Chapter 3 From module to programme

I return to the theme that good assessment plans ensure that there are assessment processes associated with each learning outcome by looking more seriously at how learning outcomes are written. My point is that assessment practices and learning outcomes should be aligned, which means that good assessment is related to the design of appropriate learning outcomes. That idea connects modules and programmes, because it is increasingly expected, at least in the UK, that modules will contribute purposefully to programme goals. So, when the assessment arrangements for all the modules in a programme are audited, it should be seen that they combine to provide a systemic, progressive and thorough coverage of the goals set out in the programme specification. We need, then, programme assessment plans: partly to keep track of differentiated assessment arrangements; partly to check that they are suitably orchestrated to promote programme learning goals; and particularly to make sure that we can make tolerably reliable judgements of those complex outcomes of learning, such as professional competence, that we choose to warrant.

This last point, explored in Section 3.3 and building on ideas introduced in Chapter 1, is important. If we wish to warrant or certify complex outcomes of learning, then we need to see this as a programme-level matter and ensure that the programme contains enough varied opportunities for assessors to observe and judge achievement in respect of these complex outcomes. Here I am making a lot out of a claim I have developed earlier, namely that it is one thing to make careful local assessment judgements – Reflections 2.1 and 2.2 encouraged you to capture features of good practice in making local judgements (at module level) – but it is quite another to generalise about achievement by warranting or certifying that someone is competent or has mastered something. I shall return to this theme in Chapter 4.

3.1 Good module-level assessment practices

Here is a partial summary of the suggestions for good module-level assessment practices in Chapters 1 and 2.

1. Make module learning outcomes clear.
2. Consider drawing most module learning outcomes from the programme specification.
3. Look for a good range of assessment techniques to match the range of module learning outcomes.
4. Blend formative and summative assessment purposes, using the notion of passports to summative assessment to underpin formative tasks.
5. Try to provide practice tasks as well as developmental ones.
6. Develop grade indicators. Sometimes indicators will be precise, sometimes fuzzy; sometimes they will make fine distinctions between levels of achievement, sometimes they will be broad brush.
7. Ensure students have the grade indicators and understand them. Good teaching creates ‘knowing’ students.
8. Use grade indicators when marking and giving feedback (and expect students to do likewise in peer- and self-assessment).
Reflection 3.1 Good module-level assessment practices

I emphasise these eight points but it is likely that, as you engaged with the text and Reflections 2.1 and 2.2, you constructed your own distinctive understandings.

If so, you might find it useful to review what you have read to date and jot down your priorities for good module-level assessment. This is intended to be an ‘advance organiser’ for what is to come. (Advance organisers are thought to improve understanding of what follows because they get readers thinking and ready to engage with text, rather than adopting a more passive text-scanning approach.)

Your notes will also provide you with a reference point when you next come to revise or devise a module.

These suggestions turn out to stimulate a variety of practices because the teachers who respond to them have different amounts of time, a range of academic purposes and work with different students, resources and colleagues. Together, though, they look like comprehensive advice on good assessment practice. They certainly summarise most of the phase 3 advice in Figure 1.1. So, that’s assessment done and dusted then? (Always given that putting suggestions like these into practice demands a lot of effort and skill and is likely to take several years of tinkering until you work out a detailed approach that students can settle into.)

In a very real sense, the suggestions contained in Chapters 1 and 2 are good enough. Students who take a module operating on these principles are likely to get a pretty good educational deal: academically stimulating, varied and fair. Yet you will have seen that I have a great deal more to say about assessment, mainly about assessment and whole programmes, which you might feel has little to do with mainstream teaching. My response is that programme-level issues matter a lot to teachers. One reason has already been mentioned, namely that teaching is increasingly understood to be about ‘design-and-delivery’. A new emphasis on the undergraduate experience as a whole means that module design has, increasingly, to be dovetailed with programme design: to be a ‘constructively aligned’ (in the words of John Biggs, 1999), part of a coherent, consistent and progressive programme. This makes programme assessment issues the concern of all module teachers, because the module–programme relationship is becoming more tightly coupled than has often been the case. As I try to develop this line of thinking I need to revisit some of the issues covered in Chapters 1 and 2, beginning with those surrounding learning outcomes.

3.2 Learning outcomes: first time around

Although I have mentioned learning outcomes, I have skirted around them. Since I shall re-visit them in Chapter 4, this section is my first serious treatment of them. I am going to concentrate here on the immediate, practical concern with creating learning outcomes at module level. I am going to treat this topic conventionally and put my reservations about the phrase ‘learning outcomes’ to one side for the moment. (I have already said that I am easier with the phrase ‘outcomes of learning’.) My concern here is that you should be familiar with mainstream approaches to developing learning outcomes, mainly at module level but, as the following extracts show, with some concern for programme intentions. This lays the way for a more critical view of learning outcomes in Chapter 4, which paves the way for my review of the (very problematic)
assessment of competence and of the design of programme-level assessment plans.

This paragraph previews the five extracts that carry most of the ideas on aims and learning outcomes in this chapter. You might wonder what this has to do with assessment. Although the fifth piece tackles this head-on, I want to make the point here that good assessment — and reliable assessment in particular — depends upon having good grade criteria or indicators. They will come from course and programme goals, which are conventionally written as learning outcomes. What is assessed, how, when, where and by whom all depend on what sort of learning we envisage coming from instruction and tasks, modules and programmes, and the wider learning environment. The first extract (Extract aims'). Notice the phase 3 (Figure 1.1) assumption that tutors devise their own aims for their modules. Notice too that the author talks of 'objectives', which he seems to see as elaborations upon the aims.
EXTRACT 3.1
AIMS AND OBJECTIVES

P. Ramsden

... The purpose of expressing aims and objectives is to improve the quality of education, in two senses. The activity should enable teachers to think more critically and deliberately about student progress, and the manner of its connection with what they do in their teaching. Secondly, the results of the exercise should make clear to students exactly what they have to learn to succeed, and what they can leave aside. Aims and objectives are no magic wand, but neither are they the dangerous witchcraft that some people seem to think they are. They will not necessarily cause teachers to reflect on their teaching or improve their students' learning, but they may be a useful technique for making those ends possible.

As in so many other things about teaching in higher education, it is very unwise to be dogmatic about the techniques of using aims and objectives. The principle of being clear about the key elements of competence that students should acquire is what matters. Objectives do not have to consist of 'things that students can be observed to do', as some proponents of behavioural objectives would try to have us believe. It is quite acceptable to think directly about the concepts that a student who has successfully completed his or her course will have understood – such as 'opportunity cost' in economics, 'the Avogadro constant' in chemistry, or 'metaphor' in literature – and how they will have arrived at an understanding of them (see also Rowntree, 1981, pp. 68–85).

Aims are best thought of as general statements of educational intent, seen from the student's point of view, while objectives are more specific and concrete statements of what students are expected to learn. The medical course at the University of Melbourne, for example, includes the general aim 'To achieve an understanding of principles in the analysis of human behaviour and functioning relevant to health and disease'; within the list of objectives for the subject of behavioural science which forms part of this programme is a specific example of what this entails. 'The student should understand [f]amily structures and their impact on patient care, particularly in regard to primary medical care.'

The rationale for using statements of aims and objectives would seem to be based on three linked assumptions:

- That education is about changes in students' thinking and knowledge.
- That it is useful at the start of a course to inform students plainly, methodically, and accurately what they need to learn.
- That it is what students do rather than what teachers do, that ultimately determines whether changes in their understanding actually take place.

Writing aims and objectives, and thinking systematically about the concepts that students will need to understand and how they will understand them, becomes a perplexing task if instruction is seen from the point of view of teaching theory 1 [transmitting knowledge to students]. Looking at what the student has to do in order to learn is not part of this [theory 1] way of considering teaching.

The second extract (Extract 3.2) Learning outcomes shows that others would prefer not to use the word 'objectives' because it is replete with unwelcome connotations of behavioural; psychology. However, unease with the idea that complex learning can be pinned down in behavioural terms is quite compatible with the view that it is helpful to all concerned to try and say what sorts of learning we hope may emerge from tasks, instruction, conversations, etc.
In adopting the term ‘outcomes’ in place of objectives, Eisner differentiated between the latter, which imply a preformulated specific goal and the former which ‘are essentially what one ends up with, intended or not, after some form of engagement’ (1979, p. 103).

Otter’s definition of learning outcomes as, ‘what a learner knows or can do as a result of learning’ (1992, p. i) affirms Eisner’s perception of outcomes as being what the student achieves as opposed to what the teacher intends to teach. Outcomes, in Eisner’s terms, are broad overarching consequences of learning which do not meet the stringent criteria which necessarily apply to behavioural objectives, where the latter are unambiguous, specific statements of expected behaviour which include the conditions under which the behaviour will occur and the standards of performance which are acceptable ... This is not to suggest that the use of outcomes in planning curricula precludes a statement of intention regarding what one will end up with, but rather that, ‘the precise dimensions of the outcomes cannot be specified to the level of clarity or specificity that instructional objectives ought to have’ (Eisner, in Popham et al., 1969, p. 23).

Subject-specific outcomes relate directly to, and result from, the content that is taught in a given context. An example of such an outcome taken from a second year undergraduate research methods module in education and the social sciences is ‘that on completion of the module the student will be able to apply knowledge of validity, reliability and triangulation to a chosen research issue’. This provides a clear statement of what the student will be able to do as a result of the learning experiences which have been planned ...

Whilst each subject-specific outcome is accompanied with an explicit statement of the assessment task by which it will be assessed and the criteria by which the outcome will be graded, this does not imply that each outcome is to be assessed discretely; the outcomes are ‘bundled’ into a composite assessment task.

This grouping of subject-specific outcomes into ‘bundles’ represents a significant departure from the assessment of objectives as advocated by Mager (1962) who demanded that the standard of performance should be explicitly stated and assessed separately for each objective.

This is explained by reference to the following five outcomes which are ‘bundled’ into a single assessment task in the research methods module referred to above.

Subject-specific outcomes
On completion of the module participants will be able to:

(i) make a clear statement of a research issue and appropriate research questions;
(ii) select and justify a research approach, appropriate for a specific research issue;
(iii) select and justify an appropriate methodology and data collection instruments for a specific research issue;
(iv) identify the major sources of literature relevant to a given research issue;
(v) apply knowledge of validity, reliability and triangulation to a chosen research issue.

Assessment of outcomes (i) to (v) will be by a written assignment. Students will be assessed according to given criteria ...

For Eisner student-specific outcomes constitute what he terms ‘personalised learning’, that is what the student has learned which has not been directly taught, learning that has occurred in an individual independent of direct teacher/student interaction, and learning which may be only tangentially related to the pre-specified subject-specific outcomes. These outcomes are not wholly predictable, and are to an extent individualised, being dependent upon the extent to which the student engages in the learning experience and takes responsibility for his/her learning. As Barnett suggests,
... the tutor hopes to set the student off on the right path, and hopes to see the student being successful. Once the contact has been made, the outcome is to a considerable degree unpredictable; the tutor cannot fully control it, for it is the student that is, or is not, successful.

(1988, p. 248)

[...] Eisner's terminology has links with more recent terms] namely, 'personal transferable outcomes' and 'generic academic outcomes' [...] They may have been taught directly, for example, 'the ability to differentiate between evidence and example', or indirectly through the actions of a lecturer, for example, 'the development of effective oral communication skills', or they may have been learned by the student as a consequence of his/her involvement in a learning experience, for example, 'how to take responsibility for one's own learning'.

Both personal and generic outcomes are perceived as being transferable to a wide range of contexts and in this respect they differ from subject-specific outcomes. The notion of 'transferability' is taken from the work carried out by the National Council for Vocational Qualifications (NCVQ) which has advocated the teaching of transferable skills 'based on the assumption that the acquisition of the core skills in some areas of competence and contexts offers the potential of generalisation or transfer to other areas and contexts which employ the same skill' (Jessup, 1990). Core skills, which generally refer to problem-solving, communication skills, numeracy, personal effectiveness and information technology skills, are now considered to be an integral element of competency-based curricula such as NVQs and the Management Charter Initiative (MCI) and are instrumental in providing the impetus and framework for the inclusion of non-traditional personal outcomes in higher education curricula ...

... [T]he use of outcomes emphasises student achievement and affirms that curriculum planning should begin with what is learnt rather than what is taught. It is acknowledged that there is a dichotomy between learning and teaching intentions, but the process of defining and expressing learning outcomes should enable lecturers to reflect upon what they intend their students to learn and thereby articulate the relationship between what they teach and what students do, in fact, learn.

This presupposes that the learning outcomes are clearly expressed, in a form which enables learners to know at the commencement of a course or module, what it is they are expected to achieve in relation to subject content, personal transferable skills and academic outcomes. But this is not tantamount to pre-specified unambiguous statements of predicted behavioural objectives which derive from a given learning experience. This is unacceptable on three grounds. First there is no intention that outcome statements should seek the unity of response which necessarily characterises behavioural objectives. Secondly there is no assumption that the outcomes derive uniquely from either the teaching objectives or the course/module content. This is not to undermine or denigrate the role of the lecturer, but rather to emphasise the role of the student in accepting responsibility for his/her own learning and to acknowledge that learning might take place in a variety of settings. Thirdly, there is no explicit expectation that the course/module must necessarily be completed in order to achieve the outcomes, some of which may be claimed through Accreditation of Prior Experiential Learning (APEL) schemes.

Equally unacceptable in higher education is the demand for behavioural objectives that indicate explicitly 'standards of the student's expected performance' ...

... [A]ssessment does remain at the core of curricula designed in learning outcome form. The more subject-specific, personal transferable and academic outcomes are clearly expressed, the more the learner is able to concentrate on what he/she needs to know in order to succeed on a given module or course. This places a greater emphasis on the specification of assessment tasks and the criteria by which judgements will be made, thereby forcing both the student and the teacher to examine and articulate the relationship between learning outcomes, assessment and the experience of learning.

Learning outcomes may well subsume a form of learning objective, but the abandonment of the descriptor 'behavioural' is absolutely crucial in allowing outcome-led design to shed the mantle of behaviourism which is antithetical to higher education. A concept of
learning embracing subject-based, personal transferable and generic academic outcomes is consistent with the attributes which characterise a graduate in the 1990s, but there are fundamental conceptual differences between outcome-led design and the traditional university approach which emphasises input and process. The challenge to designers of curricula in higher education is now to harness the use of learning outcomes to view learning from the perspective of the learner, rather than the lecturer, and thereby to enrich the quality of learning experienced by undergraduate students.

The third piece, *The language of learning outcomes* is a contribution to getting greater clarity about outcomes. It is followed by *Deriving learning outcomes from aims* (which returns us to the relationship between aims and outcomes that was introduced in the *Aims and objectives* reading (Extract 3.1)). The author gives concrete advice on writing learning outcomes that set out the teacher’s understanding of what the aims are intended to mean. They are intended to be useful to all.

The language of learning outcomes
By spelling out possible meanings of these terms we are not trying to persuade you to define them in any particular way. We simply want to suggest that it is useful to be more precise about what the terms mean in your particular course and its learning outcomes. Unless such terms are defined, students will tend to guess what you mean by them, sometimes correctly and sometimes not. Guessing what the teacher means is surely not something we should be encouraging our students to do.

**Deriving learning outcomes from aims**
Implications for assessment strategies and methods
Taken together, the materials presented to this point in Chapter 3 set out mainstream thinking about learning outcomes, give advice on writing them and emphasise the point that assessment and learning outcomes should be clearly linked together.

The next reading entitled The SOLO model: addressing fundamental measurement issues (Extract 3.3) extends this useful advice with its introduction of the 'levels' issue. Learning outcomes lead to assessments intended to shed light on student achievements in relation to a set of outcomes. Assessments are 'levelled' in that (a) they are at certificate, diploma, bachelor's, master's level, (b) there are levels of achievement within those levels – for example, in the UK we refer to First, Upper Second ... Pass levels in bachelor's degrees. There comes a point where learning outcomes have to be couched in terms of levels of achievement, which is tricky. I shall return to this in Chapter 4, where I want to challenge assumptions about the feasibility of this fine-grained and reliable differentiation, especially when it comes to the assessment of competence.

The SOLO model: addressing fundamental measurement issues reviews two influential educational ideas about levels and progression. (Progression refers to students working at higher levels of achievement as they progress through a programme.) The authors are critical of the widely used Taxonomy of Educational Objectives, volume I, written by Benjamin Bloom in 1956. I think their criticisms are right although they are not as extensive as my objections to the ways in which Bloom's work gets used. Nevertheless, Bloom's work has influenced many designers who have tried to get knowledge outcomes at the bottom of a hierarchy of levels and evaluation at the top. If I am sceptical, I am in a minority, although I share this minority status with the authors of the SOLO reading. They are impressed by John Biggs' SOLO taxonomy, which is becoming more widely known amongst educational designers. They make a good case for using it to help define levels of achievement and I agree that it can be a useful prompt for thinking about quality of performance, whether in writing, image or action. I am sceptical about their stronger claims for SOLO but recognise that here is a tool that is going to be widely used, alongside Bloom's Taxonomy.
The SOLO model, developed by Biggs and Collis (1982), proposes a structure of learning outcomes, and thus provides a clear basis for a technology of testing within learning and test theory ...

The SOLO taxonomy

Biggs and Collis (1982) developed their model from a study of learning outcomes in various school subjects and found that students learn quite diverse material in stages of ascending structural complexity that display a similar sequence across tasks. This led to the formulation of the SOLO taxonomy (Structure of the Observed Learning Outcome). The taxonomy makes it possible, in the course of learning a subject, to identify in broad terms the stage at which a student is currently operating. In this consistent sequence, or cycle, the following stages occur:

- **Prestructural** – there is a preliminary preparation, but the task itself is not attacked in an appropriate way.
- **Unistructural** – one aspect of a task is picked up or understood serially, and there is no relationship of facts or ideas.
- **Multistructural** – two or more aspects of a task are picked up or understood serially, but are not interrelated.
- **Relational** – several aspects are integrated so that the whole has a coherent structure and meaning.
- **Extended abstract** – that coherent whole is generalised to a higher level of abstraction.

Knowledge: Who painted Guernica?

Comprehension: Describe the subject matter of Guernica.

Application: Relate the theme of Guernica to a current event.

Analysis: What compositional principles did Picasso use in Guernica?

Synthesis: Imagine yourself as one of the figures in Guernica and describe your life history.

Evaluation: What is your opinion of Picasso’s Guernica?

When using Bloom’s taxonomy, the supposition is that the question leads to the particular type of Bloom response. There is no necessary relationship, however, as a student may respond with a very deep response to the supposedly lower order question: ‘Describe the subject matter of Guernica’. Similarly, a student may provide a very surface response to ‘What is your opinion of Picasso’s Guernica?’ When using the SOLO taxonomy, either the questions would be written in a different manner, or the test scorer would concentrate on classifying the responses only. An example of rewriting to maximise the correspondence between the question asked and the answer expected is:

- **Unistructural**: Who painted Guernica?
- **Multistructural**: Outline at least two compositional principles that Picasso used in Guernica.
- **Relational**: Relate the theme of Guernica to a current event.
- **Extended abstract**: What do you consider Picasso was saying via his painting of Guernica?

The greatest criticism of the Bloom taxonomy is that there is little evidence supporting the invariance and hierarchical nature of the six levels. Let us recall that Bloom claimed these six levels ‘represent something of a hierarchical order of the different classes of
objectives. As we have defined them, the objectives in one class are likely to make use of and be built on the behaviours found in the preceding classes in this list. A prior condition of the hierarchy is that there is common understanding of the various levels. Ennis (1985) argued that analysis relates to many levels.

Analysis of a chemical compound, analysis of an argument, analysis of a word, analysis of an opponent's weaknesses in a basketball game, and analysis of the political situation in South Africa seem like such different activities that we might very well wonder just what we are supposed to teach under the label 'analysis'.

(p. 45)

Calder (1983) was much more critical of all Bloom's levels, and provided illustration of the conceptual morass that followed from a 'classification filled with nebulous terms [which] makes it impossible to detect similarities in objectives in different subject areas, and frustrates efforts to develop precise principles of teaching and testing bearing on sharply delineated objectives' (p. 297). As an example, he considered the notion of 'knowledge', which includes cases where the student relates definitions of terms to specific instances as well as recalls definitions verbatim. This first instance could be confused with 'relating abstractions to concrete instances' which is defined by Bloom as 'Comprehension'. Further, he claimed that 'too many categories contain a pantechinicon assortment of skills' (Calder, 1983, p. 298). Such conceptual confusions make it difficult to expect there to be a hierarchy.

Given my critical tone, you might ask why learning outcomes, level descriptors and suchlike generic, content- and context-free statements of achievement are in such wide use. It is fair to say that critics of people such as Bloom and Biggs who have provided 'levelling tools' have not themselves come up with anything better. They sneer but don't create. Those critics have what I think is quite a good response to that charge, namely that there are no good tools to be had because there are no free-floating, content-independent pre-definable hierarchies of mental operations to be measured. Difficulty, challenge and achievement should be judged in context, with the context and the subject matter making powerful contributions to task difficulty. The idea of there being general levels of difficulty is, on their reading, an illusion, and taxonomies such as Biggs' and Bloom's are based on ontological mistakes. However well-based their reasoning may be in the psychology of situated cognition (Anderson et al., 2000), it is diametrically opposed to the spirit of times that want measurements, generalisations and certainties. My position is tainted with post-structuralism and, it has to be said that policy-makers, politicians and many practitioners are committed to managerialist, rationalist and, ultimately, simple views of human affairs. Taxonomies come from the world of determinism and simplicity, which is the world of common sense and modern managerialism. My view comes from complexity thinking and is unfashionable with those who plan, dispose and do.

### Reflection 3.2 Outcomes, aims, planning and environments

I think that it is wise to encourage you to work out a defensible position on learning outcomes. Most of the material in this pack treats them as useful aids to good educational practice. Debate is confined to:

- which outcomes are appropriate to different subjects and areas (see http://www.qaa.ac.uk/crrntwork/benchmark/benchmarking.htm);
- technical issues to do with writing them;
- technical issues to do with summatively assessing them.

I have indicated that I am not so convinced. Indeed, one of my former Lancaster colleagues who I greatly respect 'doesn't believe in them', thinks that 'they get in the way of good learning architectures' and 'direct people's energies to all the wrong places'. I elaborate on some of these issues in Section 4.2.

I recommend you to reflect on the following questions:

- Do you believe in learning outcomes – that is, do you think that we can specify in advance what we expect people to learn from a set of educational engagements?
- If you do, then how precise do you think it is useful to make them? (Bear in mind that fuzzy learning outcomes imply fuzzy assessment techniques and the absence of the sort of 'hard' data that managers like to use as performance indicators.)
- If you are not persuaded by the learning outcomes discourse, what would you use (a) to help you plan modules (b) to guide the assessment of student learning? In other words, what would you use for planning and assessment templates?
- If you are persuaded that learning outcomes have their uses, take four or five that would be appropriate to your module and think about the most appropriate ways of assessing them that you can imagine.
- If you are sceptical about learning outcomes, how are you going to get evidence of student learning and of your achievement as a teacher?
3.3 Programme assessment plans

The material reviewed in this chapter and Chapter 2 has pointed to the conclusion that a good module has a good assessment plan. The plan is evidence of good teaching, just as high scores on student satisfaction surveys may be evidence of teaching skill.

Yet this material has also pushed to the conclusion that a well-designed programme has an assessment plan as well. Module plans should relate to it. Module teachers will usually have been involved in constructing it and will refurbish as the urge or need arises.

A programme assessment plan will at least include:

1. a set of learning outcomes;
2. an indication of the learning outcomes that the department aims to assess reliably so that student achievement can be certified;
3. a map of which outcomes get particular assessment attention in which modules, showing whether this is with formative or summative purposes;
4. an audit of assessment methods in use that shows that the range of methods across the programme is appropriate for the learning outcomes being assessed;
5. grade indicators, at least for those learning outcomes that will be reliably assessed across the programme.

It is worth bearing in mind that:

1. Different learning outcomes get picked up more or less directly.
2. Some are graded while others are not.
3. Outcomes that are not graded still get attention: there is a real sense in which they are assessed.
4. Assessment tasks can readily involve more than one learning outcome. It may even become unusual to set a task with just one learning outcome in mind, although it can happen in modules where there are tests of information retention.

A programme assessment plan is likely to be shaped by the subject benchmarks and other disciplinary content-related or abilities-based reference points. Together these factors inform the choice of learning outcomes, although module teachers may add their own views, which have to be reconciled with the emerging programme specification. The assessment plan below shows the arrangements of one department for the assessment of some practical and key 'skills' across the three-year bachelor’s degree programme. It gives another example of the programme-module relationship by showing how some learning outcomes were associated with key modules so that the department could not only claim that all learning outcomes were coherently covered but also identify where evidence of achievement could be collected for each of them.
A programme assessment plan and some academic and social practices
Just as module teachers inform the programme assessment plan, so it, in turn, influences module assessment practices as is illustrated in the example above.

I think there is a stronger reason for saying that module teachers benefit from having a programme view, namely that it helps them to avoid unrealistic assessment expectations. This is an important theme in Chapter 4, so I will only introduce it here. The basic idea, already introduced, is that teachers often get themselves into a terrible mess with assessment because they (mistakenly) assume that:

1. Assessment means measurement.
2. All learning outcomes can be assessed reliably, if only the right techniques are used.
3. Formative assessment, which does not lead to grades, is inferior to summative assessment, which does.

I challenge these assumptions in Chapter 4. The aim is to free module teachers from the guilt that comes from trying to do the impossible – for example, to find reliable and valid ways of assessing critical thinking. I suggest that instead they might base what they do on a more sophisticated understanding of the assessment of human achievements, which would lead them to:

1. Work with colleagues to devise good, programme-wide arrangements for reliably assessing general learning outcomes that can and ought to be certified – warranted – by the department.
2. Devise good tasks that create plenty of local judgements of achievement.

**An appraisal story**

Ben works in the warehouse of a large retail operation which does a lot of mail order business. The pay is above average for the area, the company keeps investing in modern machinery and all the evidence is that this is a good employer.
There are twice yearly staff appraisals. Ben is rated on fourteen criteria. Ratings, which are done by those who oversee his work, are on a ten point scale. Five is an acceptable score, seven is good. His overall scores have been good and he has had pay rises as a result. His last 'health and safety' score was four, which is low. It led to conversations about what he could do to improve his health and safety rating.

In the same formative spirit he has been invited to an informal review between his last appraisal and the next so that there is time for more developmental work on the other areas that can be further improved. The firm's personnel officer has already said that Ben appears to be right for promotion in a couple of years time (he is 18) and has talked with him about taking further qualifications to prepare for a supervisory/management role. The firm is happy to help by offering a flexible hours package and adding a little more to his hourly pay.

I have tried to illustrate this with An appraisal story (above). At first sight this is nothing to do with higher education, although another look will show that there are parallels between Ben's workplace appraisal and the assessment of workplace competence in many higher education programmes. The issues this extract raises are the same as the ones we face in HE when we try to get reliable, affordable and usable judgements on complex learning outcomes, whether they are practical or intellectual.

Ben's workplace appraisal is summative, in the sense that he gets graded on performance and that his grades are linked to his pay. The company does have a formative intention, seen in the conversations about improving health and safety scores and promotion planning, but Ben sees this as a 'me and them' process in which his aim is to make the company see him in the best possible light. Good though the system is, it has encouraged him to think about looking good. He has no reason to disclose things, even though disclosure and then discussion could help him to do better in the future.

**Reflection 3.3 The meaning of good local assessments**

You might reflect upon lessons for module-level assessment that could be drawn from the story of Ben's workplace assessment. In particular,

(a) is this a good system for judging performance?

(b) how confident could we be that his appraisal scores tell us something?

If you think that the appraisal system is good and are doubtful about the meanings we can attach to the numbers that it generates, then you have identified a serious difficulty with most module-level assessments of achievement; they produce local meanings.

Arguably the problem is more serious at programme level because so many different judgements get mashed together to produce a grade point average or degree classification – I invite you to reflect on the limitations of the ways in which we commonly record and report our judgements of student achievements.
Let me be clear: this system seems to me to be as good as any and a lot better than most. It is also very hard to see how the fundamental imbalance of power (the company judges, Ben is judged) can be eliminated, which limits the likelihood of Ben being too forthcoming. (But if he decides he trusts the people doing the appraisal he may become more open, less strategic.) In many respects, this is quite similar to good assessment practice in degree modules. But what would happen if, as with degree modules, these judgements were to contribute to some general, summative judgements of Ben’s ability? In other words, how generalisable or transferable are these local judgements? Do they mean very much outside the local context?

My argument is that local judgements have local meaning. They ought not to be extended beyond the local setting or summed with other local judgements to provide a seemingly strong measure of achievement. Ben’s appraisal only has local meaning because:

1. Although the judgements are based on clear criteria, we do not know what local meaning is attached to them. We do know from research into the use of assessment criteria in higher education that they get interpreted in quite a wide range of ways (Wolf, 1997; Ecclestone, 2001). Unless something has been done to make sure that all assessors are using the criteria in pretty similar ways, we ought not to assume that the same scores indicate the same achievements.

2. The scores on each scale are put together (aggregated) to get an overall performance grade. This is, of course, adding the number of coconuts to the number of walnuts to get a figure for the total of nuts in a bowl. People often try to correct this error by weighting the different scales, but that introduces questions about what the best weighting is? Should health and safety be heavily emphasised or lightly weighted? Plainly, weighting is not an objective exercise.

3. Ben is known to his appraisers. Are they grading to his performance or to the person? And, as sometimes happens in academic life, are they grading strategically, by marking low (or high) in order to influence future behaviour rather than to reflect present achievements accurately? (In academic life, this might mean marking tough at the beginning of a module in order to panic students to work harder.)

4. How far do his grades reflect his individual achievements rather than the culture of the workplace? Take his health and safety grade. Two things depress it. He has driven his VW Golf GTi too fast in the car park, although now that the suspension has been lowered and the company has put some speed ramps in he drives more sedately. He also clipped some racking with the forklift truck he was driving. There’s something odd here. Why does a man with a 146 b.h.p. car crash a forklift that struggles to 10 m.p.h.? Interestingly, more experienced workers have also knocked into the racks. It turns out that these accidents mainly happen when they use the older forklifts that do not have mountings for the portable computers that the warehouse workers use to tell them how much stock to get from which bays. This means they have to balance the computers while they are driving. If the computer slips, they try to grab it and are likely to lose control of the truck. So, what does Ben’s H&S mark say about him and what does it say about the community of practice in which he works? What he does is what he has learned to do in that workplace, from that community of practice. The point is, of course, that the H&S mark is meaningless outside the local context, which is a remark I would make about many assessment verdicts.
that get presented as if they were reliable and generalisable warrants to competence.

The appraisal judgements are entirely fit for the local purpose of having conversations about performance in a particular activity system. They also seem to motivate Ben. But, despite the appearance of objectivity and reliability, they ought not to be generalised. They do not tell us about Ben’s ability (and purists say that it is never possible to be sure about ability): they tell us how his supervisors judge his performance in a context. They cannot be compared with any other measures of achievement because these judgements are based on a particular grading scale used in local ways with local meanings. For similar reasons, they ought not to be aggregated to get some global performance figure, even if we can trust that the judgements are ‘fair’.

Let me bring this back to my second point, that good module designs provide tasks that lead to plenty of local judgements of achievement. Ben’s case has been used to suggest that many – perhaps most – judgements are best treated as local judgements: they make sense at the module level but there are problems with the idea that judgements can be combined, both within modules and between them. So, if we want secure judgements of achievement that are based on data from different tasks and different modules, we need to work with colleagues to devise good, programme-wide arrangements for reliably assessing those learning outcomes that the department needs to certify. This involves orchestrating module and programme goals, learning outcomes, grade indicators, the meanings constructed around the grade indicators, assessment practices, ways of recording achievement and approaches to combining assessment judgements from different pieces of work and different modules.

In this chapter, I have repeated the suggestion that modules will contain some tasks that will involve summative or formative assessments that have local meanings and others that have been designed to be summatively assessed so as to contribute to programme-level reports of achievement. What is new is my conclusion that module teachers can benefit from appreciating this because it gets them off the hooks of two things:

1. Trying to reliably assess learning outcomes that will defy reliable assessment. In other words, they can settle for creating good, local assessments without worrying about the impossible question, ‘Are my judgements objective and reliable?’ (no, on both counts).

2. Working alone to find reliable ways of assessing learning outcomes that can and ought to be warranted.

This prepares the way for Chapter 4’s more formal treatment of assessment and measurement theory. The aim is to explore what can and cannot be assessed in different ways, which puts teachers and designers of modules and programmes in a better position than their colleagues trapped by the three assessment misconceptions listed above, namely:

- Assessment means measurement.
- All learning outcomes can be assessed reliably, if only the right techniques are used.
- Formative assessment, which does not lead to grades, is inferior to summative assessment, which does.

My view that good teaching involves good design is familiar by now. I want to develop the claim that good teaching is system-sensitive, by which I mean that individuals have a sense of the relationship between what they are doing and the
needs of students, the priorities of the programme and the context of higher education. In this pack I am developing that claim with reference to assessment, arguing that one way in which good teaching is system-sensitive is in the design of module assessment arrangements that mesh with programme assessment plans. Where earlier chapters have started with the module, Chapter 4 looks at these claims mainly from a programme perspective.