chapter 9

EVALUATING: TEACHING AND COURSE EVALUATION

*Happy are they that hear their detractions and can put them to mending
(Benedict in Shakespeare's Much Ado About Nothing, Act 2, Scene 3).*

In this chapter we extend the notion of evaluation beyond collecting
information for the accountability and improvement of courses and
teaching to include the development of reflective practitioners, self-
knowledge and the nature of learning for the future. We introduce a wide
variety of ways of evaluating educational processes, which we relate to the
context of more sophisticated conceptions of quality assessment,
appraisal and accreditation, together with research on learning and
assessment. We suggest that a closer integration of learning with evalua-
tion can make the process less an imposed distraction and more an
opportunity for linking research and teaching and making educating our
students a field for lively, shared inquiry and development.

INTRODUCTION

Evaluation can make all of us feel anxious and defensive, whether it is
through examination, appraisals, reviews, observations, student ratings
or even friendly critics. As academics, we can even feel anxious when we
are just being judged by ourselves. But as Shakespeare suggests, anxiety

can change to pleasure provided we can put our detractions to mending. This suggests that we should link the critical process with a constructive one, just as we would link feedback with student work, as discussed in Chapter 8. To do this, evaluation needs to be well balanced. To achieve balance, evaluation must derive from many complementary sources, since none is adequate in itself.

Good evaluation requires thoughtful rationale and, as such, will be time-consuming. Of course, there are external incentives. For example, the Quality Assurance Agency (QAA) in the UK and the European Association for Quality Assurance in Higher Education (ENQA) both use a careful process of subject review to evaluate programmes by monitoring curriculum and examining core skills. While there is no such parallel national agency in the USA, in many colleges and universities, the evaluation of teaching is linked to reaccreditation by such entities as the Higher Learning Commission (HLC) and to promotion, tenure, salary and retention decisions. The professionalization of teaching in higher education is also beginning to require reflective commentaries and portfolios for achieving accredited teacher status. But are these incentives and the traditional need to improve our courses enough?

In the Introduction to this book, we stressed that there have been two important developments in higher education: the emphasis upon generic skills – especially learning to learn – and the change of focus from teaching to learning. What has been seriously underestimated is the role of evaluation activities in enabling students to understand more about the way they learn from the different styles of teaching and the different environments and learning resources they encounter. This can help students become more effective lifelong learners when the constraints and supports of formal courses are behind them and they take full control of their own learning.

Hurried completions of brief institutional questionnaires may contribute little by themselves, but many of the evaluation methods we discuss in this chapter have much to offer in helping students to understand themselves and their responses to different ways of learning. Reflective exercises in evaluation can become important features of courses rather than conformity to institutional demands but, like the development of study skills and writing skills, they need to be regularly integrated activities.

As in the previous chapters, we consider evaluation to be a particular 'genre' of teaching in higher education with respect to the matrix of learning. Evaluation needs to reflect all aspects of a course and, at the same time, draw on the distinctive contributions of a very wide range of methods and approaches. While some evaluation methods may be more appropriate to

certain contexts or dimensions of this matrix, we highlight a handful which relate to several aspects of the critical matrix and, as such, illustrate how evaluation can achieve balance.

We are not attempting a comprehensive review of all the issues and possible methods of evaluating teaching. In this discussion, we draw primarily on those categories of evaluation which are concerned with the effectiveness of teaching, including doing small-scale studies of the courses or teaching with which the reader may be chiefly concerned. Before we consider the particular methods, however, a brief overview of some of the key aspects of evaluation research might be helpful.

Aspects of evaluation

Figure 9.1 gives an idea of the complexity of the issues and dimensions of evaluation that would need to be considered in any major research study and suggests some which would be important in more local evaluation reports. The arrows are meant to indicate the variety and range of factors that influence and shape the sort of approaches that might be taken towards evaluation research.

One especially important aspect is the educational ideology or values associated with different approaches. The technology of evaluation has often been described as ideology in disguise and it is worth considering whether the recent emphasis on the behavioural and market-driven approaches is at odds with major developments in social research which have tended to moved away from a more positivist paradigm towards a more qualitative anthropological paradigm.

Illuminative evaluation, pioneered by Parlett and Hamilton in the 1970s, is an approach where the investigator studies a programme or course by examining how it operates, focusing on describing and interpreting, rather than measuring and predicting – essentially taking a case-study approach to evaluation (Gray, 2004; Russell et al., 2004). Qualitative inquiry delves into a deeper impression of the experience of an educational programme, allowing us to appreciate the more complex context of education: we learn not just about the achievement of specific objectives, but the unintended outcomes as well.

We have become more aware of the importance of perceptions and of the motivational effects of programmes. Courses that might appear to be very efficient in one way may be very inefficient in another and liable to generate alienation and distortion of deeper, more holistic learning. If, for example, we look at the impact of problem-based learning, purely through
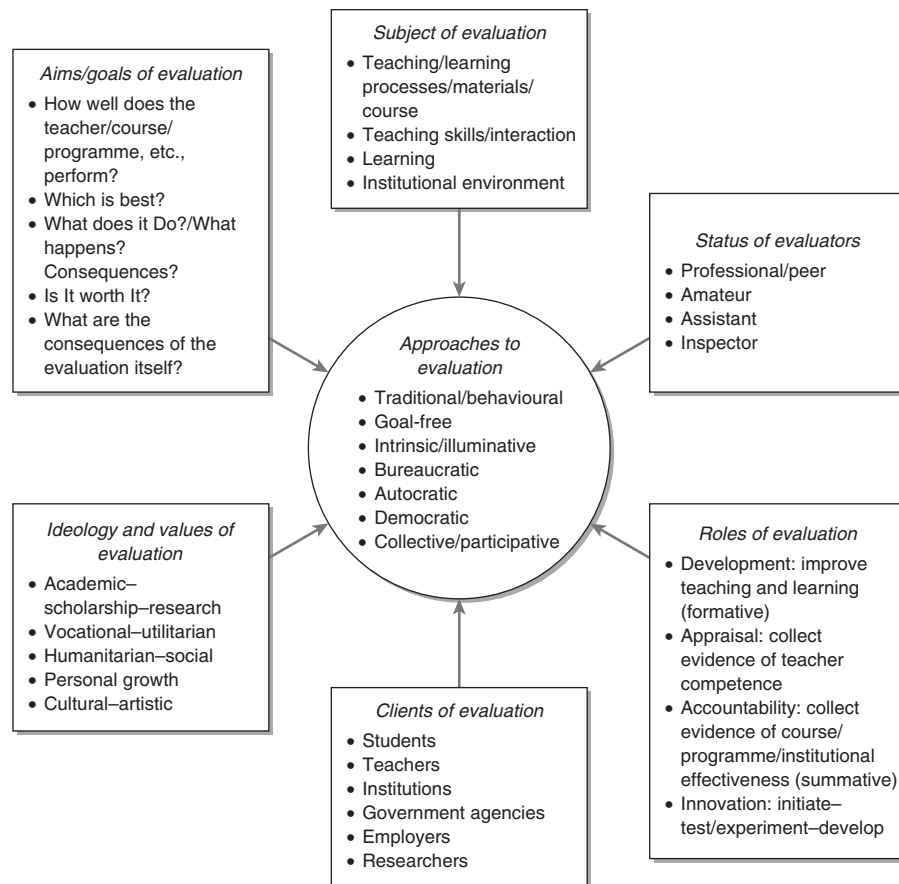
Figure 9.1 *Aspects of evaluation*

the perspective of formal test examination results, we may feel that it is not worth the effort. On the other hand, if we focus on the nature of the discipline and professional values and commitment, we may feel otherwise. This is not to suggest that measurement should not be made in these more diffuse areas but the difficulties involved may dissuade researchers from trying to give an intellectually respectable account of it.

Much educational research now gives qualitative accounts of the experience of learning and teaching, using the actual words of the faculty and students involved. This has led to a deeper appreciation of the different ways in which students learn and value their different experiences. Faculty may be less inclined to generalize about students now, and may perhaps be more aware that average ratings often conceal differences which are important in developing better courses and teaching methods.

The emphasis with illuminative approaches to evaluation is very much on the educational processes – what it is like actually to participate as a teacher or a student. It is not simply employing a different set of methods or methodology. The starting point is different. We now appreciate that evaluators may not take the formalized plans or descriptions of the courses or other instructional systems seriously, since actual practice often diverges significantly from what has been officially proscribed. A central preoccupation is the learning milieu, and how complex social, cultural, institutional and psychological variables interact in multiple ways, creating unique and intricate patterns within different teaching and learning contexts (Parlett and Hamilton, 1977; Gray, 2004).

Since measures of accountability will affect the financing of programmes as well as issues such as hiring, promotion, tenure and salary decisions, it is important that we do not allow simplistic scores and ratings to dominate our understanding of what we are doing. We may need to show evaluations of our teaching – not only with respect to individual performance reviews and, in the UK, institutional audit by the Quality Assurance Agency (QAA), but also for the formal accreditation of our teaching through, for example, formal courses and programmes accredited by the Higher Education Academy (HEA) in the UK, and the increasing number of certificate programmes for graduate students in the USA (Aravamudan et al., in press). Hopefully, such bureaucratic pressure will not drive us to take the line of least resistance, providing the simplest data that we can manage. In this chapter, we explore a range of other methods that might be taken. In emphasizing the need for multiple perspectives of evaluation, we hope to contribute to providing a richer account of the value of what we are doing, not only to those to whom we are directly accountable but also to ourselves.

TEACHING AND COURSES: METHODS OF EVALUATION

To get at these multiple perspectives, instructors can draw on a great variety of evaluation methods. Table 9.1 offers a summary of the strengths and weaknesses associated with some of the most common types of methods. Full descriptions of each follow in the sections below.

Student questionnaires and course ratings

Typically administered at the end of a term, standardized course questionnaires or surveys – also called student questionnaires, student ratings,

Table 9.1   *Relative strengths and weaknesses of common evaluations*

	Questionnaires (General, specific sessions, project)	Interactive teaching (Buzz and small groups, CATs)	Group discussion	Informal feedback	Focus groups	Observation (Peer and structured observation, teaching squares)	Student work
Relative strengths	Broad coverage of opinion	Integration of teaching and evaluation	Flexible, explore issues as they arise	Explore issues at a personal level	Students identify aspects of personal interest	Faculty perspective may complement or clarify students' perspective	Close integration of work with evaluation
	Quantifiable for comparison	Frequent, immediate link with remedial action	Can set tone to encourage criticism	Explore significance of the course as a whole for individual students	Individuals feel they can contribute without being influenced (silent majority has its say)	May be less inhibited in certain respects	Links with areas of high concern
	Easy to administer	Specific learning problems identified, linked to specific teaching	More time and encouragement to think, less 'off-the-cuff' responses	Get to know students better	Letting off steam, without raising emotional temperature	Can be reciprocal learning experience	Enables students to learn self-evaluation
	Close link to particular teaching – immediate	Not explicit teaching evaluation so less inhibition	More responsive to students' perceptions and perspective	Help students to understand their own response to the course	Provides teachers with rich information with minimal time, effort	General opportunity to share ideas about teaching	Enables students to critique their own learning
	Easily confidential or anonymous	Students learn about own learning in concrete situation	Can focus on really important themes	Check on data from other sources	Identifies the group's priorities		Enables students to explore own interests and motivations
	Systematic coverage of themes	Students share problems	Can discuss faculty views more easily		Can identify learning progress towards ability to adapt to and engage in change		
	Longer time perspective		Develop a dialogue				
	Opportunity to explore links between different aspects of course						

Table 9.1   *(Continued)*

	Questionnaires (General, specific sessions, project)	Interactive teaching (Buzz and small groups, CATs)	Group discussion	Informal feedback	Focus groups	Observation (Peer and structured observation, teaching squares)
Relative weaknesses	Frequency use may alienate students May be too simple or too complex Ambiguity difficult to eliminate Encourages 'off-the-cuff' responses rather than serious reflection Can encourage complacency Time limited Even with open questions teacher perspective dominant Needs skill in questionnaire design and analysis	Students may feel over-controlled Exclusively concerned with teacher's learning objectives Too specific, not linked with general review Tendency to avoid teaching process	Dominant personalities can be over-influential Initial tone of discussion can make it difficult to change to a different point of view Coverage of issues may be limited by time available Difficult to assess distribution of opinion Lack of anonymity may be inhibiting Skill in group discussion may be needed	Can be biased by personal impact Coverage of students poor or very time-consuming, biased sample unless careful	Aspects of interest to teachers may be omitted May not provide causative information More time-consuming for students	May be intimidating, especially if status difference May lead to under-emphasis of student view point May be mutual support for undesirable or restricted view of teaching Can encourage too much introspection May emphasize work as learning exercise rather than 'real-life production'

course evaluations and teaching evaluations – are the most traditional and pervasive type of teaching evaluation in higher education. First used in the USA in the late 1920s (Ory, 2000), they have been increasingly used as part of the general move towards accountability in Britain and other countries. These questionnaires are what faculty typically think of as traditional evaluations and may be used for diverse units, including whole courses, individual teaching sessions and specific projects. They may be primarily quantitative or qualitative, or a hybrid of each.

Course evaluations, whether they are simple surveys or more complex questionnaires, often concentrate on the 'intellectual support' aspect of the learning matrix, to the detriment of other aspects of the matrix. Teachers interested in doing more – who wish, for example, for their students to think critically, to relate the course to other problems in the field or students' own experience or to question assumptions or conceptions – may suffer if the questionnaire has no way of evaluating whether these aims are being achieved. To get more comprehensive, richer feedback about the course and their instruction, teachers must carefully consider the questions they ask their students.

General questionnaire

Table 9.2 gives an overview of one of the most common and traditional types of general teaching evaluations used widely in Australian universities: the Course Experience Questionnaire (CEQ). The CEQ has been used for several decades to measure the quality of teaching and serves here as an example of what individual teachers might wish to consider in their own questionnaires (Ramsden, 2003). It currently uses scales that measure good teaching, generic skills, clear goals, appropriate workload and appropriate assessment, although new scales have been developed to measure student support, learning resources, learning communities, graduate qualities and intellectual motivation (Griffin et al., 2003). Their potential applications are included in Table 9.2 and discussed more fully below.

Of the original five main categories, the first three – as in many questionnaires – are mainly concerned with providing support. Scoring well on these three scales suggests faculty are both aware of student concerns and generally supportive of their students. The fourth category, however, begins to address wider issues of understanding – although not exclusively – and the fifth category addresses the issue of

Table 9.2 *Categories and examples of questions in the expanded Course Experience Questionnaire*

Scale	Examples of questions
Good teaching	The [professor] made a real effort to understand the difficulties I might be having with my work
	The [professor] normally gave me helpful feedback on how I was going
Clear goals and standards	It was always easy to know the standard of work expected
	I usually had a clear idea of where I was going and what was expected of me in this course
Appropriate workload	The sheer volume of work to be got through in this course means that you can't comprehend it all thoroughly (negatively scored)
	I was generally given enough time to understand the things I had to learn
Appropriate assessment	To do well in this course all you really needed was a good memory
	The [professor] seemed more interested in testing what I had memorized than in what I had understood (negatively scored)
General skills	The course developed my problem-solving skills (variations include analytic, written communication, team-building skills)
	As a result of my course, I feel confident about tackling unfamiliar problems
Student support	Relevant learning resources were accessible when I needed them
Learning resources	Where it was used, the information technology in teaching and learning was effective
	Course materials were relevant and up to date
Course organization	The course was well organized
	There was sufficient flexibility in my course to suit my needs
Learning community	Students' ideas and suggestions were used during the course
	I learnt to explore ideas confidently with other people
Graduate qualities	I learnt to apply principles from this course to new situations
	The course developed my confidence to investigate new ideas
	I consider what I learnt valuable for my future
Intellectual motivation	The course has stimulated my interest in the field of study I found my studies intellectually stimulating

Sources: Adapted from Griffin et al., 2003; Murdoch University, 2008 (http://www.tlc.murdoch.edu.au/eddev/evaluation/ceq/questions.html)

encouraging independence. As a general questionnaire designed for widespread use, it does not, however, address the range of learning issues particular to individual learning and teaching situations. In this respect, teachers should feel free to extend the design of their questionnaires to an even wider range of learning issues focused on what is appropriate for the students on their particular courses or parts of a course.

In relation to intellectual support, for example, instructors might want to know more about student experiences of the resources available through the course or the intellectual level of the work or, indeed, the extent to which students are able to gain a sense of achievement. In relation to intellectual independence, teachers might wish to augment

CEQ-type questions with those about whether there is an opportunity to identify medium and long-term tasks, or become involved in creative work or design and/or how far students have been able to increase their confidence.

Similarly, when considering the personal dimension of learning, if we are trying to evaluate how supportive our courses are, we might use questions that delve into the students' sense of security, their enjoyment of class work or their interpersonal relationships with faculty. Questions that emphasize independence might also be extended to explore the more personal aspects of choice, including the students' perceptions about their own responsibilities and independence. Students might feel, for example, that their personal identity was undermined by having to conform to more surface ways of learning simply to deal with a heavy course load. Simply coping with course requirements may preclude students from feeling they can work in a way that contributes to their sense of developing personal identity or even occupational identity. Questionnaires may not include such terms, even though they are clearly important for motivation and for enabling students to become independent professionals later in life.

In considering the social dimension, questionnaires might address the development of peer learning communities and the general academic and departmental culture they are learning within. To what extent are there opportunities for supportive peer-working groups and those that run independently of the faculty? Teachers, with some reason, often feel they are not responsible for the students' social life or their accommodation, yet these can be vital to student learning. Many students, moreover, are clearly concerned about the social relevance of their courses but this again is usually not considered.

In relation to the 'interpersonal', peer assessment activities, learning from alternative perspectives and peer teaching sessions which focus on understanding, self-knowledge or communication skills can all be important aspects of student life which evaluation questionnaires might help us to understand. Group projects, peer-managed learning or problem-based learning often focus upon interpersonal skills and processes. These are much more valued than they used to be but they seldom appear in evaluation questionnaires. Such issues might, as we shall see, be better addressed through different forms of evaluation but quite often the 'hard' data of questionnaires carry more weight than that from less formal methods.

Questionnaires for particular sessions or projects

Faculty might opt to administer short questionnaires or surveys during the term, which can help gauge:

- student understanding of difficult sections of the course;
- whether the course is being taught at the right intellectual level;
- whether there is an overload of material;
- whether the lecturer is going too fast or slow; and
- whether more general student perceptions are related to particular sessions rather than averaged over a course.

Questionnaires can be helpful in exploring some of the problems of group work as well, although often the most important source of feedback is time out on discussing some of the problems. Questionnaires about group work can also be particularly helpful as a basis for assisting the group to reflect upon its own processes (Jaques, 2000; Jaques and Salmon, 2007).

With both group work and lecture sessions, the ratings quite reasonably focus upon judgements about the quality of teaching but it can be helpful to ask more descriptive questions about the structures and purposes of the session. If, for example, the teacher is intending to focus a session around a particular problem – clarifying the problem, presenting particular forms of analysis and evaluating different solutions – it is essential to know whether students perceive the session as such. Often it simply appears to them as an ordered presentation of data or information. Similarly, if a lecturer is focusing on a comparison of different interpretations or approaches or theories, it would be important to know whether the session was being interpreted in this way. A mismatch between student perceptions of a session and the teacher's intentions is an essential issue which questionnaires can disclose.

Faculty perceptions of traditional evaluation

Traditional evaluations – particularly the course questionnaire and student ratings – have long caused anxiety, concern and resentment, particularly when faculty believe they are used unfairly or are too heavily weighted in issues concerning promotion, retention, tenure and salary decisions (Johnson, 2000; Ory, 2000). All too frequently, questionnaires are used for administrators (who wish to maintain accountability) or for students (who wish to communicate their opinions of an instructor or the instruction to

their peers), and not for faculty. Not surprisingly, many faculty are suspicious of questionnaires and have sought to attack their validity and reliability (Calkins and Micari, 2008).

On the one hand, research on student ratings has disclosed – and mostly dispelled – at least sixteen myths related to student evaluation. These include the beliefs that:

- students cannot make consistent judgements about the instructor or course because they lack the necessary experience or maturity;
- only faculty colleagues with strong publication records can evaluate their peers' teaching;
- students ratings are just a popularity contest, favouring the friendliest professors;
- questionnaires generally lack reliability and validity;
- student ratings correspond with their real or expected marks in the course; and, finally,
- student ratings are primarily summative and not useful for improving instruction (Aleamoni, 1999).

On the other hand, there are factors that have been shown to influence traditional evaluations, including:

- the instructor's age and student perception of teaching experience and expertise in the subject area;
- specific instructor personality traits;
- student achievement level; and
- students' approaches to learning (Entwistle and Tait, 1990; Shevlin et al., 2000; Bosshardt and Watts, 2001; Wolfer and Johnson, 2003; Sprague and Massoni, 2005).

Indeed, research has found that some of the biggest disparities in ratings are found between different disciplines. Ramsden (2003) reports research from the CEQ which indicates that the visual and performing arts are the most highly rated subjects, with the health sciences and engineering rated the worst. Social sciences come in the middle, rated higher than the natural sciences but below the humanities.

Gender, too, may have an impact on student ratings. Early on, Martin (1984) found that sexist stereotypes negatively impacted on evaluations of the effectiveness of female teachers, while Baker and Capp (1997) have suggested that an instructor's pregnancy and gender may have had an adverse

impact on student ratings even in a feminist studies class. More recently, Miller and Chamberlain (2000) contend that students' misattributions of their female professors educational attainment and university rank may subtly bias course evaluations.

Despite their potential limitations, however, faculty certainly can use course questionnaires in a formative way, in order to improve their teaching. To provide meaningful feedback, questionnaires must both be well designed and carefully interpreted.

Designing course questionnaires

The design of questionnaires is, as Oppenheim has suggested, an intellectual exercise in which we are constantly trying to understand our goals (2000). Unfortunately, questionnaires and scales are not always constructed in accordance with normal psychometric processes (e.g. establishing validity and reliability). All too often, questionnaires are simply a collection of isolated *ad hoc* items that can be reported in terms of particular questions or as overall scores, both of which are dubious if we are trying to diagnose problems and 'put them to mending'.

Questionnaires can be constructed in terms of scales or major themes, such as the themes suggested in Table 9.2, and tested to see whether these themes are a reality. In general, questionnaires vary a great deal in how much background information is collected. It may be important to know, for example, whether high ratings or a particular category of responses are restricted to certain types of student: for instance, women, non-traditional students or those in danger of failing the course.

To help diagnose problems, a questionnaire might offer a range of course objectives and ask students to comment on whether those objectives were reflected in the course. This works best when they do not assume the teachers want all the objectives rated highly. Matching teacher ratings and student ratings on these can be particularly interesting. Students and teachers often have different perceptions, and where discrepancies are seen it is useful to clarify them early in the course. Other questionnaires have asked students to compare one course or one lecturer with another and comparisons are often more useful than attempts at absolute measures. Other useful variants can include asking students how important they think certain characteristics of the course are, again perhaps comparing them with the teacher's views. Assuming that all the characteristics are equally important can often make interpretation very suspect.

Table 9.3 *Contrasting responses to a supervision course*

Question	First student ('Sunil')	Second student ('Pilar')
What were the most successful parts of the course?	'The two rather formal lectures'	'I enjoyed all of the course. I found the whole course helpful'
What were the least successful parts of the course?	'Working in small groups and reflective triads'	'None, it ran very smoothly, the frequent changes in teaching, learning style maintained interest.I have learnt a lot'
What was the most surprising element of the course?	'How much time was wasted,just as much if not more can be got by reading the book'	'That supervision at so many different levels can have so many similarities'
What changes would you suggest?	'Replace the "games" with solid,sound lectures and whole-group discussion'	'None! I expected more guidelines in the beginning but accepted my role in deciding my responsibilities and negotiating other people's'
What is your overall impression of the course?	'Not a fruitful use of [course time], uninspired and uninspiring'	'Excellent. I will recommend it to colleagues. I now feel much more confident about supervision and supporting colleagues'

Interpreting written remarks

One of the most challenging and frustrating things for teachers is to inter-
pret student written comments, especially when students seem to have had
widely disparate experiences of the course. Without a thematic framework,
student remarks can seem random, unconnected, contradictory and lack-
ing discrimination, leaving faculty feeling frustrated, annoyed or simply dis-
missive of the students' ability to judge the course or their instruction
(Lewis, 2001). Table 9.3 offers a typical set of such disparate remarks, in
this case found in a course on supervision, which reflect the extreme con-
trasts many teachers often experience.

Although the course referred to in Table 9.3 had been advertised as
interactive and hands-on, the first student ('Sunil') came to the class with
quite fixed ideas about what courses should be about, expecting the teach-
ing to mainly be about the transmission of information. He viewed acquir-
ing a quantity of information in a short period of time to be a key criterion
for success in the class. The second student ('Pilar'), on the other hand,
came with an open mind and was keen to make the class as useful and

Table 9.4 *Study orientations and evaluation*

Reproduction orientation	Meaning orientation
1. *Focus on efficiency*	3. *Focus on communication*
Basic lecturing skills	Quality of explanations provided
Provision of clear goals and standards	Use of real-life illustrations
Systematic organization of course	Use of humour and enthusiasm
Workload and level of difficulty	Empathy for students
2. *Focus on organization*	4. *Focus on independent learning*
Interesting and relevant content	Assignments providing choice, resources
Level at which material is pitched	Full explanations in feedback
Pace at which topics are covered	Assessment related to course aims
Clear structure within lectures	Advice on study skills and strategies

Note:
Categories 1 and 2 are endorsed more strongly by students with reproducing orientations.
Categories 3 and 4 appeal more strongly, on the whole, to those with meaning orientations.
Source: Adapted from Entwistle and Tait, 1990

enjoyable a learning experience as possible. Individual discussion can assist both in identifying such mismatches between expectation and what is happening on a course, as well as helping students to adjust their expectations appropriately. Alternatively, it may enable students to choose courses which more closely match their expectations.

Approaching a set of student written comments as a process of qualitative inquiry can help make the process of interpretation not only more palatable, but also useful for assessing the strengths and weaknesses of the course and instruction, and for learning from one's instruction. Lewis (2001), for example, suggests that instructors categorize the responses either by simply comparing the responses directly with course ratings (e.g. 'What were the specific comments of students who numerically rated the class above average?') or by using a more complex matrix which compares individual course ratings with specific teaching areas or learning objectives ('What did students who rated the class highly say about their critical engagement with the course material?').

Considering the approaches to study that students take may also help understand the approaches students take to evaluation. Table 9.4 shows one way that remarks can be categorized.

Students taking different approaches to their learning in a course will tend to focus on different sets of evaluations items (Entwistle and Tait, 1990; Entwistle et al., 2000). Students taking more surface types of approaches to study, with reproducing orientations towards their learning, will tend to focus on items assessing the efficiency and organization of the teacher's

teaching. These are areas which could help or hinder the student's ability to memorize and reproduce the content which the teacher is teaching.

Students taking deeper approaches to their study, with an interest in constructing their own meanings from what the teacher is teaching, while still interested in the issues of categories 1 and 2, will, nevertheless, tend to focus more on the teacher's ability to communicate in a way which stimulates and offers the potential to make meaning from the content being taught (category 3), and on the teacher's provision of opportunities for the student to develop the capacity to learn independently on the course (category 4).

Interactive teaching

We all learn a great deal from our experience of teaching, but some learn more than others. Similarly, some use ways of teaching that offer more in terms of understanding student responses than others. Even with traditional, non-interactive 'transmission' lectures, we can observe certain features in our students that tell us whether or not learning is occurring. We may see signs of attention and non-attention, for example, although the actual quality of learning may be difficult to discern from the expressions of faces or the activities of pens, or the use of laptop computers. Reflecting on our lecturing, even keeping reflective diaries – an activity encouraged by many accreditation programmes – can make us more aware of our own activities and be useful for many lecturers, but it may not tell us very much about learning.

Background knowledge check

One of the best ways for learning about learning is through interactive lecturing (see Chapter 4). As a method of teaching, it improves learning by enabling students to consolidate their thinking and relate it to their own experience and their knowledge of the field (Bain, 2004). It can also function as a good method of evaluation, helping lecturers see whether the way they are teaching is appropriate to the level and interest of the students.

While focusing upon the learning of specific content of the session, it can also enable a lecturer to understand where the students are coming from and the sorts of expectations, assumptions and even hang-ups they might have about the topic of the session (Lieberg, 2008). Beginning a lecture with a quick exploration of what the students already know will not only convey the sense that the lecturer cares about and values their experience, but also serves

as a valuable method of 'evaluation-in-action', indicating whether the lecture is appropriate for their current development. We all make assumptions about what our students already know, but rarely do we test out those assumptions.

Buzz groups

Sometimes, rather than posing open questions to the group, a quick brainstorming session may be appropriate, eliciting and outlining some of the students' major concerns and preoccupations regarding the course. It can contribute to creating a shared agenda that can be an excellent basis for understanding whether our teaching is producing effective learning. During the session short buzz groups, with or without setting specific tasks or problems, can also provide a concrete basis for evaluating how much and what sort of learning is occurring (McKeachie, 2006).

Open tasks, such as asking students to 'discuss what we have talked about so far and raise any problems or other issues you would like to share', are also useful at getting at more unexpected problems. And specific tasks calling for the application or interpretation of some of the content can be useful for checking on what sort of learning is occurring and how far students are changing. This kind of evaluation is excellent for adding to our understanding without contributing to our workload. Teachers and students may, however, feel they are covering less content if used too frequently. On the other hand, evaluation feedback from this source may indicate that the course is going too slowly, or that the content is not being pitched appropriately.

Clarifying conceptual knowledge

This type of evaluation is mainly concerned with issues of intellectual support, helping to clarify problems and explore misconceptions. Engaging students with their responses can, for example, lead to a clarification of what is expected and the criteria for making progress. A study of engineering students (Cox, 1987) found that a major problem of comprehension was not the difficulty of the material but the speed of lectures. They did not allow students to consolidate their learning. The opportunity to reflect when you feel you understand is not widely appreciated in the rush to cover the ground. Rushing on can create a sense of 'retroactive inhibition' blurring earlier learning by passing on before the 'ink has dried', so to speak.

Asking the group as a whole to explain or discuss a diagram or graph or perhaps even a quotation, again, might enable the lecturer to

understand where the students are coming from and some of the conceptual problems they might be having. It provides the lecturer time out from their monologue as well, valuable time to evaluate-in-action. Such strategies can also emphasize intellectual independence, but we might also ask for more creative responses and applications. Students often talk of taking down notes which they do not understand and which often undermine their confidence. Providing students with the time and opportunity to explore and express their own ideas and ways of understanding can help them to become more independent learners (McKeachie, 2006). It also enables teachers to understand how far they are assisting their students to become more independent as opposed to conformist learners.

Other CATs

Classroom assessment techniques, discussed in reference to the assessment of student learning described in Chapter 8, can also be useful for evaluation. Often quick and easy to create and implement, CATs can help evaluate aspects of particular sessions (Angelo and Cross, 1993). The 'minute paper' is a good example of a simple but effective evaluation (McKeachie, 2006; Lieberg, 2008). Here, students are asked to answer just two questions at the end of the session:

- 'What is the most important thing you learnt during the session?'
- 'What is uppermost now in your mind at the end of the session?'

Not only can teachers (and possibly the students) get a grasp on the students' knowledge, comprehension or ability, but they can also reflect on their teaching or instruction, to gauge whether they may need to rethink their explanations, assignments or activities if their students do not seem to be learning as they would like.

Generally, then, within the *personal dimension*, if the questions and the problems are not too difficult or posed too aggressively, such exploration can enable students to feel a closer personal relationship with the staff. It helps develop a sense of security and even enjoyment. Interactive teaching generally encourages students to behave and feel more as engaged people and not simply passive recipients. In addition their sense of 'personal independence' may be strengthened by a sense of participation in exploring their own relationship to the subject. Interactive teaching can contribute to deeper learning by encouraging students to relate a topic to their own experience and interests.

Interactive teaching can also address the social dimension of learning, helping to generate an atmosphere in which staff and students work together. Teaching is not that of a remote teacher doing their own thing with a group of unrelated others, but rather a mutual exploration which facilitates participation. It can assist in reducing the impersonality of higher education which many students experience. Where periods of group work are interspersed with more formal teaching, this can also help to develop 'interpersonal' skills. Students often learn from each other things that it is difficult for the lecturer to help them learn. As one student said: 'when another student speaks you prick up your ears, it's something different.'

Group discussion

Very often questionnaires raise more problems than they solve, especially in relation to divisions between those who rate the course highly and those who rate it poorly. Background questions can sometimes help solve this problem, but may become too cumbersome or numerous to manage efficiently. In these areas, group discussions can be an effective form of evaluation. They not only help focus on the background data, but also analyse certain likes and dislikes. They can be very formal with pre-specified agendas and topics or can arise more informally.

Less formal open questions, for example, can explore more fully the issues behind good and poor ratings and the disparity between students. The responses are frequently very enlightening but, unfortunately, they are also often ignored. Although spontaneous meetings are helpful, it is useful to prepare students with an idea of what you would like to hear from them. Questionnaire responses provide a basis for discussion but may also be too restrictive. Discussion that is cursory or unreflective may repeat some of the 'off-the-cuff' likes and dislikes often obtained from questionnaires.

Table 9.1 lists several advantages of group discussion, particularly where flexibility and the development of a genuine dialogue are essential. In such situations, teachers can encourage deeper criticism and work through the nature and implications of these criticisms. It also provides an opportunity to focus on the essential themes and to be more responsive to student perceptions and perspectives. A few dominant personalities, however, can unduly influence discussion, making it difficult to hear diverse points of view and discern more subtle differences. For some, the lack of anonymity is also inhibiting, often making it difficult to assess the distribution of opinions.

Reflective triads

In this activity, groups of three students reflect together on their learning at the end of a session, which offers a useful way to stimulate evaluation through discussion. They not only remind students what has happened but they also encourage deeper thinking about the nature of the session and their own learning. In large classes, reflective triads can help students think through their ideas and what they wish to say without too much interference (unless, of course, they happen to be with the one or two dominant personalities in the group). It is often helpful to begin the discussion with a review of the course (McKeachie, 2006). It is important that the discussion is not seen as just a collection of views and opinions but, rather, puts those opinions to test. Ideally, the reflection session should provide time to work through criticisms and help students to understand intellectually what the course is doing and what it is not doing.

The timing of evaluative discussions is also very important *vis-à-vis* student learning. Discussing aims and objectives after the experience of a course is very different from discussing them at the beginning or during the course. Provided some discussions are held halfway through the course or at other times before the end, discussion can be very supportive for students who may have concerns, anxieties or misconceptions about what the course is doing.

`Independence can be encouraged by not treating students as passive consumers for whom evaluation is conducted simply to improve the quality of the product. Students need to believe they are active participants in the evaluation and its role in course development. As such they will be expected to respond intelligently to the positive features of the course as well as its faults and problems. As in the example of the general medical practitioners (see Chapter 3), responsive discussion after a poor start can be extremely useful in encouraging a strong sense of commitment to making the course develop in a way that both the students and teachers feel is useful. More formal courses may not have the opportunity substantially to restructure but there is always some flexibility. Students should not feel the only point of them being engaged in evaluation is to improve future courses.

In the personal dimension, discussion can be very supportive in providing students with a sense that faculty value their views and priorities and encourage them to express these. Appreciating the alternative perspectives within the group will also contribute to developing student independence. Without the opportunity to hear about these, students often feel their own opinions are

what everyone else thinks. Discussion should be an opportunity for them to learn more about their own responses to the constraints and opportunities of the course. Finally, as in any group process, if the discussion goes well, it can be experienced as an enjoyable social event contributing to the interpersonal context of learning. It can help generate a genuine dialogue among students and teachers in which each learns to share and appreciate other points of view and learning.

Student-generated statements combined with group discussion

This method usefully combines some of the characteristics of questionnaires with that of group discussions. In this case, the questionnaire elements are statements generated by students that may concern various elements of the design, content, methods and environment. Individual students are asked to write four to six statements and three recommendations related to some or all these areas on separate slips of paper. Pairs of students then look at each other's statements, discuss them and select the four or six most interesting statements and three most appropriate recommendations. These pairs then join another pair and do the same selection in a group of four.

Teachers can also introduce and test out their own concerns in statements that are then rated by the class in conjunction with the student statements. Asking the group to rate the statements and recommendations as to how far they agree with them can also produce useful numerical data. If there is time, rated collections can be circulated to the total group for more general discussion. Especially important themes or issues for the course can also be taken up in subsequent sessions, pursued in small groups and/or integrated with student presentations.

Students are not only introduced to the variety of each other's alternative perceptions but they also have to engage with this in making decisions as to which they should select for the next move. In our experience they frequently enjoy the process and find it provides the basis for moving beyond 'off-the-cuff' comments and perfunctory general discussion. It enables them to think more deeply than they do when filling in questionnaires and is often very effective in generating engaged discussion about learning across the intellectual, personal, social and practical dimensions in a very interactive way. It also produces a wide range of interesting issues that may not have been reflected in teacher-designed questionnaires. Time can be a consideration here but, as with some of the other methods

mentioned in this chapter, it can also be an effective learning and teaching activity (in contrast to a bureaucratic task) and is worth spending the occasional half an hour pursuing. The details of this process can be varied to suit the particular situation, discipline or institution as is appropriate.

Student work

Ironically, evaluations of teaching and the teacher drawing upon student work (and its assessment) offer valuable opportunities for teacher evaluation. Assessment provides a mutual object for engaging students about problems or issues in their learning and from which teachers may gain valuable insights about their course. As Ory (2000) reminds us, the original Latin root of 'assessment', *assidere*, literally means 'to sit beside'. Thus:

> *Assessment as 'sitting beside' promotes a developmental perspective. It is not a single snapshot but rather a continuous view. It facilitates development rather than classifying and ranking the faculty by some predetermined measurement such as a student rating item or number of publications. It encourages breaking away from the winner-loser mind-set, comparing one person to another. Instead the focus is on understanding the colleague's perspective and achievements, which means the focus is on real-world performance.*

Certainly, examination results do not typically provide good opportunities for extended insights, but will often indicate gaps in course content or difficulties with understanding specific concepts. Reports, essays and projects will generally provide more in-depth opportunities: the feedback given to students – whether written or oral – should alert teachers to student problems which may need to be dealt with more comprehensively. This feedback also frequently provides the basis for discussion through which teachers might understand their students' problems on the course more fully.

This method allows closer integration of student work with evaluation and can provide an opportunity to explore areas of high concern. There is often a tendency, however, for students to want to be told what to do to improve their assignments. If teachers do this in a rather prescriptive way, then, as Hounsell et al. (2007) suggest, the student may not understand – as they bring different assumptions about learning – and it may make very little difference. If, however, the teacher explores alternatives with the students, getting them to suggest possibilities, then they may enable students to take more responsibility for the improvement.

It is a process that can enable teachers to understand better student problems and why they have them. Simply telling them what to do to improve may add little to the teacher's understanding of student learning. Although such discussions usually fall within the intellectual dimension, they are also an opportunity for addressing personal issues: for developing a better understanding of why students fail and how personal motivation and/or social issues might be more responsible than purely intellectual failings. Several processes of assessment that may help the evaluation of teaching emphasize the formative over the summative.

Portfolios and reflective commentaries

There are a number of less traditional forms of student assessment that deserve a separate mention with respect to evaluation. They display many of the benefits described above but, in addition, provide scope for evaluation over a longer period of time, comparing, for example, recent written feedback with earlier feedback. They also enable the development of a much broader perspective of student learning on a course (McKeachie, 2006).

One of the most useful assessment methods which teachers might use for evaluation purposes is the reflective commentary, which frequently draws upon material collected in private reflective diaries as well as from the course itself. Students are increasingly submitting such commentaries either as assessment or as part of assessment, often as part of extended portfolios (see Chapter 10). They usually comment on a wider set of issues of student experience and learning than simply course content (White, 2004). At different times they may focus upon all the aspects of the matrix but perhaps are most useful in enabling students to develop towards more independent reflective practitioners. They involve many personal choices and are designed to encourage independent reflection on responses. Their very breadth and depth mean they can be extremely valuable documents for teachers wishing to learn more about the impact of their courses and teaching on their students.

Journals and session reports

The use of journals or diaries and session reports can be related to and even provide the basis for much of what is written in more public reflective commentaries and portfolios (McKeachie, 2006). They are, however, not methods of assessment so much as group and class-learning activities providing extensive scope for evaluation. Students might, for example,

be encouraged to keep reflective journals over the period of a course, describing their learning experiences – including concerns, delights, responses to particular sessions – across all four intellectual, personal, social and practical dimensions.

Despite good intentions, such diaries are rarely put into practice. Linking them with specific activities and responsibilities on the course helps maintain a commitment to them. Providing time for diaries in class and even – as was done very effectively on a course for general medical practitioners – asking for individual students in pairs to report back to the class in turn on previous sessions has wide-ranging benefits. It extends and consolidates learning, generates interest among the students in the different ways in which different students experience and learn on the course and provides the teacher with substantial evaluative data.

The rich information that journals contain and the reports generated from them will frequently become part of further diary sessions. Time permitting, these reports and diary reflections can be developed as part of group projects on which groups can be asked to report back both halfway through a course and at the end of it. Reports may be written or oral or both, and can be integrated with more elaborate student presentations. It is worth emphasizing, however, that the quality of the material generated – both for individual learning purposes and for course evaluation purposes – is dependent on students being given the choice to retain those reflections they wish to remain as private and/or to report material anonymously to the course.

One-on-one discussion

As with group evaluation through discussion, individual discussion with students may provide important feedback. Such approaches may range between formal approaches, such as interviews, and informal discussion as the opportunity arises. Such discussion helps to explore both particular issues and the significance of the course as a whole at a more individual and personal level (McKeachie, 2006). It can also be a good chance to get to know students rather better as some can be very inhibited within groups. On the other hand, personal impressions might bias evaluation if generalized to the course as a whole.

Teachers cannot expect to have long discussions about the course with all students, so the sample and the lessons learnt need to be carefully judged. Individual discussions are also useful for exploring and understanding differences in individual perceptions and responses to the course. Substantial differences in preconceptions of what a course consists of, or

should consist of, can make the role of evaluation extremely problematic if not more clearly understood by the teacher.

Student focus groups

There may be times, however, when teachers would rather have an outside facilitator discuss the strengths and weaknesses of their teaching and the course. Student focus groups, also referred to as small-group instructional diagnosis (SGID) or small-group analysis (SGA), are a form of nominal group process which offers a means for collecting early and substantial feedback directly from students using a whole-class interviewing technique (Diamond, 2002).

Simply, the professor leaves the room and a trained facilitator breaks the class into small groups, and asks them to respond to a series of questions about aspects of the course that enhance or challenge their learning. The facilitator asks the groups to share their responses with the rest of the class, and then asks the class to determine as a whole which items are most important. The facilitator may also ask students to consider their role in the learning process, by asking them to identify what they can do as students to enhance their learning in the class.

While there are many variations to this process, the facilitator may also ask the students to rate the importance of the top-three items individually, to help determine if the comments or concerns are shared broadly by most of the students in the course, or if just the ideas of the most outspoken students in the class. During the session, the facilitator can address apparent contradictions in student opinions, peeling the layers off a problem to get at the source of discrepancy. Later, in a confidential consultation, the facilitator will report the students' feedback, categorizing the comments thematically and discussing what, if anything, the teacher needs to do in the course to facilitate student learning (Lewis, 2002) (see Box 9.1).

Box 9.1 *Case study: small-group analysis*

Halfway through the term, 'Professor Yu' requested an SGA be conducted in his class on 'Political decision-making' – an upper-level seminar in political science. After the outside facilitator divided the class into five groups of four, the students identified three key aspects of the course that they believed enhanced their learning: 1) the experience Professor Yu brings to the classroom (relevant stories, anecdotes, etc.); 2) his easy conversational, informal classroom style; and 3) the overall organization and structure of the course (particularly, the sequencing and progression of topics and readings). When surveyed individually, nearly all the students agreed with the class consensus on these points.

(Continued)

(Continued)

The students were then asked to explain how the course could be improved to enhance their learning. While individual responses varied somewhat, as a group the class generally agreed that it would help 1) if the individual sessions had more structure and focus (he apparently frequently meandered off-topic); 2) if the readings could be less complex and lengthy; and 3) if he could intersperse his lecture with more small-group discussion activities.

Finally, when asked what the students believed they could do themselves to enhance their learning in the class, they agreed they could 1) read the assigned texts more thoroughly; 2) bring extra materials from outside the class into discussions; and 3) come to class prepared with questions.

In the follow-up consultation, the facilitator clarified for Professor Yu the students' two uses of the term 'structure' which, without explanation, would seem to contradict each other ('class structure helpful', 'class lacks structure'). The facilitator suggested ways to make his lectures more interactive by interspersing some common classroom assessment techniques to engage students more meaningfully. She also suggested that Professor Yu consider a 'lecture script' to keep him more organized and on track during individual class sessions, or at least to provide clear segues between topics. Lastly, she suggested he reflect on the readings, not necessarily to reduce content but perhaps their density. At the next class session, Professor Yu spoke to the class and clarified a few additional points with them directly. While he did not change the readings right away, over the second half of the term he worked several minute papers, 'think-pair-shares' and more small-group work into his lectures.

Although the process requires some planning and explanation, the student feedback provides teachers with a wealth of information without taking up too much teacher time. The process helps identify the group's priorities and with appropriate follow-up can be a useful stimulus to change and provide students with the opportunity to participate in change. In addition, it can provide an opportunity for both personal expression and group interaction, enhancing a sense of social participation and enjoyment. It may even provide an opportunity for students to let off steam. On the downside, the teacher's interest in the learning and teaching encounter may not be well represented and the activity may not provide any in-depth account of why these opinions and priorities exist.

Teaching observation

Observers drawn from peers and colleagues provide a very valuable alternative perspective to those obtained from students. They will likely have the benefits of:

- having relevant subject expertise;
- having personal experience of teaching;
- experiencing the course over a longer time perspective;
- having knowledge of related courses; and
- understanding the constraints under which the course is operating.

While aspects of observation may elicit additional anxieties and concerns about one's teaching abilities, in some respects the process can also be less inhibited, particularly if the arrangement is handled sensitively and is reciprocal, each person learning from being a critic as well as receiving criticism.

There are, of course, other problems that may need to be addressed. The process might be especially intimidating if there is a power and/or status difference between observer and observed. Observers with their own 'agenda' and/or a different conception of teaching from that of the observed could seriously undermine the observed teacher's development. They might also de-emphasize the student perspective, leading to mutual support for undesirable or restrictive views of teaching. Ideally, issues such as these should be discussed prior to the observation. (See Appendix for an example of a detailed observation protocol and guidelines.)

The activity of observation, including prior and post-observation discussion and/or reports, should be supportive and challenging. It should – especially if participants are able to share some of their fears and inhibitions – also encourage independence in the teacher's learning about their teaching. 'Co-counselling' is another valuable way of enabling teachers to work through some of their teaching concerns and problems, although it entails the development of skills that have wider application than evaluation.

Under the impact of programme and quality reviews, observers are increasingly coming in to assess directly the quality of individual departments. Observations are becoming a much more common feature of these reviews, but the 'accountability' purpose behind them is likely to restrict significant mutual learning. Peer observation has the added benefit of preparing colleagues for these visits. Observations may either be very structured or more informal, depending on the purpose and nature of the observation.

Structured observation

In this type of observation, the observer – whether a peer, a new colleague, graduate student or trained outside facilitator – focuses on specific teaching areas during the session. Taking an anthropological approach to observation, the observer carefully logs the events and activities that occur during the session, keeping track of how long each aspect of the session takes. Leaving interpretation until *after* the session is complete, the observer simply records what she sees, considering key teaching areas:

- sharing learning objectives/rationale for activities;
- promoting critical thinking, engaging students, effectiveness of teaching methods, organization and clarity of session;
- familiarity with material; and
- effective use of course materials and technology.

After the session, the observer will write up her notes more concisely and offer suggestions or comments thematically (see the Appendix for an example).

Combined observation and focus group

Here, the observer will first conduct a structured observation and, when the teacher has left the room, will then talk to the students directly about his observations in a focus-group format. The observation thus serves to enhance and clarify the student remarks, and helps explain ambiguities, discrepancies and points of confusion, as well as adds a richness to the overall observation.

Teaching squares

Other teachers might opt for a less formal observation method. Here, four teachers take turns observing and reflecting on one another's teaching sessions. The process is meant to be mutually beneficial, self-referencing ('How can what I observed improve my teaching?' not 'Here's how you can improve your teaching'), and appreciative and respectful of one another's teaching methods and style (Hafer et al., 2002).

Digital recording, playback and discussion

Video-recording (now digital recording) has long been a very important part of educational development workshops designed to improve skills in lecturing and small-group work. Telling teachers what is wrong with their teaching is not, as we have seen, always the best way of achieving significant change. It is particularly true where there is a strong emotional element in what we are doing, and certainly teaching styles are quite closely bound up with a sense of personal identity.

This is not to say that comments from others cannot be extremely helpful but they are likely to raise far fewer defensive responses when the teacher can plainly see what the observer is talking about. Even viewing a recording by yourself can enable you to take a more objective stance and a greater sense of responsibility for what is happening and how you might

improve. The tape or DVD can be stopped and time given to thinking why mistakes were being made and how they could be avoided in future. Simple reflection after a teaching event can, of course, be very helpful but it is easy to forget the more worrying parts.

There is more likelihood for change if this activity is seen in a positive light. It is useful to have agreement on the kinds of areas of teaching performance and their relationship to learning. It is also usually better to begin reviewing a performance from the point of view of what was successful. If the recording is watched in the company of students, peers or educational developers a positive attitude may, indeed, be reinforced by their comments if the positive side is explored before suggestions are made for what can be changed.

In approaching the areas needing improvement, it is often better to begin with critical comments from the observed teacher rather than from the observers. We are far more likely actually to change our behaviour or attitudes in response to criticisms we have made of ourselves than we are from those of others. Nevertheless, we all have our blind spots and, given the right atmosphere, critical comments can be taken on board and acted upon, especially within the context of a revealing tape and discussion within a supportive atmosphere.

There is considerable scope within such reviews, particularly if students are involved, to explore the performance with respect to the various dimensions and contexts of the learning matrix which the recorded session is addressing:

- What aspects of learning are being addressed?
- What are the teacher's intentions in this respect?
- Are they appropriate?
- Are they shared with the students?

It is worth mentioning that the critical matrix (Chapter 2) also informs the teacher's learning in such situations. Intellectual support and challenge are essential when working with students, but perhaps the most important area is development within the personal dimension. Achieving the right support–independence–interpersonal balance in an atmosphere of mutual respect and shared responsibility is a necessary ingredient of good (and enjoyable) professional development.

Less conventional methods

There is clearly a wide range of other ways of learning about the effectiveness of our teaching – even if using 53 of them (Gibbs et al., 1993b) is asking a

lot of teachers! Teachers should feel free to explore (with colleagues and students) other, less conventional approaches to evaluation that may address specific issues important to their course or enhance learning in an innovative way. Indeed, they may regard it as part of their professional role.

One example of such an unconventional method might be using role-plays of various kinds. In this category of method the student and/or teacher steps outside their normal roles and assumes other roles, and even 'persona', to examine complex issues of the course, or even to instigate development of it. Participants in role-play are permitted to express things which are very difficult to express when they are constrained within their normal roles. Students (and/or teachers) may take on the role of, for example, a 'traditionalist' or 'radical' teacher (and/or student), interpreting that in ways which can be instructive to developing a shared understanding of what the course is about and the different ways it is being perceived. As with most role-play, debriefing about the experience afterwards is particularly important, not only to look at the reality behind the roles but also to reconcile and come to terms with some of the things expressed, some of which may have been more emotional than expected.

The widespread use of communication and information technology also provides opportunities to engage in more unusual approaches to evaluation. Establishing an online course 'chat room' or discussion room for students to exchange their views on the course can be a useful way of eliciting information on a course. This may be left entirely up to the students to operate with the teacher simply 'eavesdropping' in on the conversation, or it may be one in which the teacher plays an active role in the discussion.

Similarly, the rules of the discussion may be very open to 'whatever happens' or they might be set up more formally with guidelines and specific themes preset by the teacher and/or negotiated with the students. It needs to be handled with care and the limitations of IT-based systems taken into consideration (see Chapter 7), but can be an instructive and rich supplement to more traditional methods.

Collecting concrete critical incidents of good and bad experiences is another less conventional method which can make evaluation more interesting and help to reveal the unexpected. Such methods generally may meet with traditional resistance to change and there can be difficulties in interpretation as well as finding the time actually to make them happen. On the other hand, developed and used creatively, they can be very useful for addressing issues that are unseen and/or ignored by more customary forms of evaluation.

EVALUATION OF ACADEMIC OUTCOMES AND CHANGE

Ultimately the success or failure of a course depends on whether or not students change in the way desired by them or their teachers. In addition to the direct measures of student learning and their work (discussed above and in Chapter 8), there are several indirect measures which are frequently collected and used by institutions for evaluating their impact on students. These indirect measures include:

- alumni, employer, student surveys;
- exit interviews of graduates;
- focus groups;
- graduate follow-up studies;
- retention and transfer studies;
- length of time to degree;
- standardized achievement scores;
- graduation rates;
- transfer rates; and
- job placements.

Of course, the relationship between the type of change suggested in these measures and the quality of teaching is highly problematic. Notoriously, large and small-scale studies of teaching methods and teaching resources often fail to show the relationship to academic performance as measured by traditional measures, let alone indirect measures. This is likely mainly due to a restricted nature of the assessment systems, but compensation for inadequate teaching must be another important factor – the course succeeds in spite of the teaching rather than because of it.

An added difficulty in most normal teaching situations is that there is very little pre-testing so that assumptions have to be made in order to credit students with actual gains in academic achievement. In general, it would seem that pass rates, graduation rates, job placements and academic standards are relevant to the evaluation of teaching but the relationship is a difficult one to interpret. At present they are mainly useful as warning signals when there are large fluctuations over time or between similar courses.

An important limitation on their use that deserves more attention is the infrequency of follow-up assessment of past students. When it is done it is usually very restricted. Ideally, it should address:

- how what was learnt on the course was or is being used;
- how efficiently it is being used; and
- whether it was relevant and effective.

While such an approach is not essential for all types of courses and pro-
grammes, it would be a strange course that did not expect to have some
positive intellectual, personal, social or practical impact on the future life
of the student.

We suggested earlier that problem-based learning, for example, may not
lead to a greater improvement in the student's actual knowledge or compe-
tence in specific skills and techniques than traditional courses, but it does
seem to have a significant effect upon their attitudes and their desire to go
on learning. Despite a long history in psychology of attitude measurement,
it is still not common within academic courses, despite the fact that
employers are increasingly interested in many of the attitudes towards
learning and lifelong learning which students will bring to their future
work. In the long run, these may be more important than knowledge and
the development of specific skills. While maintaining a degree of scepti-
cism about the accuracy of our attempts to measure these attitudes, the
actual attempt to do so may, nevertheless, be a useful way more fully to
understand the sort of attitudes we are attempting to encourage.

It is important not to regard the assessment of attitude change as com-
pletely separate from academic assessment. Less highly formalized methods
of student assessment – such as diaries and portfolios (see Chapter 8) – can
reveal a great deal about the more emotional and personal changes in
students' attitudes brought about by their educational experience. They
have an important role to play in enabling students to become more aware
of the relationship between their education and their developing sense of
identity, and enabling teachers to become clearer about the way in which
teaching relates to the personal and social concerns of their students.

CONCLUSIONS

Despite social pressures transforming accountability into a system of
accounting that favours quantitative evaluation methods over qualitative
ones, there is a clear need for a wide variety of complementary methods. No
single method is likely to have the necessary range or depth for evaluating
the complex processes and outcomes of university teaching. The emphasis
on institutional audits, and programme review plus the explosion of 'league
tables' and ranking systems both national and international (Liu and

Cheng, 2005; Usher and Savino, 2006), has raised the profile of evaluation in recent years, but it is also an essential feature in the appointment and promotion of individual staff. It is essential that its processes and methods live up to high standards of academic work expected elsewhere in the academy and that they contribute to the maintenance of the academic values which place learning and knowledge at the centre of higher education. Critically reviewing our evaluation methods with respect to the extensive and complex wealth of that learning is at the heart of both future educational development and the practices of the reflective professional.

An essential part of that review, however, will take the reflective professional beyond simply an examination of which methods most comprehensively address which aspects of student learning. It will (as much of the above discussion suggests) transform the conception of evaluation (found in most of the literature) from its focus on assessing teachers and teaching quality in terms of student learning to an engaged process which itself facilitates student learning.

The most significant developments in the evaluation of teaching will come not from teachers thinking about their own courses as delivering quality or from students as consumers expressing their judgements about the quality of the courses provided for them, but by an integration of evaluation into the learning process. In this conception, evaluation is, itself, an important part of a student's learning and self-knowledge, helping them to explore the strengths, weaknesses, inhibitions and styles of their thinking and working in relation to the constraints and opportunities of the course. Just as the assessment of students' academic attainment has become increasingly integrated into actual learning activities, so the evaluation of teaching may develop away from retrospective and external judgements towards the constant reflection upon the significance of the educational experience and the transition to becoming genuine reflective professionals.

Final questions: Dressel's (1976) comment at the conclusion of the previous chapter that 'only when the students become competent evaluators of their own goals, experiences and accomplishments do they become truly educated' is equally applicable to teachers. Indeed, the two are mutually interdependent. This raises questions for the reflective professional. What does this mean in practice? How can I integrate evaluation into the learning process? What methods of evaluation work best for my teaching context? Will the methods I use allow me to judge clearly my teaching and its impact on student learning? And finally, and most simply of all, how can I learn from my teaching and seek to improve?