Chapter 2 Assessing students taking your module: practice issues

This chapter is not as theory driven as Chapter 1. My voice is more muted because once design issues are resolved, albeit provisionally, there is quite a lot of agreement about the business of assessing students' work. This allows me to cover the main features of this topic by means of extracts from the work of acknowledged authorities in the field. My role is mainly to do with marshalling the extracts, providing advance organisers and making suggestions for reflection. This approach may mean that you will be doing some jumping from page to page but the Reflections are there to encourage you to synthesise the suggestions made by the scholars whose work is used in this chapter.

This is not to say that responding well to students' work, whether in the form of performances, writing, exhibitions, electronic productions or conversations, is an easy matter. It almost always involves creating some judgements about quality that are recognisably akin to judgements other people teaching on the same programme will be making. Although I shall argue that the problems are less when assessment is for local, often formative, purposes than when it is intended to produce more reliable and generalisable judgements, problems do not vanish simply because the stakes are relatively low.

In Section 2.1, I present a series of suggestions about ways of grading students' work, putting a lot of weight on the idea of criteria-related assessment while arguing that criteria are not panaceas. They may help us to judge well and to align our reasoning when we judge with other people's reasoning but they cannot guarantee reliability. I return to this theme in Chapter 3, when writing about learning outcomes, and Chapter 4, when deconstructing the notions of learning outcomes and competence.

Lest this seem to be very negative, when my intention is really the positive one of saying that when we recognise the limits of assessment we can make a better job of it, then Section 2.2 describes and even itemises good assessment practice, with Section 2.3 contributing another eleven suggestions as desirables, rather than as essentials. Chapter 3 contains Reflections that encourage you to construct your own list of features of good assessment practice at the level of a single module.

---

Reflection 2.1 Good module-level assessment plans

You might find it a useful exercise to reflect, now, on what you reckon are the most important features of good assessment practice at the module level.

Be prepared to be disconcerted by some of the suggestions I shall make during this chapter. For example:

- I am fascinated by the idea that if different learning outcomes and tasks have to be assessed in different ways, then it follows that we need not expect all assessments to be objective, let alone done by the teacher.

- If we cannot be certain what people learn, at least as far as 'soft skills', self-theories and metacognition are concerned, then it may be that we should think more about good learning opportunities and tasks and less about assessment tasks. Perhaps good assessment is less assessment.

- If we take employer views seriously, the more self-assessment tasks we set, the better, which might imply that the less grading we do, the better.
2.1 Grading work in your module

Regardless of whether assessment purposes are formal and local or summative and generalisable, it is in everyone's interests that the standards to be employed are known, clear, shared and used. Of course, there needs to be greater certainty, agreement and precision when an assessment is high-stakes and reliability matters. But even with low-stakes, formative assessment conversations, people benefit from having some common language, even if it is fuzzy and meanings are negotiable.

But where do these standards come from? I need to introduce a distinction between norm- and criteria-referencing before I can answer that question.

Norm-referencing is comparative, telling us that this student is better than another, similar to a third and not as good as a fourth. The student's rank, not mark, determines the grade and the distribution of grades is kept fairly constant. We need, of course, to have some standards in order to get the marks in the first place but those standards or criteria are not central to norm-referencing in the way that they are to criteria-referencing.

Many of us believe that good teaching makes a difference to the quality of student learning and hope that it shows up in the form of better marks. We object strongly if we are told that we must not give 'too many' good marks because that would 'devalue our standards'. This use of norm-referencing appears to be designed to deny that good students, good teaching and good learning can make much difference.

Criteria-referencing (commonly called criterion-referencing) is fundamentally opposed to norm-referencing in that it countenances situations in which everyone fails (rotten teaching, lazy learners, bad assessment) or everyone gets grade 'A'. The theory is simple. Identify what counts as successful performance or good attainment, specify it precisely and judge evidence of achievement accordingly. In most assessment situations levels of achievement are described, each with their own criterion or level descriptor.

I am assuming that it is better to use criteria-referenced grading than norm-referenced grading. So, where do the criteria, or indicators, come from?

The material entitled Creating grade indicators below provides some clear and practicable answers, to which you should add the suggestions in Setting criteria and standards which follow it. The title, 'Creating grade indicators', is mine. The author preferred to talk of 'criteria'. It is common to talk of grade criteria and, less grammatically, of criterion-referenced assessment. I prefer to talk of 'grade indicators' because the term reminds me that assessment is about judgement, whereas 'criteria' is a term that, for me, is redolent of measurement and objectivity. You might want to look ahead to the piece in Chapter 3 (Section 3.2) entitled The language of learning outcomes because it provides some rather useful, albeit rather technical, advice on writing learning outcomes, criteria and indicators. If you have access to a good library, you will also find useful advice in: Walvoord, B. E. and Anderson, V. J. (1998) Effective Grading: a tool for learning and assessment.
Creating grade indicators

*Identify what marks and grades mean*
This makes it pretty clear what is required and what is not! However, such a specification may not help marking, because descriptors like these are usually multi-dimensional; students may succeed on some dimensions and fail on others, making an overall judgement difficult.

- You can specify a rating scale for a specific criterion, for example ‘Use of colour: A B C D Fail’.
  
  This tells students something that will be taken into account in marking. However, a problem with such scales is that they do not specify standards. In this case it is unclear what an ‘A’ or a ‘Fail’ rating would signify in terms of ‘Use of colour’.

- You can specify both ends of a rating scale, for example:
  
  ‘coherent, logical, ............, muddled, bitty’.
  
  This is useful for the end points, but what does the third slot along mean?

- You can specify every point on the scale, for example:

  **Group facilitation**

  A Wide range of interventions used sensitively and effectively.
  
  B Varied interventions, positive impact, natural.
  
  C Narrow range of interventions, limited impact, sometimes clumsy.
  
  D Few or no appropriate interventions, negative impact.
  
  This would be useful for students and would also make marking more consistent.

- You can link marks explicitly to each rating scale so that total marks are a sum of ratings rather than a global judgement which might be made independently of the separate rating scales. This could improve consistency of marking at the cost of giving too little scope to overall features of students work that could not be easily captured in separate criteria ...

**Case Study: Setting criteria for assessing performance**

... It is interesting to see how in this case study criteria have been specified and used effectively to assess the highly creative skills of orchestral conducting. The same criteria are used in a structured way in both mid-course formative assessment and end-of-course summative assessment.
I am now going to summarise Walvoord and Anderson’s (1998) advice in order to complement what is recommended in the material you have just been reading. They suggest that teachers construct criteria by doing a primary trait analysis (PTA) for each assessed task they set. PTA, which has similarities with Gosling and Moon’s advice (2001, pp. 29–30). This involves:

1. Returning to module learning outcomes to remind yourself of the module’s declared intentions.
2. Identifying the aspects of the task that will govern your marking – the things you will value and those you will penalise.
3. Writing a scale of between two and five points to describe different levels of performance. The descriptions should be precise and concise.
4. Piloting this on a sample of students’ work, revising and then applying it to all of the assignments. (p. 69).

As long as they are carefully devised, there can be little objection to these marking schemes or sets of grade indicators as local guides to what is being valued, especially if they are securely anchored to module learning outcomes that, in turn, relate to a programme specification. But they are local judgements. For example, judgements about the weighting between elements are the tutor’s judgements. So, in assessing a pupil midwife’s competence in ante-natal clinic, is it possible to weight the following elements in any meaningful way outside the actual context: interpersonal skills, taking case history, conducting a physical examination, keeping records, and showing knowledge of common ante-natal problems? In another example, Walvoord and Anderson’s indicators for an original biology experiment seem to give equal weight to ten elements: title, introduction, scientific format, material and methods section, non-experimental information, experimental design, operational definitions, control of variables, collecting data and communicating results, and interpreting data (pp. 197–201). I would have weighted them differently and, if I were using these indicators regularly, I would mark only one or two elements carefully in any one week.

A second reason for treating them as local indicators is less obvious but more important. Suppose that a programme learning outcome was that students should be able to design sound experiments in biology. The score that Walvoord and Anderson’s indicators would give provides one, rather ambiguous, piece of evidence. A reliable verdict on the student’s work, judged against this outcome, would need several – ideally quite a lot – of scores from different experimental tasks across the whole programme. One task by itself is not a reliable indicator. And if different tutors used different sets of grade indicators or
criterias, then they would overlay the signal about experimental understanding with other sources of noise. When assessment is understood as local practice, then the advice presented in this section, which is perfectly in keeping with phase 3 thinking about assessment (Figure 1.1), is good advice.

Once the aim is to make judgements on programme learning outcomes there is a strong need for some programme-wide criteria — for a set of indicators that all tutors have agreed and can use in judging student performance on experimental design tasks. Tutors also need to come to an agreement about how they apply the indicators, there needs to be some second marking or monitoring of their grading practices, marks need to be clearly and openly related to the criteria and students need to have the indicators and understand them as well.

A department in which I worked had generic grade indicators for essays, a sample of which is printed below and titled Extracts from a social science department's grade indicators. Students had them, teachers used them throughout the programme, there was second marking and monitoring of grading practices, and feedback to students tended to be written in terms of the grade indicators. Yet it did not entirely solve the problem of getting affordable and tolerably reliable assessments of essay-writing because:

1. the indicators openly say that they aid but don’t replace tutor judgement;
2. they plainly need interpretation;
3. only the people teaching a module are in a position to spot fresh thinking as opposed to recycled notes (only they know the process standards).

Extracts from a social science department’s grade indicators

Table An alternative marking scheme

Copyright material removed
There are two good, recent papers on the problems of marking essays reliably – one by Fleming (1999) and the other by Ecclestone (2001). What is true for essays holds true for most other assessment methods.

If this is a problem with something as well established as essay marking, consider how much more serious it is when authentic assessments are being used to assess professional competence, perhaps through workplace performance. I return to this issue in Chapter 3. As far as module level assessments are concerned, perhaps the most important thing to think about is the cost–benefit balance of reducing uncertainty about the quality of assessment judgements. To put it another way: greater certainty that a judgement is a secure judgement comes at a cost, as I showed when writing about how reliability can be increased (in Section 1.5). The question that teachers should ask, especially those who feel swamped by work to be assessed, is how confident they need to be about the goodness of their judgements of particular learning outcomes in relation to particular assessment tasks. With some outcomes of learning and some tasks, certainty is hardly to be had and it is not a good use of time to fret about it. With others, notably tests of information and facility with algorithms, certainty is
cheap. And then there are mid-range outcomes and tasks, where greater confidence could be brought but the question ought to be whether it is worth doing so. The extract below, from the conclusion of a report I presented to the Assessment Working Group of the Engineering Professors’ Council, amplifies this thinking.

Assessment processes for different learning outcomes and tasks

(Report to the Assessment Working Group of the Engineering Professors’ Council, December, 2000)

The output standard for Engineering has rightly been formulated to express what Engineering graduates should understand and be able to do. It is recommended that they be assessed in one of three main ways:

1 There are some simple achievements, such as information retention, that can be cheaply and reliably warranted. There are others that departments can warrant with somewhat less confidence but at a greater cost. As far as some of the output statements go, HEIs can afford to publish fairly secure judgements about achievement. Notice that these are judgements about achievement in respect of some of the output standards.

2 Other attainments are costly or impossible to assess with reliability, unless they are subject to assessment processes that destroy their very essence. Departments can reliably attest that students have had plenty of engagements of the sort that should make for considerable achievement. They can – and should – also make reliable statements about the contexts of attainment, allowing stakeholders to see whether performances could be associated with high levels of support, structure and guidance or whether they are indicative of a good, professional engineer in the making. Departments can make highly reliable statements about process standards that should make for high levels of achievement in respect of output statements that do not readily lend themselves to summative assessment.

3 Students should learn about claims making, about making well-supported claims to achievement based on evidence collected and appraised throughout their undergraduate years. The implication is that they will create and learn to develop records of achievement (or portfolios) and become skilled at the self-evaluation necessary to use them in their claims making. This implies that portfolios should be subject to formative feedback, although it is hard to see that summative assessment would be feasible, ethical or prudent.

2.2 Good module-level assessment – necessary features

This section comprises twelve things that I claim to be fundamental to good module-level assessment practice. I present them as a list, amplifying some points with extracts from other authorities on assessment practices. It is followed by a similar section (section 2.3), listing practices that I think are desirable but which I consider to be less vital than the dozen covered here.

But before this, materials concerned with How can I prepare students for assessment? provides a context for this summary of necessary and desirable assessment practices at module level.
How can I prepare students for assessment?
1. There is a module assessment plan

It should say what the assessment tasks are; what their purposes are (local or generalisable, formative or summative); how summative assessments are weighted; and, ideally, it should close the circle by showing how module learning outcomes will each be assessed. Assuming that a module has half a dozen learning outcomes, there will need to be a variety of assessment methods in use.

You may need to explain the thinking behind the assessment pattern, particularly if your scheme is a distinctive one.

The following section entitled *What effect does assessment have on your students?* together with extracts from Miller and Parlett’s research and Snyder’s writing on *The Hidden Curriculum* talk about what happens when students do not have this sort of information and are left to try to work out for themselves what the ‘rules of the game’ are. Not only is this unfair to students and probably unethical as well, it also harms their learning in the sense that they are quite likely to interpret the cues wrongly and put effort into learning that the teacher neither intended nor wanted.

*What effect does assessment have on your students?*

I was interested in how rock sequences get laid down and that is what I thought the course was going to be about. Dr Ryan is into all this alternative methodologies and theoretical stuff. But I bet that the exam will only be about recognising rocks because that is all we ever do in the labs and they want to make sure we get good marks. (Student)

What is assessed is not always the same as what is taught or what teachers say will be assessed, and all assessment systems have the side-effect of changing what students pay attention to and learn. Students are perfectly capable of distinguishing between their own interests, their teachers’ intentions, how they experience the course, and what they think will count in assessment. In this ambiguous context of varied and sometimes conflicting goals they have to work out what it is sensible to focus on ...

*Cues about assessment*

The extract that follows, by Miller and Parlett, also explores the way that students respond to the assessment system. The authors interviewed and observed students, this time at Edinburgh University. They noticed that students differed in the extent to which they were ‘tuned in’ to the hidden curriculum, and that lecturers at Edinburgh differed in the extent to which they helped their students to tune in. Miller and Parlett also have quantitative evidence that students who were most finely tuned in, the ‘cue-seekers’, gained the best degrees. Clearly students’ awareness of the
nature of the hidden curriculum, and the extent to which they focus
attention on it, is important for their success. Your students will attempt to
find out what the hidden curriculum consists of, regardless of what you
do.

The practical implications of this work seem to be that you have to make
sure that students’ orientation to the demands of the assessment system
work in favour of your goals, rather than against them. Your assessment
tasks need to demand the use of the skills and knowledge you are really
hoping to develop, and students need to be clear about what is being
demanded. This may involve being explicit in your expectations, and
drawing the cue-deaf students into discussing and becoming more aware
of your expectations. If everything is explicit and congruent and students
understand these expectations and act on them then the curriculum will no
longer be hidden.

You may not have control over either the formal curriculum or the
assessed tasks – in which case your feedback when you mark students will
carry a heavy burden in explaining your marks and why students have
done well or badly. These explanations are their cues. They will frame the
way your students understand the curriculum they are working with and
will have a profound effect on how and what they study. You may need to
say not just ‘I’m impressed’ but ‘the reason I have given it an A rather than
a B is ... ’; and not just ‘try harder’ but ‘in order to get a B rather than a C
you would have had to ...’.

You have already (in Section 1.1) read a summary of Snyder’s work on the
hidden curriculum and an extract from his paper follows Miller and
Parlett’s. Despite studying very different universities in two different
continents with different cultures, and even with different assessment
methods, they were studying very much the same phenomenon and were
aware of this. These two extracts illustrate the value of seeking literature
from other higher education systems to provide insights into one’s own.
Before you read Extract 2.2 from Snyder, think about the extent to which there might be a hidden curriculum for the course you teach, what this might be based on, and what impact it might have on your students.

- Is there a hidden curriculum?
- What do you suspect your students believe will gain them good marks on your course?
- Do you suspect that your students are 'selectively negligent'? If so, what do you think that they neglect?
- What implicit messages might your assessment tasks, tests or assignments be providing about what students should spend their time on?
- When you give marks and write comments on students' work and hand it back, what cues are you giving to students about what really counts, either deliberately or accidentally?
- What do you think your students actually spend their study time on, and to what extent is this dominated by your assessment?
2 Assignment specifications are clear and full

The module assessment plan may tell students exactly what they have to do for each piece of work that will be assessed, summatively or formatively. If not, full specifications should be given to students along the lines of The specification for a summatively assessed assignment below:
3 Students will be clear about what counts as legitimate co-operation and what counts as plagiarism.

The following extract, *Are assessment practices immune from fraud?* has more to say about this. You may also want to refer back to Section 1.5, where a different extract from the same journal article informs the section headed *Balancing reliability with side effects*. The paragraphs that follow the Newstead extract suggest some strategies appropriate where a student appears to be nearing or crossing the plagiarism boundary.
EXTRACT 2.3
ARE ASSESSMENT PRACTICES IMMUNE FROM FRAUD?
S. E. Newstead

(Adapted from Newstead, Franklyn-Stokes and Armstead, 1996)
On the edge of plagiarism

'Two essays were almost identical'

Your institution and/or Department should have rules for students on plagiarism, and guidance for staff on what to do if they suspect plagiarism. These rules should deal with students copying each other's work. (They should also deal with copying from books or other sources.) Help students to understand from the start of the course what the institution means by plagiarism. As you do this you'll need talk about plagiarism sensitively but clearly: there can be cultural differences in what is required and expected of students, and the boundaries between co-operation and plagiarism are less clear than you might hope.

If you suspect plagiarism of any type, tell a senior colleague, perhaps the course leader, as soon as possible, and follow the published advice and guidance to the letter (this advice also applies to other forms of suspected student cheating).

'This essay is pretty much my lecture notes'

Some students may expect high marks for giving you back what you gave to them. After all, it must be good: look who it came from! Again, the solution is briefing. Talk to them about the relation between what you say to them in the lecture and what you expect from them. For example:

It's not enough simply to repeat back my lectures, however brilliant they are! I'm interested in the use you make of what I say. I want to see you relating what I say to what you read and research for yourselves, critically analysing the ideas you hear and read and making your own sense of them.

At the start of the course, talk to them about the relationships between your lectures, the readings, and the work they do.

If you receive work which parrots your lectures, give it a low mark and explain why.

4 Students know the criteria, indicators or standards that will be used in the assessment of each task

Where the assessments are local and formative, the indicators will often be fuzzy. Where they are local and summative indicators are likely to be more precise, which is also true of general, summative assessments. It is helpful for everyone to print the indicators for each task and give them to students. In the case of written, graded coursework, the grade indicators can be printed on the coversheet that students attach to the assignment. Examples can be seen in A cover/feedback sheet, below and in Extracts from a social science department's grade indicators, above (see Table 2.2).
A cover/feedback sheet

(Adapted from sheet prepared by the Psychology Unit at Oxford Brookes University by Dr R. Paton and Dr S. Fearnley)
5 Students understand the criteria, indicators or standards that will be used

It is one thing to present students with indicators, another for them to understand what you want the indicators to mean, especially where the indicators are local and summative. (If they are general and summative, then all except first year students should have had some experience of their meanings through their earlier assignments. If they are local and formative, then the whole point is for there to be conversations about how to interpret and apply the indicators.)

Help students to understand the meaning of indicators by: discussing them in class; showing them marked examples of work from previous years; doing mock grading exercises in class; using the indicators for formative peer-assessment.

6 Refer regularly to your indicators, marking schemes or model answers when you assess (or when students formatively assess each other)

You might go back to Section 1.5 to refresh your understanding of reliability issues before reading the following Extract 2.4 The psychology of student assessment, which draws again on Newstead’s research and contains some quite terrifying material on our normal inadequacies as assessors. It should be quite enough to make us swear to mark with assignment specification and grade indicators by our sides. Some advice on doing the marking, which follows the extract, gives just that!
Copyright material removed

(Newstead and Dennis, 1994)
Of course, when students are classified for a degree, the marks for more than just one essay are taken into account, and as the number of different marks increases so does the reliability of the overall mark. Hence the misclassifications of students overall will not be as alarming as the results on just one piece of work might suggest; nevertheless, the discrepancies between markers are clearly worrying, and indicate that our marking is less reliable than it ought to be.

Are standards consistent?

It is difficult to say definitively whether standards are consistent since we have no centralized standard to use in any comparison. Nevertheless, there are data relevant to this.

Over institutions?

Smith (1990) studied the variability of psychology degrees over different institutions. The proportion of good degrees ranged from 29 per cent to 77 per cent – and the differences between institutions were fairly consistent over time. There was no obvious reason for these differences, since they seemed to be relatively independent of resourcing and of entrance qualifications. It is difficult to escape the conclusion that different standards apply in different institutions.

This problem is not unique to psychology. Chapman (1994) has studied variations in geography and has obtained similar results.

Over time?

It is usually assumed that degree classifications remain constant over time – in other words, that a First Class honours degree means the same thing now as it did 10 or 20 years ago. Once again, the statistical evidence suggests that this might not be the case. The Universities' Statistical Record (1982 onwards) shows that the proportion of Firsts that are awarded nationally has increased by more than 50 per cent over the last 15 years from six per cent to over nine per cent. ... This increase has been fairly consistent over the years, but seems to be particularly dramatic since 1985.

Chapman (personal communication) has studied the number of good degrees in eight disciplines (not including psychology) and has reached the conclusion that the modal degree is now an Upper Second, not a Lower Second as it was just a few years ago.

Of course, it could be that students are getting better over time, or that teaching methods have improved to such an extent that students are performing much better than their counterparts in the past. But even if this is the case, it does not really explain why the proportion of good degrees has gone up so much at a time when the participation rate has increased so dramatically. About 30 per cent of 18-year-olds now enter higher education, compared to only five per cent some 20 years ago, and it is difficult to believe that the proportion of these modern-day students who merit a good degree is so much higher than the proportion of their much more highly selected predecessors.

Nor is the change in the number of good degrees particularly surprising. Universities and the departments within them are increasingly judged by the exam results they achieve, so is it surprising that markers sometimes err on the side of generosity?

Over disciplines?

Employers usually assume that a good degree (i.e. a First or Upper Second) means the same thing in different disciplines. Such an assumption is important since so many of today’s graduate employment opportunities are positions for which graduates from any discipline can apply.

Data from the Council for National Academic Awards database (e.g. CNAA, 1983) suggest that differences exist. For example, over the years 1971 to 1986, a student had a 50 per cent chance of obtaining a First or Upper Second in philosophy and history, but only a 30 per cent chance in accountancy. Recent data released by the Higher Education Statistical Agency (HESA, 1995) suggest that the problem might be even worse now. Just over 40 per cent of students in maths and education obtained good degrees, compared to more than 60 per cent in engineering and technology (see Table 2). These differences may be explicable in terms of intakes and resourcing, but it does seem unlikely that this can explain differences of such magnitude.
confirm this pattern, though data we have collected at the University of Plymouth (Hoskins and Newstead, 1995) suggest that there the difference exists only at the bottom end of the scale (see Table 4).

<table>
<thead>
<tr>
<th>Table 4 Percentage of Firsts and Thirds obtained by males and females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Rudd (1984)</td>
</tr>
<tr>
<td>Males</td>
</tr>
<tr>
<td>Females</td>
</tr>
<tr>
<td>THES (1995)</td>
</tr>
<tr>
<td>Males</td>
</tr>
<tr>
<td>Females</td>
</tr>
<tr>
<td>Hoskins and Newstead (1995) (Plymouth Students)</td>
</tr>
<tr>
<td>Males</td>
</tr>
<tr>
<td>Females</td>
</tr>
</tbody>
</table>

A related study by Yorke et al. (1995) has demonstrated that in the modular schemes of six different universities some disciplines were consistently marked more highly than others; for example, Sociology and English produced higher marks than Law in every institution studied (see Table 3).

<table>
<thead>
<tr>
<th>Table 3 Mean rankings of marks in eight disciplines in six institutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sociology</td>
</tr>
<tr>
<td>Fine Art</td>
</tr>
<tr>
<td>English</td>
</tr>
<tr>
<td>French</td>
</tr>
<tr>
<td>Business Studies</td>
</tr>
<tr>
<td>Maths and Statistics</td>
</tr>
<tr>
<td>Computer Studies</td>
</tr>
<tr>
<td>Law</td>
</tr>
<tr>
<td>2.3</td>
</tr>
<tr>
<td>2.5</td>
</tr>
<tr>
<td>2.8</td>
</tr>
<tr>
<td>4.7</td>
</tr>
<tr>
<td>4.8</td>
</tr>
<tr>
<td>5.8</td>
</tr>
<tr>
<td>6.0</td>
</tr>
<tr>
<td>7.0</td>
</tr>
</tbody>
</table>

(York, Cooper, Fox, Haines, McHugh and Woolf, 1995)

Note: A low ranking indicates that this discipline received relatively high marks

Are markers biased?

The most widely investigated bias in recent years has been that of gender. It is a highly publicised fact that males tend to get more Firsts than do females, but that they also get more Thirds – a trend which is particularly marked at Oxford and Cambridge. Rudd (1984) presented summary data which indicated that males obtained twice as many Firsts and considerably more Thirds than did females. Recent data from the Higher Educational Statistical Agency (THES, 1995) The controversial issue concerns the causes of this discrepancy. A number of explanations have been put forward, including the greater tendency of women to conform, and biological differences in the distribution of abilities in males and females. The explanation that is of concern here is that put forward by Bradley (1984) who claimed that it resulted from gender-related biases in marking. She claimed that markers expected women to perform at middle-of-the-road levels and that this expectation led to a bias. At the top end of the scale this was a pro-male bias – a tendency to give good male scripts a very high mark; and at the bottom end of the scale it was a pro-female bias – a tendency to mark down males who performed at this level. In other words, females are expected to perform at average levels while males are expected to perform either very well or very poorly.

Bradley investigated this by looking at disagreements between the marks awarded to projects by the supervisor and the second marker which crossed a degree class boundary. She argued that the supervisor would know the student's true ability and would be less likely to show the bias, but that the bias would be more obvious in second markers. She found, as predicted, that second markers tended to be more extreme than first
markers for male students, but not for females (see left hand column of Table 5).

**Table 5 Marker disagreements on projects**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>63%</td>
<td>32%</td>
</tr>
<tr>
<td>Females</td>
<td>41%</td>
<td>47%</td>
</tr>
</tbody>
</table>

This bias is disturbing, though we have evidence that it may not be as widespread as feared. In studying the marks awarded to psychology projects at Plymouth, Newstead and Dennis (1990) found no evidence of the bias found by Bradley (see the right hand column of Table 5). One possible explanation of these differing findings is that the publicity given to women’s issues and to the problems of bias in assessment during the 1980’s has led to the bias disappearing (Finn, 1994) – though there are of course a number of other possible explanations.

The bias that students worry about most is that resulting from the personal like or dislike of the marker. It has not proved easy to investigate this, but we have made an attempt through the use of statistical modelling techniques. Dennis, Newstead and Wright (1996) analysed the marks awarded to psychology projects at the University of Plymouth. Projects are given four separate marks and are assessed by the supervisor and a second marker. The structural equation model that we used enabled us to analyse which factors accounted for the observed variance in the marks. Our model used factors corresponding to: a) the ‘true’ merit of the project (i.e. factors affecting both markers); b) the first marker; c) the second marker; d) error. Approximately 30 per cent of the variance could be attributed to factors which influenced the first markers but not the second markers. There are a number of possible explanations for this, but the most likely is that this reflects the influence of personal knowledge of the student on the first marker’s mark.

This bias seemed to be more marked for male than for female students, and had the effect of elevating male marks relative to those of females. The conclusion we were driven to is that marking biases do exist, and that they are related to gender. In this, we agree with Bradley. However, in sharp contrast to Bradley, we believe that such biases are more apparent in first markers than in second markers and are based on personal knowledge rather than on gender stereotypes.

Some advice on doing the marking

Work your way through the pile of scripts

A large pile of scripts can be very intimidating. Before you tackle the whole pile, you may want to get a feel for the job of marking these scripts, for the standard of work they contain, and for how you will mark and grade these scripts. Here’s one way to do this.

1. Take the first five or six scripts from the pile.
2. Read them quickly, and pencil in a tentative grade.
3(a) Either: skim quickly through the rest of the scripts, and sort them into grade piles (A B C D E, or whatever grading system you are using). If you have time, take one script from each pile and get a second tutor’s opinion on its grade.
3(b) Or: skim through the rest of the scripts and put them, again very quickly, into rank order. Then work out where you think the grade boundaries lie. If you have time, take one script from either side of each boundary and get a second tutor’s opinion on its grade.
4. Mark all of the scripts in detail, using the marking scheme.
5. When you have finished marking, check again the first three scripts you marked under step 4. Check if your standards have drifted up or down as you marked. If you find discrepancies, you may need to adjust your marks for consistency across the whole batch. (Which ones do you adjust? Your marking scheme should help you decide which marks are the more valid.)

Using assignment attachment forms

I suggest above that you give detailed feedback that makes explicit the standards which you use to make judgements. However, giving such detailed feedback can take a long time. Here’s a way to give detailed feedback much more efficiently. It involves investing some time before you start to assess a set of work.

1. Collect and list the comments that you make on student work. These comments will include overall reactions to the work as a whole. For example, ‘You have addressed each element of the question’, or ‘The overall shape is unclear’. They will also include comments about particular points in the work, for example ‘Well argued’, ‘Good use of source’, ‘I don’t understand this part’, ‘Original idea’, ‘This calculation is wrong’.
2. If you haven’t marked a similar piece of work, spend a few minutes thinking about the question and noting at least some of the likely strong and weak points of student answers.
3. Also write up the overall requirements for the piece of work you are about to mark, grade or give feedback on. These may include the learning outcomes, the assessment criteria, and any requirements for structuring (for example headings or length) or presentation (requirements for introduction, summary or referencing).
4. Describe the overall qualities you would expect to see in the work: clarity, wide range of sources, sources well chosen, a critical or creative approach – whatever is necessary and appropriate.
5. Put all these together in a table, with space for ticks.
6 Add space at the top for the student's name, course, assignment details and any other essential data. Leave space at the bottom for individual written comments.

7 Print as many copies as there are pieces of work to assess.

8 As you mark, tick the general comments which apply to the piece of work you're marking.

9 For specific comments, place the number of the comment next to where that comment applies in the student's work; it's much quicker to write '7' than 'This is an ingenious and appropriate approach'.

[You have seen an extract from a cover/feedback sheet earlier in this section, and your institution or discipline may have other examples, if you ask]

You can also use a model answer as an attachment, perhaps commenting on where the student's answer followed or diverged from this model answer. The applicability of this approach will vary among subjects.

Give oral feedback

Speech can lead to dialogue, in which understanding can be checked. Speech can offer a direct, immediate, personal communication. We can generally talk faster than we can write.

For these reasons, face-to-face oral feedback can be very useful, for example in a tutorial, during lab work or in studios and field trips. Most of the principles described above apply equally to oral feedback, especially being specific, being honest, referring if possible to the student's previous work, showing respect, starting and finishing positively, and making suggestions for improvement.

You could try this four-step approach to oral feedback:

1 Give the feedback, following the principles described above.

2 Encourage the student to ask questions about your feedback, to ensure that he or she understands it.

3 If you feel it necessary, ask students to paraphrase what you said to them. This provides assurance that they heard and understood accurately what you said (whether or not they agreed with you – we'll come to that next).

4 Finally encourage them to react to your feedback, building on what you said, exploring how they will use what you said for future work, questioning or disagreeing with your feedback, and offering a clearer account of what they meant.

This is formulaic, but it's an effective formula. It may sound long-winded, but with practice it can be done in a couple of minutes. And it makes the student appropriately active, which boosts the value of your oral feedback.

Be sensitive in the way you give feedback:
• different students have different tolerances for feedback, whether you are giving praise or making suggestions for future improvement
• students also vary in the extent to which they are happy receiving oral feedback within earshot of other students.

7 Justify grades or feedback and do it in terms of the declared criteria indicators or standards

If you don't, how are students to work out what is being rewarded and what is limiting their performances?
8 Mark high-stakes assignments as fairly as possible

This extends point 6. In these cases, second marking is normal. Second markers should be particularly inquisitive about the match between the suggested grade, the first marker's comments, the student's work and the grade indicators. All four should be pointing in the same direction.

Note that fewer and fewer universities expect external examiners to adjust marks or resolve disagreements between internal markers. My experience is that it is quite hard to see what external examiners are expected to do and harder still to see what they can do about practices that disturb them. Some policymakers were suggesting in 2001 that new quality assurance regimes ought to be policed by an enhanced, more professional external examining cadre some 10,000 strong.

9 Give timely, clear, useful feedback that relates to the marking indicators

This ground is covered by Giving feedback, Extract 2.5, Written feedback on student writing and Internalising feedback (see below).

I want to emphasise two points. First, there should be clear — that is to say specific — advice on how to improve on similar tasks next time. The clutter of nit-picking comments matters less than one or two headline points about improvement. Secondly, say less rather than more. Concentrate on the headline points because saying too much can obscure them and be overwhelming as well.

Giving feedback

Guidelines on the content and style of feedback

Be strategic in your feedback — focus on those comments which will have maximum positive effect.

How do you decide where to focus your feedback to your students? You need to take account of:

- what you know of each student and their present stage of development
- what good qualities and what confusions in their work are the most important at this stage
- what feedback is likely to have the greatest good effect now.

This is rather general advice, but we can now go into more detail.

Make the feedback specific

'Good' simply offers encouragement, whereas 'Clearly expressed, well argued, excellent selection of sources' say why the work was good.

If you can say why you feel the work has each of these qualities, and what your criteria or indicators are for clarity of expression, quality of argument and excellence in selection of sources, so much the better. For example, you might have expanded this feedback by saying:

From a long reading list you chose sources which contrast well with each other. You explained why you choose some sources and not others. You justified your own views, showing how you arrive at them. You gave accurate accounts of views with which you disagree. In the more technically difficult passages, you reduced the sentence...
length considerably, which aided my understanding. You also helped me to follow your argument by giving periodic reviews of what ground you covered, and indicating where you would take the argument next.

This approach, of course, greatly increases the length of your feedback commentaries. One way to manage this is explained later.

**Make the feedback honest**
Tell it as you see it – as long as you also follow the rest of the advice given here!

**Clarify the status of the feedback**
If you are expressing your own view, say so: ‘I didn’t find this at all clear, because it ... ’.

If you are stating a fact, source it: ‘On this topic Smith actually says ... ’ (give the reference).
Feedback, as well as the students’ work, should be academic writing.

**Make the feedback personal**
If possible, make the feedback refer to their previous work. Your records should show you how this student did in their previous piece of work with you. Refer back to this:

I found your presentation much clearer this time. You were also much more fluent in your handling of quotations, partly by selecting short quotations and partly by weaving linking comments among them. However, you still need to work on the narrative – I still found this rather jerky, especially in the latter half – you need to find ways of connecting sections of the essay and showing the links. Section 1 of Chapter 9 of Smith (reference) does this very well – see how she does it, then try to find a way that works for you.

**Show respect**
You are giving feedback to another human being. Treat them like one, however many suggestions you may have to make for future improvement to their work.

**Write feedback which the student will understand**
Your understanding of, abilities in and vocabulary about the discipline are more sophisticated than those of your students. Don’t try to move the students from where they are to where you are in one step. Instead, build on things they have shown that they know and can do.

You pointed out one instance where the theory predicts an outcome which you didn’t find in practice. You could also have taken this approach with the other theories you considered, relating them to what you discovered ...

**Write feedback which will help the student to increase their understanding in the discipline**
Giving feedback should be part of your teaching. It may be the only part of your teaching where you can work one-to-one with your students. You need to lead them forward at a pace which they can handle.
Your approach to the question was mainly descriptive – the question called for analysis. Analysis here means asking, and trying to answer, questions such as “Why might these problems occur?” and “What different models or theories might be used to account for them?”.

**Say why as well as what**

Your paragraphs are very long, sometimes a page or more. These long paragraphs often contain several separate ideas – I’ve indicated these on page 5. Try using shorter paragraphs. This will break up the look of the paper, and will signal the transition from one idea to the next.

**Respond to what you see in the work**

You may have hypotheses about how and why the student has gone wrong in a piece of work. However, you should start by reacting to what you see in their work, not with what you think may be going on inside their heads.

I was surprised that you made no reference to last month’s field trip in this paper. I feel that this seriously weakened the paper, and left your concluding assertions dry and not backed by evidence.

**If you need to guess why the student has done something, make your guess explicit**

It may be that you thought that this paper had to be written without any reference to your field work. In fact the brief for the paper did ask you to review your field work experience before producing it.

**Guidelines on structuring your feedback**

**Start with positive comments**

There will always be some positive things to say. Find them and comment on them.

**Describe problems or errors in the work, always with ideas on what the student can do about them next time**

For example:

Your initial calculation of flow conditions contained a serious numerical error – you didn’t show your working in enough detail to enable me to see how you made this. Having made this error you were bound to select the wrong model. Next time: (1) Be extra vigilant when the initial steps in a problem clearly affect everything that follows; double-check, work it another way; (2) Try to visualise the physical situation described in the problem. If you had done this here, you might have seen that the flow would be turbulent rather than laminar, and hence that your answer was fundamentally wrong.

**End with an encouraging summary**

Your choice and use of models (after the initial serious error) were sound, and your reasoning and calculations accurate.
**Guideline on the timing of feedback**

**Turn student work round fast**
Students move on – to the next topic, the next assignment, the next course, the next term or semester. They need feedback quickly, in time to use it. Put feedback time into your diary the day after the work is due to reach you. Get the feedback out to the students within a week of receiving it – sooner if at all possible.

**Use feedback for a range of functions**
You can use feedback to correct students’ mistakes, to improve writing and presentation, to clarify the nature of tasks, and to give encouragement.

**Correcting mistakes**
You should say as exactly as you can what is wrong and where the errors lie, and show a correct answer or approach, with a source in the course materials.

You have confused induction and deduction here. Induction means using individual observations to build a theory or model; deduction means drawing specific conclusions from a theory or model. The handout for lecture 7 gives more on this.

You’ll need to do this with sensitivity – too much of it at once can be overwhelming. If there are many errors, you may feel you should concentrate on those which are the most important.

**Improving writing and presentation**
I found this difficult to read. At a technical level, you should use a larger font – perhaps 11 point rather than 8 point or 9 point. You also need to make it a little more formal – it read to me like transcribed speech. I’ve worked through one paragraph to show the kinds of changes I am suggesting.

**Clarifying the nature of the task**
You were asked to discuss Monet’s work in relation to that of his contemporaries. In your essay you wrote about Monet and Degas, and made some interesting comments on their work. However, you only considered one other artist, but the question called for more than one by saying, ‘... his contemporaries ...’. And you did not make any links between their work, as the question asked by saying, ‘... in relation to ...’.

**Giving encouragement**
This work shows real improvement over your previous report. You have got to grips with the subject matter, and you have substantially overcome your previous problems with linking and relating ideas.
Mary Lea was specially commissioned to write this extract. Further details of the case study can be found in Lea, M. and Street, B. (1998) 'Student writing in higher education: an academic literacies approach', Studies in Higher Education, 23(2), pp. 157–72.
Marking and giving feedback involves a great deal more than simply correcting students’ work and allocating a mark. This chapter has raised issues concerning how reliable your marking may be, how students interpret and respond to your feedback, and how students respond to cues you may provide about what counts in terms of marks. They present assessment in its social and organisational context and explain how your feedback is interpreted by students in relation to their personal aims and their understanding of what is required. Students’ focus of attention is dominated by assessment and the way they perceive its demands and you have a central role to play in influencing their studies through the way you mark their work and give feedback.

**Internalising feedback**

Students may read and even remember feedback, especially general feedback about how to tackle a particular kind of task, but still not act on it:

*(Hounsell, 1987, p. 116)*

It is a common experience of teachers to find students making the same mistakes again and again despite frequent feedback on the same point. In lab report writing students can do things like failing to label the axes of their graphs for years on end. If you asked them to give some sensible advice about writing lab reports they would tell you to label the axes of graphs – they just don’t do it themselves! Students often fail to internalise such advice and make it part of their own internal quality control. They hand in work which they have not checked, leaving all the quality control up to you as their teacher. Hughes (1995) describes an innovation in assessment on a Pharmacy course. He had found that his own feedback on his students’ weekly lab reports had little impact on their average marks; giving feedback appeared not to be very effective. He introduced a system in which students marked each other’s reports from the previous week at the start of the next lab session. This greatly increased the quality of student work. It was not that student feedback was more accurate or comprehensive than his own had been, but in marking others’ reports, students had internalised what makes a good report and applied this to their own work. Even copious and careful feedback from their teacher had not resulted in this internalisation of criteria and internal quality control.

If you asked your students how much time and attention they gave to the written feedback you write on their work, what might they say? How carefully do they note your points and to what extent do they act on them in their next piece of work? How well do they understand your comments? What kind of comments do they most value? How do you know?

10 If you assess group products, make sure that students understand the rules of this game

Will each group member get the same mark? If not, on what basis are differential marks going to be calculated? Will ‘easy riders’ get 0? Assessing groups, below, contains useful advice here.
Assessing groups

Assessing individual students can encourage competitiveness and discourage collaboration: students may avoid sharing books or discussing their work for fear of others benefiting. In extreme cases students may hide library books or behave in other socially dysfunctional ways. But students can learn a great deal from each other through peer tutoring ... and by working on joint tasks or projects. Groups can share resources such as computers, equipment and books. The performance of students in teams almost always exceeds the performance of students tackling the same tasks on their own. Groups require less supervision than the same number of students working alone, and assessing groups can save a great deal of time. The main problem is that individual students need marks that reflect the quality of their own learning, not an average of the quality of the learning of others. And if students know they are going to get the group’s mark regardless of their own effort, they may not contribute, thus lowering others’ marks and boosting their own marks in an unacceptable way. However, there are some simple and well documented techniques that enable teachers to set group learning tasks and assess groups with confidence. The main alternatives are:

- observing the group members in action and moderating group marks accordingly (this may be possible in laboratories but not for independent work);
- dividing the assignment so that students are each responsible for a component and gain 50% of their marks for that component and 50% for the whole piece;
- multiplying the mark you give by the number of students and asking the students to divide these marks among group members in a fair way and with unanimous agreement;
- asking each student to rate the relative contribution of each other member of the group to each component of the assignment (for example the literature review, the data collection, the analysis, the write up) and using these ratings to moderate marks;
- setting a ‘project test’ which asks questions about the group assignment under exam conditions;
- using vivas to moderate the group mark.

There is research literature about the impact of these alternatives on both student learning behaviour and on their marks (for an overview see Gibbs, 1995).

Case Study: Assessing group work

In this case study some imaginative methods are used to apply social pressure to get students to work effectively in groups and to make it in their interest to succeed as a group ...
('Group reports and group presentations in a special study module' in Hounsell et al., 1996, p. 45)

11 Peer- and self-assessment are valuable, not least because they help students to internalise assessment criteria

(This is illustrated in Assessing groups, immediately above.) Provide opportunities for students to practise it in your module. I advise you to side-step a long and, I think, tedious literature on the reliability and fairness of peer-assessment by using it mainly for formative tasks.

12 Design for depth, not for breadth of content coverage

You have already read the claim that student learning can be damaged by modules trying to cover too much ground, a point also made by Chambers (1992). It is better to cover less material but with a strong emphasis on understanding, application and the development of reflection, social practices and metacognition.

You must also assess accordingly, which is likely to involve what are known as 'authentic' assessments of achievement (Cuming and Maxwell, 1999). These are worthwhile, complex tasks that are as true to life as possible and which are likely to shed light on more than one learning outcome at a time. The specification for a summatively assessed assignment (above, in this section, 2.2) can be regarded as an authentic assessment within a social science research methods course. In Chapter 10 of Being a Teaching in Higher Education (Knight, 2002a), I provide a more complex example of organising one course's assessment showing how a single task can engage with several learning outcomes.
Reflection 2.2 Advising a new colleague on managing module-level assessment tasks

I would like to encourage you to draw together your understandings of the considerable amount of advice with which you have engaged so far in this chapter.

Tasks that encourage 'deep' approaches to learning (see Section 1.3) encourage learners to re-work the material available to them, not simply to re-present it. With that in mind, I invite you to review Sections 2.1 and 2.2 as if you were going to give advice to:
(a) a new colleague
(b) about teaching in your subject area
on
(a) assessment practices that you recommend them to follow
(b) assessment 'traps', things that you recommend them to avoid.

I suggest that you restrict yourself to advice that would fit on one side of paper, concentrating on the most important things you can say.

The greatest value is to be had from this activity if you talk over your ideas with a colleague in your programme team, or a fellow participant in any teaching and learning programme you might be engaged with. (The same principle applies to student learning, too.)

2.3 Good module-level assessment – desirable features

This section follows the same format as Section 2.2. By designating this a collection of 'desirable' rather than 'necessary' features I am being rather arbitrary and it is quite likely that you would want to see some movement between the two lists. Reflection 3.1, near the beginning of Chapter 3, invites you to consider prioritising your own principles for good module-level assessment. The following list of eleven things I value supplements Section 2.1 and may inform your developing values.

1 Efficiency

From Knight (2002a, p. 158), I provide 10 suggestions for speeding up feedback, below. You may notice that some of these reiterate practical suggestions you will have encountered in Section 2.2 and reinforce the systemic or holistic approaches recommended throughout this pack.

Creating faster feedback

1 In good learning cultures, ones in which they know the 'rules of the game' and understand the criteria to be applied, students are less likely to make a complete mess of assignments, meaning that there are fewer occasions when massive feedback and coaching are necessary.

2 So too when students have worked collaboratively on projects and conversed with one another about drafts.
3 In good learning cultures, students know the grading criteria because they are printed on assignment cover-and-feedback sheets. Again, this helps to reduce the incidence of badly wrong work.

4 Cover/feedback sheets can speed up feedback when students have to identify the indicators that best describe their work (when they have to assess themselves). Sometimes the teacher needs to write little more than, 'I agree'. Having an idea of the student's judgement of an assignment can also make it easier to give feedback because it precisely identifies any gap between the teacher's and the student's judgements; feedback can be concise because it is targeted.

5 Limit what you say. Most people find up to three major suggestions enough to deal with. Cover/feedback sheets can encourage concision by restricting the space for comment.

6 Consider creating a bank of the feedback statements that you frequently use and then draw on it when you give feedback.

7 If there are lots of errors, mark only the first page and then return the work for correction. It saves your time and it forces students to think about how to do better.

8 Sometimes, rather than explaining exactly what's wrong, it's better for student learning to direct them to sources where they can find out for themselves.

9 How far can peer assessment be substituted for tutor assessment? (But bear in mind the time it may take to set up the peer assessment in the first place.)

10 Can self-assessment replace routine second marking?

The following section entitled Submission and marking of assignments suggest that some different efficiencies may be possible with the use of electronic submission and marking systems.

Submission and marking of assignments

From a student's point of view, the option of submitting assignments electronically can be very attractive. If work is produced in electronic form, it makes sense to send it that way, too. Electronic submission can mean very simply: a file attached to an e-mail message sent to the teacher. It can also be part of a more complex system which is perhaps linked in to an institution's student records.

For you as a teacher, there will be some advantages, for example, legible writing, but there will also be disadvantages. Depending on your precise administrative and technical set-up, the constraints you impose on your students, and your own working style, you may have to deal with the following:

- converting computer files between different formats;
- moving incoming assignments into appropriate folders;
- printing out assignments;
- marking on paper and then adding comments to the electronic version.
Questions of authenticity are always raised in the context of electronic submission, but the problem is not very different from what happens in non-electronic submission. Plagiarism is recognised as a potential hazard, and the usual vigilance must apply.

In terms of automatic marking, computers are used mostly for objective tests, that is to say tests in which each question is designed to have an unequivocally correct answer. Objective tests are particularly good for informal self-assessment to give students regular feedback about their progress. Approaches to computer-assisted assessment (CAA) fall into two categories:

- automated marking of paper forms, using an optical mark reader (OMR) - questions are presented on paper, and students complete an OMR-readable answer form;
- computerised marking in which questions are presented on a computer screen, students make their response, and receive feedback and marks.

The main benefits are high reliability in the assignment of marks compared with human marking, faster turn-around time, opportunities for repeated practice based on randomised questions from large question banks, and built-in analysis of student responses. Problems associated with CAA include difficulties in phrasing questions such that there is an unequivocal answer, and 'surface learning', that is, learning of surface detail or mainly through guesswork.

A final word about marking with C&IT. If your students are doing some computer-marked tasks or questions, be sure that you give them some additional feedback that will help them to interpret their results. Marks and feedback produced by a computer can appear rather stark and may demotivate some students.

2 Mesh your assessment plan with the programme specification and its assessment plan.

Students benefit because they can see how your module contributes to the programme and, if they wish, choose it because of the learning outcomes you will promote and assess. It is good for programme coherence.

3 Capitalise on learning in other modules.

For example, if students learned to do oral presentations in Year 1, consider (i) using modified versions of the Year 1 assessment sheets, (ii) getting them to prepare for your assessment by recollecting their Year 1 experience.

4 Where possible use programme grade indicators.

For example, the piece entitled The specification for a summatively assessed assignment (which appears as a Box in Section 2.2) contains extracts from the default indicators used to assess written work in one department. Students benefit from this because they get to understand better what the indicators mean and they should benefit from feedback on work done in other modules that was written in terms of those indicators. You should benefit from this too, not least because you need to spend less time familiarising students with the standards that you will apply.
5 If possible, follow the programme policy on the balance between formative and summative assessment.

Innovators often meet resistance from students, who rightly wonder what the benefits of difference are. If there is a good programme assessment policy, explicitly align your module with it. If your module is distinctive – perhaps you have a lot of formative tasks leading to one or two graded, authentic tasks – then you will need to help students understand the thinking behind your unusual practice.

6 Print the university’s rules for calculating degree classifications

This allows students to see where they stand and what they might aim to achieve.

7 If you dare, set tasks that get students to apply and transfer their learning to novel problems.

While there is a lot of agreement that novel tasks are good ways of assessing what students have really learned, there are two forces working against this assessment strategy. First, these are more demanding tasks and marks are likely to drop. This looks bad, unless novel tasks are used formatively. Secondly, learners tend to try to bargain with teachers to re-cast demanding, novel and ambiguous terms into less ambiguous and easier forms. They want tasks where they know that hard work following clear rules will get good grades (Doyle, 1983). Expect student resistance unless your colleagues also set authentic tasks like this. (For an example of an institution that runs on authentic, complex tasks, see Mentkowski (2000) on Alverno College, Milwaukee).

8 Make space for students to reflect on their learning achievements and needs.

In Chapter 4, I will argue that some outcomes of learning should not be warranted or certified by the higher education institution but should be claimed by students who can back up their claims through portfolio evidence. Good module assessment practices encourage students to link their learning with their ongoing portfolios.

9 Consider allowing students to nominate which pieces of graded work are to be counted in classifying the module grade.

The box below summarises one approach to flexible weighting of assessment components.
10 Evaluate your assessment arrangements, possibly as a part of your annual module review.

There are helpful suggestions in *Learning about and from your assessment*, below.

**Learning about and from your assessment**

*Identify relationships between marks and feedback*

As you mark and grade and give feedback, collect together the marks or grades and the associated comments or feedback, then identify and check the relationships between your feedback and your grades.

Ask yourself:

- Which kinds of comments do I tend to give with which marks or grades?
- Why?
- What does this show me about what I value in student work, and about what errors and failings I think are important?
- Have I got this right? Am I valuing the appropriate things in students’ work?

This way, you can continue to refine and improve your assessment.

A mark or grade is also a form of feedback. Students sometimes put a fair amount of energy into interpreting written comments to see what mark or grade they imply. You should therefore:

- provide an indicative mark or grade with your feedback, even when a mark or grade is not formally required and will not be carried forward to compute their overall performance (make sure you label this mark or grade as indicative rather than actual, or students may become confused);
- alongside the mark or grade, tell them why they gained this mark or grade, and suggest what they would have to do to raise this mark or grade in the next piece of work.

*Learn from your assessment*

You can’t help learning from the assessment you do. Here are some ways to increase that learning.

**Overall learning from the marks and grades you award**

Overall marks and grades don’t give you detailed feedback on your assessment or on your teaching. You need to see which questions were answered well and which poorly. A simple table or plot showing the number of answers to each question gaining each mark or grade will do for a start.
Table 2.3  Distribution of answers by grade

<table>
<thead>
<tr>
<th>Grade or mark band</th>
<th>A (100–70)</th>
<th>B (69–60)</th>
<th>C (59–50)</th>
<th>D (49–40)</th>
<th>E (&lt;40)</th>
<th>Number of students answering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question</td>
<td>1</td>
<td>12</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>11</td>
<td>7</td>
</tr>
</tbody>
</table>

(Adapted from Moore, 1993)

Why did a third of those answering do so badly on Question 1 whilst half scored a respectable B? Why did students perform so spectacularly badly on Question 3? Why did so few attempt Question 2? More generally, what is the overall grade distribution?

As you begin to analyse the questions themselves, and the queries arising from your analysis of students’ responses, you will discover which questions about students’ marks and grades and which consequent lines of enquiry are particularly useful for you.

You should investigate assessment each time you undertake an assessment. This can be a simple process. You could:

- ask students why they found one question more difficult than another;
- ask students why they think they gained the mark they did for each question, rather than a higher or lower mark;
- where they say they didn’t understand a question, ask them to rewrite the question to make it more readily understood;
- ask them what a better question would have been on particular topics, and why.

Discuss with them criteria for assessing questions (assessing the questions themselves, not the answers). These criteria for a good question might include clarity of expression, clarity about which section(s) of the syllabus it addresses, validity, scope for allowing students to express and defend their own views, etc.

Learning from the feedback you give and students’ reactions

As you mark or grade or give feedback, note what the students do well and note the main errors and confusions. (Your feedback to students should include positive comments as well as comments on where they have gone wrong and on what they should do differently next time.)

Keep copies of your feedback. After a while, clusters of your comments on learning and of errors and confusions will become apparent ... Analyse the feedback that you give. Ask questions such as:

- On what elements or aspects of their work did I: make positive comments?; make negative comments with suggestions for future improvement?
- What comments do I make often?
- Into what clusters or patterns do these comments fall?
• What did these clusters show about the elements or aspects of student answers which I value as I mark?
• Are these really the most important aspects of students' work?
• What else could I profitably be commenting on?

Of course, this analysis won't tell you on what aspects of students' work you don't comment. That would take longer and be more difficult, though it would also be very informative. Do it if you have time.

**Ask students if they find your feedback helpful**
You could do this in a class discussion, or with a simple questionnaire, or with short meetings with a sub-set of the class. You could ask, ‘Did you find this feedback helpful?’, ‘Which parts of the feedback were helpful/not helpful, and why?’, or ‘What use did you/will you make of the feedback?’

**Ask them how you could make your feedback more helpful to them**
You could use the same methods.

**Act on as many of their comments as possible**
They will learn that you take their comments seriously. Your feedback will become progressively more useful to them. You will become better at giving useful feedback.

Where you can't act on one of their suggestions, tell them why.

**Learn from colleagues' feedback on your assessment**
Ask a colleague [who might be the person who second marks the work you assess summatively] to look at some feedback or marking which you have done.

Ask them what they would have said about the work, or how they would have marked it.

Where they differ from you, explore why.

**Play fair with students**
In part this just repeats yet again an important message about making sure that everyone knows 'the rules of the game'. More is said about fairness in *Maximising fairness* (below).

Why have I put fairness in the 'desirable' section and not in the necessary one? My thinking is that fairness is necessary but we can only approximate to it. The standards set in *Maximising fairness* are high. They should be heeded but will rarely be attained.

**Maximising fairness**
Ideally, assessment systems should be equally fair to all students, but unfairness can arise for many reasons.
• Exam questions may vary in predictability or difficulty, producing very different average marks and mark distributions. Students' marks may depend more on their choice of question (or even their choice of revision topic) than on what they have learned.
• Where students are undertaking project work or are in work placements, the difficulty of the task and the opportunity to do well in the context may vary considerably between students.

• Some forms of assessment suit some students better than others. In particular, some may perform poorly in time-constrained exams but perfectly well on extended open-ended assignments such as projects. The type of assessment you use and the weighting you place on these components will make a great deal of difference to who succeeds on your course and who does not. The extreme demands made by some forms of assessment can be moderated, for example by using open book exams or seen exams (see Section 1.2).

• Forms of assessment which have a social component or which place an emphasis on assertiveness, such as seminar presentations and group project work, may disadvantage women or students from educational or cultural backgrounds where these skills are either not practised or not valued.

• Some assignments that are expensive in materials, equipment, access to computers and the internet, or in travel and subsistence costs may disadvantage students who are less well off.

• Some assignments require extended periods of time in a library or a laboratory, or require blocks of time away on field trips or work placements. Part-time students and those living at home or with family responsibilities may find such demands difficult to cope with.

• Assessments on Fridays, Saturdays, Sundays or other holy days may be problematic for particular religious groups.

• Some forms of assessment assume that students are mobile, that they can view detailed visual material without aids, that they can write legibly and at speed, and so on. These assumptions can disadvantage students with disabilities.

• Where students are divided into sub-groups for studio work, labs or seminars, their teaching schedules may differ considerably. Some groups may have been given less time to complete assignments, and less warning that books may be required from the library. They may be taught at less congenial times of the day – and perhaps by less fresh (or even less able) teachers. Any of these factors might affect their chances of doing well.

• Where students' work is shared out between teachers for marking, some markers may be tougher than others, or may have specific preferences that students not in their seminar or supervision group would be unaware of.