the student's understanding of what it means to be an engineer or a doctor and what the course seems to be providing can be crucial both for motivation and for intellectual development.

Higher education is often less about learning specific topics and more about developing constructive ways of approaching them. Our own identities are very much an expression of the successive choices we make in life. This is extremely important for students beginning to know themselves (Weinstein, 2003). If there is little choice in the courses and the assessments they have to cope with, then they can easily feel they are being shaped and moulded by others. In extreme cases, students might rebel against what they perceive and experience as external control and the lack of choice, even to the extent of failing. Indeed, failing in such cases – where narrowly constructed parental and academic expectations are experienced as oppressive – is often seen as the only way for some students to exert choice and develop themselves personally. 'Success' simply acknowledges the pressure to succeed and conforming to it. Choice and opportunities for self-direction might alleviate this desire to frustrate the aims of others by failing.

Learning contracts or agreements

Providing students with the opportunity for informed choice and the time to take that seriously is an important part of the teacher's role. As noted above, project work provides one way to do this, but many other assignments can also offer choice among alternatives. The development of both informal and formal learning contracts – where teacher and student agree procedures and areas of inquiry – can also encourage students to feel they have a more personal role in their education. These can also provide an important platform for personal reflection. In arts subjects and, to some extent, social sciences the expression of personal perspectives has always been valued but, in the sciences and medical sciences, this has been difficult. With the decline of positivism and a willingness to take more open positions in many areas, the possibility of developing personal views is increasing.

Reflective commentaries

Supporting students to reflect on their choices and judgements enables them to feel they are developing personally while in higher education. In some disciplines, it may appear difficult to find areas of personal expression. Even in very technical projects, however, asking students to write a section

in reports on their personal response to the experience can be a useful way of encouraging them to understand and extend their own responses and experiences of learning.

Encouraging identity exploration

Many of the issues discussed under personal support are designed to enable students to have a more secure sense of personal identity. This is also, of course, an essential aspect of developing independence. Changes in the nature of higher education in the past three or four decades have, however, emphasized 'independence' aspects of a student's identity more acutely than it once had. In the earlier more homogeneous and restricted 'elite' system of higher education, students could expect to develop their identity through 'identification' – stressing an apprentice role. In the radical changes in the latter part of the twentieth century and the development of mass higher education (Trow, 2001; Thelin, 2004), the focus of student 'identity' shifted towards exploration, gaining self-knowledge and intellectual autonomy.

The traditional examination system corresponds to the 'apprentice' role and the more modern, open and diverse forms of assessment corresponded to the 'exploration' route (Cox, 1973). This distinction matched the parallel changes in society in relation to conceptions of the structure of knowledge, occupational structure, teaching, participation, adolescent development and cultural and social life. The 'traditional' cultural life involved fewer boundaries and expectations, identity was more prescribed, there was more consensus of values and duties, and clearer artistic styles and conventions. The more 'modern' social life – with its blurred boundaries and expectations, lack of prescriptions about identity, lack of consensus upon values and duties and lack of conventions in artistic expression – is even more prevalent in the new millennium's 'supercomplex' world.

Ironically, academic pressures – lack of resources and rising faculty-student ratios – threaten to return faculty and students to some of the more formal, traditional approaches to assessment of the past (Hu, 2005). We still find within some of the more formal subjects, fears that under the pressure of overloaded curricula there is no scope for individuality and creativity, only for learning the right answers and reproducing them. On the other hand, PhD guidelines may recommend reflective commentaries on the experience of writing a PhD and conclusions about what they have learnt from doing it (Sloboda and Newstead, 1995). Even the personal pronoun 'I' is not automatically rejected, and contextualizing writing in terms of particular personal responses is seen as more academically respectable.

Encouraging this generally within assignments might weaken the tendencies towards bureaucratization and the commodification of higher education and assessment.

Enhancing self-knowledge in groups

Encouraging the interpersonal within the personal dimension is very close to considering the interpersonal within the social dimension. An important distinction to recognize are the aspects of self-knowledge – crucial to working within groups – which group work and assessments/reports on group work can emphasize. We learn a great deal from reflecting on our own behaviour; this is particularly the case when we reflect on it in terms of the responses, reactions and interactions of other people.

While it is still somewhat unusual for courses to emphasize self-knowledge and interpersonal skills in their assessments, recent interest in key skills and transferable skills has, as we noted above, raised their profile. At present, this might be more in theory than in practice but assessment, even formative assessment, may help to make it a more important aspect of education.

Accurate and reliable assessment in this area is still problematic. Nevertheless, the development of reflective commentaries which, for example, ask students to look at how far they have changed in response to a course and what they have learnt about their own participation and reactions to various events and activities during the course can be invaluable. Careful programmes in peer-tutoring and peer-mentoring might also provide a valuable interpersonal role in assessing self-knowledge. Encouraging students to look at their own strengths and weaknesses and their own ways of benefiting from a range of different relationships can contribute to their overall formative assessment.

THE SOCIAL DIMENSION

Fostering informal peer interaction

Traditionally, there has long been an emphasis upon the value of college life at university and the importance of participating actively in the social life and student societies. Seminars and group work, however, do not have a reputation as valuable social experiences. More recently, however, there has been more emphasis upon syndicate learning, peer-managed groups and group projects. The development of more reflective study has often focused upon peer support but, generally, there has been little emphasis on the assessment of the social dimension of learning.

Assessment has always generated emotional problems for students, and even with less emphasis placed on finals, considerable anxiety is still felt by many students undertaking and completing assignments. Isolation can be a serious problem for foreign students and for those who feel less able and may be liable to fail. Frequently this is simply not true but it is usually difficult for teachers to convince them of this. The opportunity to discuss drafts and problems and the way they are approaching their work with fellow students, however, can be very useful in allaying the anxieties that can make their pessimistic predictions self-fulfilling.

Many students provide each other with support in informal 'learning communities', usually without assistance from faculty. Unfortunately, the students who most need such support often cannot do this. The formalization of such 'learning communities', while not simply focused on assessment, can play an important role in supporting and fostering student approaches to assessment.

Innovative approaches with new technology, for example – particularly through the use of online chat rooms and discussion groups – have been valuable in helping courses become more robust learning communities rather than merely a collection of individuals. Such groups, for example, might engage in the kinds of formal marking exercises and/or activities setting up criteria for self and peer assessment, discussed above. The social dimension of coming together, however, extends broader its pragmatic and utilitarian reasons. The enjoyment of social interaction can often provide an important context enabling students to discuss their problems and worries.

Fostering independence within groups

Encouraging independence within a social dimension seems rather paradoxical. Nevertheless, developing students' interpersonal skills might merge the somewhat incongruous aims of encouraging independence and providing students with emotional and intellectual support. Working in teams or groups – attempting to solve problems, develop knowledge or design new objects and processes – requires mutual group support and encouragement within a strong emotional and intellectual climate. At the same time, students also need to make independent contributions and not simply reflect the prevailing opinions and ways of working. They need to learn to preserve their sense of individual identity while, at the same time, working in harmony with the group.

Contributing to the working climate of the group demands accepting support and making contributions which others cannot or are less able to

make. Assessment can be an important aspect in achieving this. Higgins et al. (1989) report on an initiative in a department of chemical engineering to encourage problem-solving skills. They developed a two-week programme for the first weeks of the first term as an introduction to university life with respect to both the problem-solving and the social perspectives. Teams of about six worked on general problems and then problems in chemical engineering. At the end of the programme they presented two kinds of written and oral reports, the first concerned with solving the problem and the other on how the group functioned as a group. From the beginning, assessment was used to indicate the value of the social dimension of learning and students were encouraged to reflect upon the nature of their interpersonal contribution to the group and its ways of interacting to solve problems.

Students' commentaries on their group participation need not be formally assessed. In some universities, while it is a condition of assessment, the commentary is not itself assessed. Group work, of course, is generally designed to encourage openness and risk-taking, and assessment can be an inhibiting factor to such social interaction, especially if the criteria are left vague. Nevertheless, assessing a student's actual contributions to a group can be managed effectively although it is very different from encouraging students to write about their experience and simply assessing their writing. If, for example, reflective commentaries are an integral part of the learning on the course, the task can be less threatening and criteria can be developed which are more attuned to individual variations within the group and the context in which they work.

THE PRACTICAL DIMENSION

Supporting practical skill development

There have been many changes in practical assessment over the past few years and there are many disciplinary differences. We begin each section, therefore, with some general characteristics. Essential aspects of giving support in practical work include providing:

- clear detailed instructions/briefs/checklists;
- close tutor/technician supervision;
- clear support documents/materials/equipment supplied and specified;
- clear criteria for correct methods and solutions; and
- faculty assessment for both process and product.

Objective, structured clinical examinations

While traditional practical examinations are rare now, variants of the objective, structured clinical examinations are common in the health disciplines (Hounsell et al., 2007). The exams consist of about 20 short, clearly defined practical tasks that represent the key objectives of the course. Students move from one problem/task to another at specific 'assessment stations' every five minutes. They might require taking the history of a patient and diagnosing a problem, interpreting test results, interpreting radiographs or slides of tissues, setting up or using equipment, making dissections or putting in stitches in simulated wounds. Each task can be assessed in a variety of ways and might involve simulated patients and/or the completion of short-answer questions. Generally a trained observer or a trained interviewer conducts the assessment. Although setting up the tasks can be time-consuming, the system can be very efficient – in two hours 20 students can be assessed on 20 different tasks. Careful attention will need to be paid to security issues – particularly for summative assessment – although the tasks can be varied for subsequent groups of students.

Objective, structured clinical examinations are mainly effective in assessing the specific goals of practical work. They can improve individual technical skills and even develop problem-solving skills. How far such exams can improve understanding of scientific inquiry, reinforce good practice and/or nurture professional attitudes is more problematic, but may be lessened with the developing trend towards competency-based evaluation in the medical and health fields (Carraccio and Englander, 2000).

Performance evaluation guidelines

Another way to introduce highly structured and supportive practical assessment is to document activities with comprehensive report sheets. Brown et al. (1997) describe details of performance evaluation guides and a self-assessment manual of standards developed for practical work in dentistry. Levels of practical competence are defined very closely for failing, passing and excellent grades. A failure, for example, is defined as: 'unacceptable outcomes as a result of treatment or lack of treatment which has already caused irreversible damage to the patient's aural environment, or will cause severe damage in the future' (1997: 107). Further criteria employed in the assessment include both the 'aesthetics' and the 'structural and biological integrity' of the task.

Such comprehensive reports can be made for observing groups of students who are doing such things as setting up apparatus, with such headings as reads instructions, checks layout of apparatus, checks instructions in relation to apparatus, seeks advice from the demonstrator. Each can be given an estimate of the proportion of time spent on each activity. Box 8.1 illustrates how creative work can be assessed in terms of acquired skills and overall aesthetic.

Box 8.1 *Case study: assessing creative work*

When Cameron, a professor of fine art specializing in sculpture, first started to teach at the university he struggled with assigning final marks to his students. How can one assess creativity and originality, he wondered, since art is so subjective and personal? Previously, he had taught in a selective art school, where most of the students were highly motivated and very talented and where very little attention had been paid to marks. In his current teaching context, most of his students did not major in art and very few expected to be professional artists.

He decided that, rather than trying to assess his students' work in terms of pure creativity, he would break down the assessments by considering his learning objectives. He wanted his students 1) to learn to appreciate art within its larger social, cultural and political context; 2) to acquire basic sculpting skills and techniques; 3) to create a sculpture incorporating different techniques; and 4) to be able to articulate their artistic vision of the sculpture.

To support their learning in class, he demonstrated specific techniques during class, and then let his students practise, sharing their progress as they worked. Outside class, he promoted their independence by requiring them to visit different sculptures and analyse the vision of different artists. At the end of the course, the students all showed their work to the rest of the class and an assortment of other faculty and students, enhancing the interpersonal. The final grade, which had so worried Cameron, was no longer based on a single project but, rather, on his students' participation (including feedback on each other's work), critical reflections, their level of technical skills and their written and oral explanations of their artistic vision. In the end, he still reserved a small part of the mark for the student's imagination and creativity, but a student did not need to be a 'genius' to get a high mark.

Reports on laboratory work, of course, are common forms of assessment now. While they may be concerned with more open forms of assessment, they can be very closely structured according to rigorous guidelines. Such detailed and rigorous guidelines and criteria are often obligatory – particularly with respect to health and safety procedures – and have the added benefit of improving the overall reliability of the assessment.

There are disadvantages, however. While such guidelines need not restrict independent activity, in general the more detailed and prescriptive, the less scope there is for individual and independent response and expression. Students can end up reproducing practical techniques without understanding the principles behind them. Detailed criteria and guidelines may not be sensible if used to the exclusion of more open forms of assessment.

Practice tests

Another method to support practical assessment is to allow students to complete sample exams or practice tests that are similar to the types of exams used in the course. Snooks (2004) offers a method in which students take short tests based on the required reading, which they then compare and discuss answers with a partner, and then the rest of the class. The process is designed to enhance critical thinking, question analysis and test-taking skills; to reduce student anxiety; and to imitate a real-world professional context in which colleagues share and discuss written work.

Recorded skills

In some fields, especially in medicine and other professional areas, students might benefit from having their skills recorded electronically as they perform a task or skill in a real or simulated environment, to review their strengths and weaknesses later (Silvestrone, 2004). These can be very effectively combined with peer and self-assessment approaches informed by detailed discussion with respect to the development and use of the assessment criteria.

Encouraging independent engagement

As in the other dimensions, encouraging practical independence involves a higher degree of student involvement in the assessment process. Generally it will include:

- a wide choice of practical tasks or problems;
- student involvement in planning and decision-making;
- student responsibility for finding and providing support materials;
- student involvement in criteria for success;
- student self-assessment;
- time for development; and
- risk-taking and creative responses.

All these characteristics take time but they should be seen as contributing not only to the assessment and development of practical work but also to the general development of students. Such characteristics can apply to many different forms of practical work as well as straightforward work in laboratories.

The development and use of portfolios (see above) can be a useful way of enabling students to work independently, permitting them to assemble evidence

of their practical achievements and skills – including their laboratory work. Broad practical outcomes can be negotiated and established with students regarding both what would be included and what criteria would be used for assessing them.

The types of problems and tasks considered appropriate are important in widening the scope for independent work. Brown and Knight (1994) report a range of alternative possibilities for practical work including the production or generation of artefacts and products, designs, drawings and plans, design and build, games and simulations, and IT-based work. Wherever design and production are involved there is also considerable scope for creativity or at least for significant differences between the various student products which highlight possibilities for more independent and engaging practical work. Broadening the scope of this work can also contribute to overcoming some of the dull routine that can characterize work in laboratories and other practical settings.

Fostering interpersonal practice

Developing interpersonal skills through the conduct and assessment of practical work requires, again, a substantial focus on students as operating in ‘learning communities’ or, at least, within active groups. General aims here would be to:

- establish a reasonable proportion of group practical work and projects;
- focus on group process and development as well as group tasks;
- emphasize group negotiation in executing and completing tasks;
- include the group in the setting of assessment criteria; and
- involve the group in actual assessment.

These aims or characteristics are similar to those we have discussed for other areas of assessment but they take on an added significance in practical work, particularly when it involves a more emotional engagement. Such real-world assessment is often called ‘authentic assessment’. Here, the assessment reflects how the skills would be used in the real world: the problems are unstructured and may require the student to make an informed decision from several options and may allow for feedback and second chances (Svinicki, 2004).

The actual concrete experiences of practical work can engage people as whole people rather than simply intellects. Experiential learning (as we saw in Chapter 2) involves observing, feeling and acting as well as

thinking – areas in which the development of interpersonal skills can be particularly powerful. Of course, these skills can be developed through reading, discussing and watching videotapes, but the more realistic the setting, the greater the possibility for deeper learning. Real-life settings can be the most effective but can often be difficult to arrange. Partial versions of actual experience, however, can be effectively developed. Brown and Knight give an interesting example of a practical assessment using real people in real situations:

Surveying students are assessed on how they undertake negotiations between a client and a housing association for the lease of premises. Local professionals are involved in the setting of the assignment, using actual local properties and, when possible, role play themselves, before contributing to the student's assessment (1994: 85).

Similarly, how a student performs before an audience in a creative field, such as music, might offer another real-world assessment (Stanley et al., 2001). Such activities are not restricted by discipline and provide a useful and valuable compromise between students actually operating in the 'real' world – as in practice placements and fieldwork – and in the more formal environment of the seminar room. Interpersonal issues can easily be lost in such activities unless care is taken to ensure particular concern that students do reflect upon the nature of the interactions and begin to develop more sophisticated understanding and responses. As with seminar work in general, the topic or task can completely take over from the development of interpersonal understanding and the ability to respond to the demands of the process.

CONCLUSIONS

The approach we have taken towards the issue of assessment in this chapter has been to offer information and arguments from which readers will draw their own conclusions and make their own professional judgements. There are no ultimate prescriptions or rules for the practice of assessing students in higher education. It is, rather, a developing genre of the language of learning and teaching. It needs, as we have indicated, to be reconsidered and reflectively practised within the context of both our changing understanding of higher education generally and of learning in particular. In this respect, it is a multifaceted and multidimensional phenomenon positioned at the heart of learning and teaching and its scope for innovation and the improvement of student learning should not be underestimated.

For intellectual and academic reasons, assessment needs to be carefully and accurately balanced and to emphasize the key elements of a course. For the student personally, and for the development of his or her sense of identity, the control and certification functions of assessment need to match the intellectual, personal, social and practical demands of a course. In this respect, assessment needs to be less a rite of passage and more a significant and relevant personal achievement.

This relevance needs to be apparent not just in terms of today's needs, reinforcing the status quo, but in terms of the demands of tomorrow's increasing complex and indeterminate challenges. Graduation must promise a stake in the future, not just the past; its rituals need to take students beyond traditional culture to a world of change and uncertainty. A forward-looking quotation from the past might be a useful way of ending: 'Only when the students become competent evaluators of their own goals, experiences and accomplishments do they become truly educated (liberally educated, one might say) and capable of engaging in the individual fundamental processes essential in a democratic society' (Dressel, 1976: x).

Final questions: as the last quotation suggest, assessment need not be simply an assessment of student learning, but it can model a critical student learning experience in itself – self-assessment. In designing and constructing their assessments, teachers might carefully reflect on the learning experience as well as the learning assessed. Do my assessments match, or constructively align with, the deeper learning objectives I have identified for my students? Do the assessments provide the kind of evidence which allows me to determine whether they are achieving the kind of learning outcomes I expect? Are my students engaged with the assessment process? Does success on the assessment provide them with a meaningfully sense of accomplishment?